

## **PGA DIVISION COMMITTEE – UNIT REVIEW BOARD ON RESEARCH DATA AND INFORMATION**

### **SECTION I: BOARD PROGRAM OVERVIEW**

#### **A) Mission Statement**

The mission of the Board on Research Data and Information is to improve the stewardship, policy, and use of digital data and information for science and the broader society.

#### **B) History and Evolution**

BRDI was founded almost 3 years ago following discussions with the internal NSF data committee and the OSTP Interagency Working Group on Digital Data, which was formed about a year earlier. The Board grew out of a long-standing Academy committee, the U.S. National Committee for CODATA (the Committee on Data for Science and Technology). CODATA is an interdisciplinary committee of the International Council for Science (ICSU). The U.S. CODATA was founded in the 1970s and has been directed since 1992 by Paul Uhlir, the current director of BRDI.

Because of the growing importance of digital data in the research process, the sponsor of the U.S. CODATA at NSF, Lou Brown (GEO, now retired) believed it was important to raise the level and visibility of the U.S. CODATA activity at the Academy to board status. Moreover, the CODATA committee was focused primarily on international activities and there was a general consensus by sponsors, committee members, and staff that greater and more sustained attention to the various research data issues on an interdisciplinary basis at the Academy would be beneficial. Finally, the formation of the IWGDD at the OSTP signaled a higher level of interest among the federal science agencies and the need for a counterpart structure at the Academy.

Initial funding was made available in September 2008 and the Board held its first meeting in January 2009. Because of a sudden staff discontinuity at NSF in early 2009 and a hiatus of funding from the NSF for two fiscal years, the Board's major core funding was not restored until March 2011. The low level of funding as well as the uncertainty during this period made it difficult to achieve substantial progress during the first two years and it was not until this year that a vigorous level of effort was possible to be undertaken.

The current funding stability and future prospects for core activity and ad hoc project funding support has greatly strengthened the Board's position and has led to a recent ramping up of work, both existing and planned. Unfortunately, the first group of appointed board members did not share in the sense of accomplishment that the current activities and plans promise to deliver.

#### **C) Accomplishments**

We divide our summary of accomplishments according to national and international activities, since most of our projects address multiple substantive areas in our mission statement and are largely interdisciplinary. Details of the Board's individual projects are provided in Sections III.A and D.

## ***BRDI Projects with a National Focus***

Because the predecessor of the new Board, the U.S. CODATA, was focused primarily internationally, one of the principal motivations for creating BRDI was to provide a greater focus on national issues. We have used several different mechanisms to examine, advise, and network with other organizations at the national level.

Perhaps the most significant and long-lasting accomplishment of BRDI is to provide a new, high-level forum at the Academies for discussion of emerging issues in the field of scientific data and information that affect U.S. governmental science and informatics agencies, as well as the U.S. (and the international) research communities. We have done so using a variety of approaches. These include instituting a series of in-depth, ad hoc symposia and workshops on emerging issues of concern to the sponsors and the broader community (on automated knowledge discovery, data attribution and citation practices and standards, and intellectual property issues in databases), consensus studies (we are about to begin a new study on the future career opportunities and educational requirements for digital curation), and short public symposia on various themes of broader public interest. The latter topics include: author deposit mandates for federal grantees, scientific data for evidence based policy and decision making, the changing role of libraries in support of research data activities, the value of shared access and reuse of publicly funded scientific data, and crowdsourcing: improving quality of scientific data through social networking).

The Board also has implemented a strong outreach and communication strategy, targeted at different constituencies. For the sponsors, we provide substantial time at the Board meetings for briefings and interactions (both individual sponsor liaisons and representatives of interagency groups), we are starting a new BRDI Sponsor Forum (the first one will be this fall on the data management mandate plans required for all science agencies under the America COMPETES Act), and we meet individually with the sponsors once or twice per year for focused briefings. With regard to the broader community, we have implemented a sustained communication and outreach effort with sponsoring organizations, academic and NGO entities, and the general public. For example, we invite the directors of three academic or non-governmental organizations working in the scientific data and information policy and management arena, to speak with the Board about issues and activities of common concern, we invite them to the other Board symposia, and disseminate a semi-annual *BRDI Told Me* Newsletter to several hundred (free) subscribers. We also have begun implementing a Young Scientists Program to involve younger graduate students and post docs in the Board's work, which as far as we know is a unique initiative at the NRC.

## ***BRDI's International Projects***

The Board also acts as the U.S. CODATA representative, which cooperates with ICSU's interdisciplinary CODATA and plans the various international activities. We have a small U.S. CIODATA Executive Committee, which is composed of a small subset of BRDI members. The BRDI/US CODATA international activities have included international workshops and symposia on the socioeconomic effects of public sector information on digital networks (2009, with OECD), designing the microbial research commons (2009, with the Board on Life Sciences), a symposium on the data sharing plans and scientific benefits of data sharing through GEOSS (2009, with the Board on International Scientific Organizations—BISO), and the case for international sharing of scientific data—a focus on developing countries (2011, with BISO and ICSU).

The U.S. CODATA began a successful series of informal meetings in 2000 with the Chinese Academy of Sciences and the Ministry of Science and Technology, which led to the development of MOST's Scientific Data Sharing Program in 2003 and a major workshop in Beijing in 2004. These activities led to the organization of a more formal bilateral Roundtable on Scientific Data Cooperation in 2006, which BRDI has continued. The Roundtable holds annual meetings in which academic and governmental institutions from both countries participate in discussions focused on thematic topics of mutual interest, and these meetings have resulted in a bilateral USGS-CAS MoU and some data exchanges.

The Board/US CODATA also is very active in international CODATA activities, including the biennial conferences and General Assemblies, and the Board recently initiated a U.S. Forum on CODATA-World Data System Cooperation, which will hold its first meeting in conjunction with the next BRDI meeting.

Finally, the Board director has complemented the Board's work with many related articles, invited presentations, and chairing or participating on advisory boards and committees, mostly at the international level. He was awarded the international CODATA Prize in 2010 and the NRC Special Achievement Award in 1997 for his work in this field.

#### **D) Funding**

As noted at the outset, the Board's funding situation did not become stable until March 2011, when the National Science Foundation provided core funding through a 3-year grant in a lump sum of \$1,050,000. The NSF first gave start-up core funds of \$275K in September 2008, which were used in the first year on the expectation that funding would continue in 2009 and 2010 at \$350K/year.

In addition to the National Science Foundation (\$350K/year), BRDI's core activities are currently funded by the National Institutes of Health (NLM, NIBIB, and NIEHS at a total of \$50K/year), the National Institute for Science and Technology (\$75K/year), the Institute of Museum and Library Services (\$60K/year), the National Oceanic and Atmospheric Administration (\$35K/year), and the United States Geological Survey (\$35K/year). The funding of the latter three agencies came as a lump sum for a 3-year period, but was used up in the first year of funding because of the budgetary shortfall caused by the almost 2-year hiatus in NSF funding, as explained above. The Library of Congress has provided resources in kind by providing a program officer on detail for the last three years. Finally, the Defense Technical Information Center (DTIC) also provided core funding in the 2008-2010 period (\$50K/year), which was terminated due to a substantial budget decrease for DTIC in FY 2010.

The Board has also received ad hoc project support, but usually at a lower amount than other similar activities at the NRC, primarily because of the full-time program officer through the IPA arrangement from the Library of Congress. This has allowed us to perform projects at a reduced amount, since perhaps one-half of the labor costs of some projects have been provided through this mechanism. The Board's core funds, which are generally much higher than for other NRC boards (except in the 2009-2010 period), have made up some of the other shortfall. More specifically, the following is a summary of ad hoc project funding, in chronological order from the oldest to most recent:

*US-China Roundtable on Scientific Data Cooperation.* This activity was initiated in 2006 through the U.S. CODATA, prior to the formation of the Board, and is terminating in 2011. The

NIH National Institute of Biomedical Imaging and Bioengineering has provided funding at the amount of \$75K/year, and a small supplementary request from NSF for the U.S. CODATA core support of \$30K in 2007 also was used for this activity. The project is now ending because of modest results and an unfavorable budgetary climate for bilateral research projects with the Chinese (described further in sections III.A and C, below).

*International Symposium on Designing the Microbial Research Commons.* The symposium itself was held in October 2009. Funding for the preparation and travel of this meeting came largely from the initial NSF grant in 2008 and a small Department of Energy grant of \$20K, also for travel costs. The subsequent editing and report production work, which was delayed by the budget uncertainties in 2010, was provided by the Library of Congress IPA and the Board's core funding in 2011 from the NSF.

*Symposium and Workshop on the Future of Scientific Knowledge Discovery in Open Network Environments.* This meeting in March 2011 was funded in September 2010 by NSF at \$100K. The remaining amount has been covered by the Library of Congress IPA and some core funds.

*International Symposium on the Case for International Sharing of Scientific Data: A Focus on Developing Countries.* This was a joint project with the Board on International Scientific Organizations, which contributed \$150K in direct costs in the 2010-2011 period. The meeting was held in April 2011. Most of the project's labor expenses have been covered by BRDI's core funds.

*Developing Data Attribution and Citation Practices and Standards.* The Sloan Foundation provided the bulk of the funding (\$281K) for this 2-year project. The IMLS contributed \$50K, and the international CODATA \$10K and Microsoft Research \$5K.

*Consensus Study on Future Career and Educational Requirements for Digital Curation.* This study is beginning in September 2011. The NSF has provided funding in the amount of \$100K and the IMLS will provide \$300K this month, with the remaining \$155K still forthcoming.

#### **E) Funding strategies, including approximate anticipated/desired sponsor base and levels**

The anticipated funding levels over the next three years are based on a mix of core funding, currently projected to be approximately \$700K/year; one IPA (~\$300K/year), ad hoc project support of ~\$400K/year, and in-kind contributions from collaborating entities of perhaps ~\$100K for a total of about \$1.5M.

The Board's current level of effort is based on 3.5 professional staff (director, senior program officer at 50%, program officer on IPA, and senior program associate). Funding for support staff at about 25% is anticipated with the consensus study now being initiated and this is expected to increase to 50% next year, assuming that the existing level of core support and new project funding will be forthcoming, for a total number of 4 full-time staff. We have used and will continue to employ, some off-site consultants and editors for discrete tasks that take several weeks to complete and cost in the \$5K-10K range.

This level of funding and staff will allow us to undertake a project profile per year of:

- Two Board meetings and related activities (strategic planning, sponsor briefings and fund-raising, outreach and communication, monitoring of substantive issues and

developments, various coordination activities and staff meetings, and professional staff development);

- Two public mini-symposia with a Meeting ReCap, in conjunction with the Board meetings;
- Two half-day meetings of the U.S. Forum on CODATA-World Data System Cooperation, in conjunction with the Board meetings;
- One consensus study, with report;
- One or two 2-day symposia/workshops, with report;
- One sponsor forum;
- Planning and participation in the international CODATA Conference and General Assembly every other year, and related CODATA activities in the interim; and
- Various small ad hoc meetings with external visitors and delegations.

Because of its interdisciplinary scope, the funding strategy is to develop and maintain a core set of multiple federal agency sponsors, and ad hoc project support from a mix of government agencies, foundations, and a small amount from corporations and the international CODATA. The current funders are all expected to continue, and DOE has indicated an interest to contribute \$100K/year. It is realistic to expect 1-2 other agencies to become core sponsors, while 1-2 existing sponsors may either reduce their funding or drop out altogether, due to the tight federal budgets. The same or other federal agencies are very likely to support the ad hoc project activities, and the director and senior staff have very good relations with all existing sponsors and some additional ones.

Among private foundations, there is strong support from the Sloan Foundation. Other private sources that appear likely to support future ad hoc projects and with which we have established various relations include: the Ford Foundation, the Mellon Foundation, the MacArthur Foundation, the Google Foundation, and Microsoft Research. Other foundations and corporations are a distinct possible source of support.

Over the past three years and before with the U.S CODATA, the director has endeavored to form collaborative relationships with other organizations, both within and outside the NRC. Joint projects with other NRC entities and some resources have come from: the Committee on Science, Engineering, and Public Policy (COSEPUP), the Computer Science and Telecommunications Board (CSTB), the Board on International Scientific Organizations (BISO), the Board on Life Sciences (BLS), and the Board on Earth Science and Resources (BESR). External collaborations and in-kind contributions have come from CODATA, OECD, and UNESCO, as well as national scientific institutions in China, Brazil, South Africa, and Japan.

This collaboration has included young scientists as well, with several Fellows over the years from the Christine Mirzayan Fellowship. However, both the Board and the U.S. CODATA before it have used PhD candidates and post-docs as part-time paid or volunteer consultants (see Section III.A below for more information).

## **F) Strengths and Areas for Growth**

The interdisciplinary and multi-sectoral nature of the Board's focus is both a strength and a weakness. It is a strength in that it can view topics holistically and not just from a single discipline perspective. This is important for the consideration of data and information policy and management of digital content issues and problems, since many are relevant across science and require an integrated focus to address them successfully. Examples of this are the data attribution

and citation project or the study of future digital curation workforce and education requirements, although there are many other such topics (see Sections III.A and D). It thus fills a niche, especially at the Academies, that would otherwise be served only narrowly or on ad hoc basis, if at all.

The Board's collaborative approach is essential for getting domain expertise and reaching communities for diffusing advice and solutions. Serving as a focal point or forum for discussing these issues across the federal agency informatics community, and bringing in the academic, NGO, and to some extent the private sector communities that work in this area is another important contribution. BRDI is also in a unique position to provide a forum and leadership to the international scientific communities in seeking solutions to problems that are now increasing with global data sharing.

The weaknesses stem in part from the same attributes. Its interdisciplinary nature, while broad in scope, is necessarily less deep in detail, unlike a focus in a certain discipline or sector. This can make “ownership” of the issues difficult, both from the perspective of providing advice as well as from attracting support and participation. As with some other NRC activities, funding for BRDI is not linked to any particular “essential” programs or projects of the agencies; it is discretionary spending and thus more vulnerable to reduction or elimination in the current fiscal climate. The Board's struggle in getting adequate funding in its early years is partially a reflection of that, although this situation appears to have been stabilized this year and for the near-term. An inordinate amount of time by the staff was devoted to the raising of funding, with attendant ill effects from reduced productivity and stress.

## **G) Goals**

BRDI's statement of task reflects goals similar to other NRC boards:

1. Address emerging issues in the management, policy, and use of research data and information at the national and international levels.
2. Through studies and reports of the NRC, provide independent and objective advice, reviews of programs, and assessment of priorities concerning research data and information activities and interests of its sponsors.
3. Encourage and facilitate collaboration across disciplines, sectors, and nations with regard to common interests in research data and information activities.
4. Monitor, assess, and contribute to the development of U.S. government and research community positions on research data and information programs and policies.
5. Initiate or respond to requests for consensus studies, workshops, conferences, and other activities within the Board's mission, and provide oversight for the activities performed under the Board's auspices.
6. Broadly disseminate and communicate the results of the Board's activities to its stakeholders and to the general public.

The substantive scope and goals of the Board's mission statement are reflected in task #1.

## **H) Current goals and planned changes in goals**

We believe that the current goals are consistent with the goals set forth in the statement of task. As with other NRC activities, the focus of the work fluctuates with the interests of the sponsors and Board members, the amount of funding available, and other external factors. There are no immediate plans to change the Board's mission or task statement, especially since the group was formed a scant three years ago.

## **I) Strategies to meet those goals**

The Board members, sponsors, and staff devote a lot of time to discussing the mix of activities, both at the semi-annual meetings and in-between. There is a list of topics and types of activities (mini-symposia, major symposium and workshops, consensus studies, forums) that the Board develops, guided by its mission and task statement and discussions with the sponsors. That list is then fleshed out and further discussed with sponsors to see which ones are of interest to them.

A list of metrics from the Board's strategic plan is presented in section III.C.

## **J) Ideas for how the PGA Committee might be able to help the unit fulfill its goals**

The Board members and staff believe that external review is a valuable process that stimulates useful introspection and objective external review, both of which should be constructive and lead to improvements. The Board is interested in particular to get the PGA Committee's perspective and advice of both substantive projects and procedural matters, experts to involve, and sources of support.

## SECTION II: FUNDING AND SPONSOR INFORMATION

Table 1: TOTAL INDIRECT AND DIRECT PROGRAM ACTIVITIES

### Board on Research Data and Information

*By Expense Type*  
(\$ in thousands)

<b>Expense Type</b>	2007 (Actual)	2008 (Actual)	2009 (Actual)	2010 (Actual)	2011 (Proposed)
Regular	179	296	616	584	729
Flow-Through	66	71	74	69	72
<b>Total</b>	<b>\$244</b>	<b>\$367</b>	<b>\$690</b>	<b>\$653</b>	<b>\$802</b>

Note: all fiscal year data represents the time period of January 1st to December 31st.

Table 2: TOTAL INDIRECT AND DIRECT PROGRAM ACTIVITIES

### Board on Research Data and Information

*By Account Type*  
(\$ in thousands)

<b>Account Type</b>	2007 (Actual)	2008 (Actual)	2009 (Actual)	2010 (Actual)	2011 (Proposed)
<b>Regular Program Activity</b>					
On-Site Labor	63	73	163	157	160
Off-Site Labor	3	35	81	98	114
Consulting Fees	--	--	9	--	9
Overhead	46	66	147	129	137
Travel	24	58	57	63	122
Other Operating Expenses	14	13	56	42	60
General and Administrative Costs	<u>29</u>	<u>51</u>	<u>102</u>	<u>95</u>	<u>127</u>
	179	296	616	584	729
<b>Flow-Thru Activity</b>					
Grants and Fellowships	63	68	70	66	70
General and Administrative Costs	<u>3</u>	<u>3</u>	<u>4</u>	<u>2</u>	<u>2</u>
	66	71	74	69	72
<b>Total</b>	<b>\$244</b>	<b>\$367</b>	<b>\$690</b>	<b>\$653</b>	<b>\$802</b>

Table 3: TOTAL INDIRECT AND DIRECT PROGRAM ACTIVITIES

**Board on Research Data and Information***By Funding Source*

(\$ in thousands)

<b>Funding Source</b>	2007 (Actual)	2008 (Actual)	2009 (Actual)	2010 (Actual)
<b>Federal Government</b>				
Department of Commerce	--	--	68	191
Department of Energy	--	--	16	4
Department of Health and Human Services	94	68	167	160
Department of the Interior	--	--	--	4
National Science Foundation	148	296	249	153
Other Federal	<u>--</u>	<u>--</u>	<u>165</u>	<u>115</u>
Subtotal Federal	\$242	\$364	\$665	\$627
<b>NRC Indirect Cost Pool</b>	\$2	\$3	\$25	\$26
<b>Total</b>	<b>\$244</b>	<b>\$367</b>	<b>\$690</b>	<b>\$653</b>

### **SECTION III: PROJECT-LEVEL INFORMATION**

The Board on Research Data and Information has undertaken or initiated a portfolio of continuing programs and ad hoc projects pursuant to its mission and task statement. The summary descriptions of past and current activities that have been either initiated or completed in the past three years are presented in chronological order in the next two sections. Section C details some of their impacts for the various Board constituencies at both the national and international levels. The final section in this part describes the programs and projects that are about to be initiated or have begun to be planned by the Board members, staff, and sponsors.

#### **A) Program/Project Details**

##### **1) *Meetings of the Board on Research Data and Information***

The meetings of the Board offer the members, sponsors, and staff to plan the various programs and projects, review the current or ongoing activities, and provide a forum for discussing the various issues and developments relevant to the Board's purview. The meetings all have included a half-day session for the planning and review of U.S. CODATA activities vis-à-vis the international CODATA; briefings by or discussions with the Board's sponsors and high-level interagency groups involved in scientific data and information programs and coordination (IWGDD, CENDI, and U.S. GEO); an outreach session focused on coordination with several external groups with similar activities and interests; a short topical symposium (discussed separately below) of more general public interest; and a closed planning and review session for the Board members.

The agendas and presentation slides for all five of the BRDI/US CODATA meetings are located at the following links:

First Meeting: January 29-30, 2009 [http://sites.nationalacademies.org/PGA/brdi/PGA\\_047585](http://sites.nationalacademies.org/PGA/brdi/PGA_047585)

Second Meeting: September 24-25, 2009  
[http://sites.nationalacademies.org/PGA/brdi/PGA\\_052793](http://sites.nationalacademies.org/PGA/brdi/PGA_052793)

Third Meeting: June 3-4, 2010 [http://sites.nationalacademies.org/pga/brdi/PGA\\_056785](http://sites.nationalacademies.org/pga/brdi/PGA_056785)

Fourth Meeting: November 30-December 1, 2010  
[http://sites.nationalacademies.org/PGA/brdi/PGA\\_059159](http://sites.nationalacademies.org/PGA/brdi/PGA_059159)

Fifth Meeting: June 13-14, 2011 [http://sites.nationalacademies.org/PGA/brdi/PGA\\_062657](http://sites.nationalacademies.org/PGA/brdi/PGA_062657)

##### **2) *CODATA International Conferences and General Assemblies***

The international Committee on Data for Science and Technology (CODATA, see [www.codata.org](http://www.codata.org)), is an interdisciplinary committee of the [International Council for Science \(ICSU\)](http://www.icsu.org), concerned with various types of quantitative data resulting from experimental measurements or observations in the natural and social sciences, and the engineering disciplines. Particular emphasis is given to data management problems common to different scientific fields and to data sharing among these disciplines. The U.S. National Committee for CODATA (the

U.S. CODATA) functions as a bridge between the scientific and technical community in the United States and the international CODATA on data issues addressed in the natural and social sciences. The U.S. CODATA is represented by BRDI.

CODATA held its 21<sup>st</sup> International Conference and General Assembly on 5-10 October 2008, in Kyiv, Ukraine, and the 22<sup>nd</sup> meeting on 23-27 October 2010. Several USNC/CODATA members and staff participated in organizing the program and individual sessions, and gave presentations there. The U.S. National Delegate and other U.S. participants in the CODATA leadership are very influential in setting the organizations agenda and leading the activities. The results of both the Conference and the General Assembly for both meetings, respectively, are available at <http://www.codata.org/08conf/index.html> and at <http://www.codata.org/archives/2010/index.html>.

**3) *Public Symposium on Author Deposit Mandates for Federal Research Grantees, 29 January 2009***

In 2007, Congress enacted legislation requiring all grantees of the National Institutes of Health to deposit their journal articles in the National Library of Medicine's PubMedCentral repository, in order to provide open availability online to the government-funded biomedical literature for the general public. There is a strong possibility of extending this legislation to other federal agencies. Because of the importance of access to the results of publicly funded research, the Board organized a "mini-symposium" to examine the pros and cons of such a policy for other types of federally funded research.

The symposium was chaired by the BRDI Chair, Michael Lesk, and the representatives of the following organizations made presentations and participated in the discussion: American Physiological Society, American Institute of Physics, American Psychological Association, Society for Science and the Public, and publisher of *Science News*, Scholarly Publishing and Academic Resources Coalition (SPARC), and National Association of State Universities and Land-Grant Colleges (NASULGC). The symposium was audiocast and made available on: [http://sites.nationalacademies.org/pga/brdi/PGA\\_047286](http://sites.nationalacademies.org/pga/brdi/PGA_047286).

**4) *The Socioeconomic Effects on Public Sector Information on Digital Networks – Towards a Better Understanding of Different Access and Reuse Policies, 4-5 February 2008, Paris, France, and Report, June 2009***

This workshop was organized prior to the formation of BRDI by the U.S. National Committee for CODATA, the NRC Board on International Scientific Organizations (BISO), and the Organisation for Economic Co-operation and Development (OECD). However, the project was completed under the auspices of BRDI. The workshop reviewed the state-of-the-art in assessment methods and to improve the understanding of what is known and what needs to be known about the effects of Public Sector Information (PSI) activities. The workshop focused on different U.S. and foreign approaches for evaluating the direct and indirect economic and non-economic benefits and costs of PSI access and reuse policies in the online environment, and on measuring the economic and social costs and benefits of the PSI. A full report of the workshop, which was published by BRDI/U.S. CODATA in June 2009, is available online at: [http://sites.nationalacademies.org/PGA/brdi/PGA\\_047287](http://sites.nationalacademies.org/PGA/brdi/PGA_047287).

## **5) The U.S.-China Roundtable on Scientific Data Cooperation**

The U.S. National Committee for CODATA and the Chinese National Committee for CODATA, under the auspices of their respective Academies of Sciences, organized a U.S. - China Roundtable on Scientific Data Cooperation in 2006, which has convened a series of annual meetings pursuant to the following Statement of Task:

- a) Provide a unique bilateral forum for government, academic, and private-sector stakeholders in the United States and China to discuss and address scientific data practices and policies, pursuant to a mutually agreed agenda.
- b) Serve as a catalyst and coordinating body for bilateral cooperation on scientific data practices and policies at the Academy and national level in each country, with appropriate recognition and representation of other thematically related bilateral and international activities.

The four areas identified for framing the scope of discussion include: scientific data policy, advanced cyber-infrastructure data applications, health and biomedical data, and environmental and geospatial data. Two of the meetings were held prior to the initiation of BRDI, but the program has been performed under the auspices of the Board, beginning with the third meeting. Details about the agendas and presentations for the third and fourth meetings are available on the BRDI website at [http://sites.nationalacademies.org/PGA/brdi/PGA\\_050409](http://sites.nationalacademies.org/PGA/brdi/PGA_050409) and [http://sites.nationalacademies.org/PGA/brdi/PGA\\_056119](http://sites.nationalacademies.org/PGA/brdi/PGA_056119), respectively.

A fifth meeting of this Roundtable is scheduled for October 27-28, 2011 in Beijing, China.

## **6) Symposium on Common Use Licensing of Publicly Funded Scientific Data and Information, 25 March 2009, Beijing, China, and 27 March 2009, Taipei, Taiwan**

These meetings were separate from the U.S.- China Roundtable activities described above. The U.S. National Committee for CODATA was the co-organizing U.S. partner for both symposia. In Beijing, the cooperating partners were the National Science Library of the Chinese Academy of Sciences and the Creative Commons Office of China. In Taiwan, they were the Academia Sinica and the Creative Commons Office of Taiwan.

This one-day symposium at each venue reviewed the rationale, practice, and issues associated with the application of Creative Commons/Science Commons "common use" licenses to scientific literature and data in government and academia, and explored the possible implementation of such licenses to publicly funded scientific literature and data. It was intended to provide an introduction to a high-level government and academic audience, and to develop some potential follow-up activities in this area. Details about both events can be found at [http://sites.nationalacademies.org/PGA/brdi/PGA\\_054916](http://sites.nationalacademies.org/PGA/brdi/PGA_054916) and [http://sites.nationalacademies.org/PGA/brdi/PGA\\_051473](http://sites.nationalacademies.org/PGA/brdi/PGA_051473), respectively.

## **7) Public Symposium on Scientific Data for Evidence Based Policy and Decision Making, 24 September 2009**

One of the stated priorities of the Obama Administration is to promote transparency and openness in governance, including in the dissemination and use of scientific data and information, and the use of factual scientific information in the formation of government policy and official decisions. BRDI held a public symposium on "Scientific Data for Evidence Based Policy and Decision Making", which provided a more detailed understanding of the policies and programs of the current Administration that support its stated policy of basing its decisions on sound science. The symposium featured presentations by high-level Administration officials from OMB, OSTP, FDA, and EPA on the use of scientific data in federal regulations and policymaking, scientific information policy for promoting better decision making and innovation, new mechanisms for public access to federal data, and examples of high-value applications of scientific data for drug approval and for environmental policymaking. These presentations were followed by a panel discussion of the invited speakers and several Board members, and provided an opportunity for interaction with the audience. Additional background, the agenda, speaker bios, and presentations are available at [http://sites.nationalacademies.org/PGA/brdi/PGA\\_052920](http://sites.nationalacademies.org/PGA/brdi/PGA_052920).

**8) *International Symposium on Designing the Microbial Research Commons, 8-9 October 2009, Washington, DC***

The opportunities to accelerate scientific discovery and resulting applications are made increasingly possible by technological breakthroughs and pioneering methods to process and integrate vast amounts of data, information, and raw materials. Microbial research, which is outgrowing its "small science" institutional structures, needs to build upon these opportunities in an attempt to develop a global microbial research commons to promote access to databases, literature, and materials through an open, digitally distributed network. However there are various "soft infrastructure" barriers that need to be addressed and overcome in order to build global microbial research commons. An essential, early step is the development of a set of design principles that address the economic, legal, and institutional dimensions of the transformation of the existing research infrastructure into what could become a globally distributed and digitally integrated research commons. The goal of this redesigned soft infrastructure would be to better manage publicly funded research resources, without compromising downstream commercial applications and fruitful partnerships between the public and private sectors, or between developed and developing countries.

In light of these considerations, BRDI organized an *International Symposium on Designing the Microbial Research Commons*, This Symposium expanded on prior international discussions on the same topic at a conference in June 2008 in Ghent, Belgium (see: <http://www.microbialcommons.ugent.be/> ). The October 2009 International Symposium addressed topics such as models to lower the transaction costs and support access to and use of microbiological materials and digital resources from the perspective of publicly funded research, public-private interactions, and developing country concerns. The symposium report is openly available at [www.nas.edu/brdi](http://www.nas.edu/brdi).

**9) *Symposium on Data Sharing Plans for GEOSS and the Benefits of Data Sharing for Science, 16 November 2009 (in conjunction with the Group on Earth Observations Plenary Meeting in Washington, DC)***

The Global Earth Observation System of Systems (GEOSS) is a major international initiative under the Group on Earth Observations (GEO): see <http://www.earthobservations.org/>. The program is being developed with existing observational systems and incorporating new systems

for Earth observation and modeling that are offered as GEOSS data components by the GEO Member Countries and Participating Organizations. A major challenge is the coordination and harmonization of data policies and procedures to facilitate the sharing and use of GEOSS data to maximize societal benefits for the widest possible range of users. The International Council for Science (ICSU), working through the international CODATA, developed a 100-page *White Paper and Implementation Guidelines for the GEOSS Data Sharing Principles*, and is now finalizing *Legal Options for the Exchange of Data through the GEOSS Data-CORE*. The BRDI director was the lead author on both publications.

The U.S. CODATA, together with the international CODATA, the GEO Secretariat, and several other organizations held a side event regarding the GEOSS data sharing principles at the 2009 GEO Plenary Meeting in Washington, DC. Additional information about this event, including the agenda, speaker bios, and presentation slides is available on the BRDI website at: [http://sites.nationalacademies.org/PGA/brdi/PGA\\_053959](http://sites.nationalacademies.org/PGA/brdi/PGA_053959). A Meeting ReCap has been published and is available there as well.

#### **10) Meeting with Chinese Microbiology Research Delegation, 7 December 2009**

The BRDI staff hosted a delegation of eight senior microbiology researchers and officials from the Chinese Academy of Sciences and the Chinese Ministry of Science and Technology to discuss possible basic and applied research cooperation with U.S. representatives of the National Science Foundation, the National Academies, and the Biotechnology Industry Organization. The discussion laid some additional groundwork for the third meeting of U.S-China Roundtable on Scientific Data Cooperation on 29-30 March 2010. For the agenda and presentations at the meeting with the Chinese Delegation, see [http://sites.nationalacademies.org/PGA/brdi/pga\\_054999](http://sites.nationalacademies.org/PGA/brdi/pga_054999).

#### **11) Symposium on Managing the Exaflood: Enhancing the Value of Networked Data for Science and Society, 19 February 2010**

The total amount of data generated each year is likely to double every two years for at least the next decade as the costs of computing and networking continue to plunge, and the number of people and data generating instruments connected to the internet soars. George Gilder has referred to this phenomenon as the "exaflood." To realize the full benefit and value of these diverse and voluminous data requires more effective data management techniques, institutional arrangements, and policies.

In view of these pressing challenges, several BRDI members and staff organized a symposium at the AAAS Annual Meeting in San Diego, CA. The symposium brought together leading research and policy experts who addressed the value proposition of the exaflood, provided compelling examples of applications that expand the boundaries of what is possible, and discussed some of the policy and management issues that must be resolved. Additional information about this event may be found at: [http://sites.nationalacademies.org/PGA/brdi/PGA\\_053368](http://sites.nationalacademies.org/PGA/brdi/PGA_053368).

#### **12) The Changing Role of Libraries in Support of Research Data Activities, 3 June 2010, Washington DC**

The institutional roles of libraries in the management and support of scientific data activities for research and for the broader society are one important area of focus for the BRDI. Three of the Board's sponsors are major federal library organizations - the Library of Congress, the National

Library of Medicine, and the Institute of Museum and Library Services - and they each provide a different role and perspective concerning these issues.

The Board's public symposium on the afternoon of 3 June 2010 featured presentations on this set of topics by senior managers from these three federal library organizations, as well as from the Association of Research Libraries, which has examined the role of libraries for research data in some depth and whose members are institutions with a great deal at stake in this debate. The symposium concluded with comments by two Board members, a university professor and researcher working in the data-intensive field of geographic information and a university professor of information science. The symposium was moderated by Prof. Michael Lesk, the Board chair. To see the video webcast, presentation slides, and bios of the speakers, please visit [http://sites.nationalacademies.org/PGA/brdi/PGA\\_056901](http://sites.nationalacademies.org/PGA/brdi/PGA_056901).

***13) Public Symposium on “The Value of Shared Access and Reuse of Publicly Funded Scientific Data”, December 1, 2010***

BRDI held this public symposium to explore some of the research, economic, and social benefits that can be derived from providing online access to publicly-funded scientific data, as well as how such benefits can be evaluated, with a view to adding to that inquiry. The event was chaired by Michael Lesk, and included presentations by representatives of the Congressional Research Service, the NIH National Institute on Aging, U.S. Geological Survey, and the Université du Québec à Montréal & University of Southampton. There were also presentations by two members of the Board. The presentations were followed by a panel discussion. A Meeting ReCap will soon be available.

***14) The Future of Scientific Knowledge Discovery in Open Networked Environments: A National Symposium and Workshop, Washington, DC, March 10-11, 2011***

Scientific knowledge discovery in open networked environments, referred to here as computer-mediated or computational scientific knowledge discovery (SKD), may be defined as a research process that is enabled by different digital computing technologies such as data mining, information retrieval and extraction, artificial intelligence, distributed grid computing, and many other automated methods. Together, these technological capabilities are supporting the emergence of computer-mediated SKD as a new paradigm in the conduct of research.

A two-day symposium and workshop was organized by BRDI to bring together key stakeholders in this area for intensive and structured discussions. The objective was to obtain a better understanding of the state-of-the-art and future trends in the study of computational SKD in the open online environment and to develop a range of options for future work in this area. The Steering Committee for this project was chaired by Prof. John King of the University of Michigan. Presentations were made by experts from governmental agencies, private sector, academia and non-profit organizations. Additional background information, the meeting agenda, speaker bios, and presentation slides are available on the BRDI website at: [http://sites.nationalacademies.org/PGA/brdi/PGA\\_060422](http://sites.nationalacademies.org/PGA/brdi/PGA_060422).

***15) BISO-BRDI-CFRS International Symposium on the Case for International Scientific Data Sharing: A Focus on Developing Countries, April 18-19, 2011***

BRDI, in collaboration with the NRC Board on International Scientific Organizations (BISO), and the International Council of Science Committee on Freedom and Responsibility in the conduct of Science (CFRS), organized a 2-day international symposium in Washington, which discussed various issues that promote or hinder data sharing in various scientific fields and made a compelling case for the benefits to science and society of international data sharing. The symposium held five sessions, each one focused on an important aspect of data sharing. There were presentations by national and international experts from governmental organizations, academia, non-governmental organizations, and industry. A proceedings report is being prepared by BISO and BRDI, and will be published in the winter of 2012. The CFRS will release a policy statement on this topic soon. The symposium agenda, speaker bios, and presentation slides are available at: [http://sites.nationalacademies.org/PGA/biso/PGA\\_061353](http://sites.nationalacademies.org/PGA/biso/PGA_061353).

**16) *Public Symposium on “Crowdsourcing: Improving the Quality of Scientific Data Through Social Networking”, June 13, 2011***

Crowdsourcing may be described as a distributed information production and problem-solving activity, today performed mostly online. The technique invites contributions on one or more specific issues or problems, either from a targeted group or the general public. Although there are many types of crowdsourcing applications in many sectors and businesses, the public research community has used the technique extensively in recent years.

Because of the growing use and potential importance of this technique to various research applications, including the improvement of scientific information resources, BRDI held a public symposium in Washington DC to explore some of the key underlying issues. It was chaired by Michael Lesk, and had presentations by experts from governmental agencies, academia, and non-governmental organizations. The presentations were followed by a panel discussion. More background information, the symposium agenda, and links to the speaker bios and presentations are available at: [http://sites.nationalacademies.org/PGA/brdi/PGA\\_062938](http://sites.nationalacademies.org/PGA/brdi/PGA_062938).

**17) *Developing Data Attribution and Citation Practices and Standards: An International Symposium and Workshop, August 22-23, 2011, Berkeley, California***

US CODATA and BRDI in collaboration with the CODATA-ICSTI (International Council for Scientific and Technical Information) Task Group on Data Citation Standards and Practices, held a 2-day international symposium and workshop and discussed a) the major universal as well as field-specific or context-specific technical issues that need to be considered in developing and implementing scientific data citation standards and practices; b) the major universal as well as field-specific or context-specific institutional, financial, legal, and socio-cultural issues that need to be considered in developing and implementing scientific data citation standards and practices; and c) the status of data attribution and citation practices in individual fields in the natural and social (economic and political) sciences in United States and internationally. The meeting background, agenda, speaker bios, and presentation slides are available at: [http://sites.nationalacademies.org/PGA/brdi/PGA\\_063656](http://sites.nationalacademies.org/PGA/brdi/PGA_063656).

This meeting is part of a longer CODATA-ICSTI Task Group activity that will develop a white paper on best practices in scientific data citation in 2012-2013. The BRDI/U.S. CODATA staff is also the staff for the CODATA-ICSTI Task Group and received funding to support that continuing activity over a two-year period.

## **18) Young Scientist Program**

In 2010, the Board initiated a systematic mentoring approach to:

- 1- Encourage young researchers to participate in Academy activities in order to get first-hand experience in areas related to their studies and interests. This could be facilitated through one-to-one mentorship activities.
- 2- Expose young researchers to issues of national and global concern.
- 3- Help young researchers expand their professional and personal networks through interdisciplinary national and international collaborations.
- 4- Encourage and enable young researchers to start and sustain their own activities in subject areas seeded by the Board's focus.
- 5- Develop a network of researchers who are able and interested to represent the U.S. in different national and global *e*-science and digital policy areas.

The Board has sought to implement these goals by mentoring Christine Mirzayan Fellows, hiring young consultants periodically on a part-time basis, and appointing Ph.D. students and post-docs as "unpaid consultants" to participate in ad hoc project activities and serve as rapporteurs at workshops.

## **B) Activities completed over the past 3 years**

The Board has held over 20 meetings and has posted the agendas, presentation slides, and descriptive information openly on the [www.nas.edu/brdi](http://www.nas.edu/brdi) website in all cases. We also have published two symposium reports and 3 meeting ReCap reports, with several other now pending completion. There are more activities shown as completed than projects described in the section immediately above, because we have listed each meeting (with links to its results) of each project separately in this section.

- 1) The Socioeconomic Effects on Public Sector Information on Digital Networks – Towards a Better Understanding of Different Access and Reuse Policies, 4-5 February 2008, Paris, France, and publication of the report June 2009.
- 2) CODATA International Conference and General Assembly, 5-10 October 2008, Kyiv, Ukraine.
- 3) Inaugural Meeting of the Board on Research Data and Information, 29-30 January 2009.
- 4) Public Symposium on Author Deposit Mandates for Federal Research Grantees, 29 January 2009.
- 5) The Third U.S.-China Roundtable on Scientific Data Cooperation, 23-24 March 2009, Qingdao, China.
- 6) Symposium on Common Use Licensing of Publicly Funded Scientific Data and Information, 25 March 2009, Beijing, China, and 27 March 2009, Taipei, Taiwan.
- 7) The Second Meeting of the Board on Research Data and Information, 24-25 September 2009.

- 8) Public Symposium on Scientific Data for Evidence Based Policy and Decision Making, 24 September 2009.
- 9) International Symposium on Designing the Microbial Research Commons, 8-9 October 2009, Washington, DC, and publication of the report in September 2011.
- 10) Symposium on Data Sharing Plans for GEOSS and the Benefits of Data Sharing for Science, 16 November 2009 (in conjunction with the Group on Earth Observations (GEO) Plenary Meeting in Washington, DC).
- 11) Meeting with Chinese Microbiology Research Delegation, 7 December 2009.
- 12) AAAS Symposium on Managing the Exaflood: Enhancing the Value of Networked Data for Science and Society, 19 February 2010 (organized by BRDI members and staff).
- 13) The Fourth Meeting of the U.S.-China Roundtable on Scientific Data Cooperation, 29-30 March 2010, Irvine, California.
- 14) The Third Meeting of the Board on Research Data and Information, 3-4 June 2010.
- 15) The Changing Role of Libraries in Support of Research Data Activities, 3 June 2010, Washington DC.
- 16) International CODATA Conference and General Assembly, 24-27 October 2010, Cape Town, South Africa.
- 17) The Fourth Meeting of the Board on Research Data and Information, November 30 – December 1, 2010.
- 18) Public Symposium on “The Value of Shared Access and Reuse of Publicly Funded Scientific Data”, December 1, 2010.
- 19) The Future of Scientific Knowledge Discovery in Open Networked Environments: A National Symposium and Workshop, Washington, DC, March 10-11, 2011.
- 20) BISO-BRDI-CFRS International Symposium on the Case for International Scientific Data Sharing: A Focus on Developing Countries, April 18-19, 2011.
- 21) The Fifth Meeting of the Board on Research Data and Information, 13-14 June, 2011.
- 22) Public Symposium on “Crowdsourcing: Improving the Quality of Scientific Data Through Social Networking”, June 13, 2011.
- 23) Developing Data Attribution and Citation Practices and Standards: An International Symposium and Workshop, August 22-23, 2011, Berkeley, California.

### **C) Impacts on Agencies, Scientific fields, and other Audiences**

The Board has existed less than 3 years and a substantial amount of the start-up time was underfunded and thus not as productive as originally planned. Many of our activities therefore have not yet produced reports or measurable impacts.

We submit an annual (or quarterly) report of its activities to its sponsors, and to the members of the Board at the semi-annual meetings. In addition to these reports, the Board plans to prepare an evaluation of its activities prior to the first four years of its operation, when several other reports are released, to assess the its performance related to its mission. It will be done according to the following criteria, and incorporate the results of the PGA review:

#### **1) Fulfillment of the BRDI Statement of Task:**

Each activity will be evaluated to see if it has met the objectives of one or more tasks in its Statement of Task.

#### **2) Sponsor Interests:**

After completion of each activity, BRDI will seek feedback from its sponsors to determine the extent to which the activity has fulfilled the interests of its sponsors, and has met their expectations. A renewal of funding by the sponsors will be one important indicator of successful fulfillment of the sponsor interests.

#### **3) Timeliness:**

Time actually taken to perform each activity will be measured against the time of completion that was estimated at the time of the planning of the activity. The difference between the two will be a measure of success or failure of the activity.

#### **4) Budget:**

Funds actually spent on each activity will be measured against the estimated budget prepared at the time of planning the activity. The difference between the two will be a measure of the success or failure of the activity.

#### **5) Follow-up:**

If the completed BRDI activity leads directly or indirectly to another follow-up activity undertaken by BRDI, or another NRC Board or Committee, then it would be an indicator of success of the completed BRDI activity.

#### **6) Other Measurable Impacts:**

The impact of a BRDI activity, nationally and internationally, will be measured by the following indicators:

- a) Policy or management changes that are directly or indirectly influenced by the activity;
- b) Mention of the activity in reports, workshops, symposia, meetings, congressional record, and in the media;
- c) Articles in scholarly journals and in the media that reference the findings, conclusions, or recommendations of the activity.

There are, of course, intangible indicators of impact as well, such as the results of outreach to various constituencies, the internalized knowledge gained by the audiences at the meetings and public symposia, or the changes in views or practices by experts that are not referenced or too subtle to acknowledge. For this reason, we devote a lot of attention to establishing relationships with different constituencies and have a suite of activities that are targeted at different audiences.

## **D) New Agency Programs and Projects Under Development**

The first study project below is almost fully funded and about to begin. The next two projects have been in the planning stages for the past few months and will be held in the coming months and years as part of the Board's core program. These will be recurring on a regular basis, as described in each project summary. Finally, the last three projects have already been discussed with the Board members and some sponsors, and are expected to be initiated in the 2012-2013 period.

### ***1) Future Career Opportunities and Educational Requirements for Digital Curation***

An ad hoc committee of the BRDI will soon begin a study on future career opportunities and educational requirements for digital curation. The Statement of Task of the study is:

1. Identify the various practices and spectrum of skill sets that comprise digital curation, looking in particular at human versus automated tasks, both now and in the foreseeable future.
2. Examine the possible career path demands and options for professionals working in digital curation activities, and analyze the economic and social importance of these employment opportunities for the nation over time. In particular, identify and analyze the evolving roles and models of digital curation functions in research organizations, and their effects on employment opportunities and requirements.
3. Identify and assess the existing and future models for education and training in digital curation skill sets and career paths in various domains.
4. Produce a consensus report with findings and recommendations, taking into consideration the various stakeholder groups in the digital curation community, that address items 1-3 above.

The study was requested by IMLS in December 2010. It will be formally initiated in October 2011, with the study committee currently being identified and vetted. The study will include 4 committee meetings, a 2-day symposium, a questionnaire to the stakeholder community, and several commissioned papers. The project is expected to be completed in 18 months, resulting in consensus report and symposium proceedings.

The project cost is \$565K, with \$300K and \$100K provided by IMLS and NSF, respectively. Another \$165K is expected from NOAA, DOE, or the Sloan Foundation.

## **2) U.S. Forum on CODATA-World Data System (WDS) Cooperation**

The International Council for Science (ICSU) is promoting greater coordination and cooperation among its various organizations engaged in scientific data and information activities. Two of the largest such interdisciplinary bodies are the international CODATA and the new World Data System (formerly the World Data Centers, arising from the 1957 International Geophysical Year). The Board and U.S. CODATA are in the process of establishing a cooperation forum with the U.S. members of the World Data System (over 20 expected by next year). The purposes of this Forum will be to:

- 1) Serve as an informal mechanism to cooperate and coordinate on issues of mutual interest to the U.S. CODATA and the U.S. World Data System members, in the context of the international CODATA and WDS activities; and
- 2) Provide an opportunity for the managers and other principals of the World Data System in the United States, the members of BRDI and U.S. CODATA, and other constituents of the two bodies to share experiences, exchange ideas, and provide a direct link for expert input on each other's activities for mutual benefit.

BRDI will organize a half-day meeting in conjunction with its regularly scheduled Board and U.S. CODATA semi-annual meetings, beginning with the next one.

This activity will be funded by the Board's core support and will not require GBEC approval for the initial planning meeting. Travel for BRDI/US CODATA members is already included in the budget and the WDS members will fund their own travel. A steering committee and formal GBEC support will be sought in the event that the Forum participants decide that a meeting ReCap report will be desirable for future events.

## **3) Sponsor Forum**

BRDI is establishing a Sponsor Forum in collaboration with all its supporting organizations. The purposes of the Sponsor Forum will be to: (1) identify the most common current and emerging needs and interests of the sponsors; (2) serve as an opportunity for the sponsors to engage in interagency discussions of key issues in each subject area; and (3) use that information to plan the Board's activities to serve the sponsors' common interests that are consistent with its own mission and Statement of Task.

The first half-day Sponsor Forum is expected to be focused on the data management mandates that all the agencies have instituted or are in the process of developing pursuant to the America COMPETES Act. That meeting will be held in the November-December 2011 time period. Future Sponsor Fora are expected to be convened once per year.

This activity will be funded by the Board's core funds since participants will be mostly local and any Board member travel costs will be minimal. No written summary is expected from the first meeting and formal GBEC approval has not been sought or needed.

#### **4) A National Symposium and Workshop on Research Data as Intellectual Property: How the Law Influences Data Sharing**

Two of the most significant impacts of digital technologies and networks on scientific research have been the great increase in the amount of data generated or collected by researchers and a tremendously improved infrastructure for sharing, aggregating or recombining data sets. With these changes have also come a series of challenges concerning preservation, authentication, annotation, and provenance. As funders and researchers contemplate methods for responding to these opportunities and challenges, they often find themselves in doubt about the legal framework that governs research data. Are data "owned"? If so, by whom? What rights do the "owners" have against those who would copy, redistribute or reuse the data without permission? Conversely, how can a researcher share data over the Internet and assure other researchers that they have any permission they might need to reuse the data as they wish? In some cases, the law has been used to impede productive data sharing by researchers. In other cases, uncertainty about the law has impeded researcher collaboration even though the law itself would pose no barrier to such collaboration.

The Board will establish a steering committee to organize a symposium and workshop in Washington, DC to bring together key stakeholders in this area for intensive and structured discussions in order to obtain a better understanding of the ways in which the law of intellectual property, licenses, and contracts affect data sharing. The discussion also will cover the policies that require data management plans for the data that support published articles, and several case studies will be presented. A report will be published within a year of the meeting.

The project will cost approximately \$240K. Funding is expected to be provided by the National Science Foundation and one or more other funders, including the in-kind staff support through the Library of Congress IPA.

#### **5. Study on Sustainability Strategies for Noncommercial Scientific Databases**

An ad hoc committee of the Board proposes to conduct a study on sustainability strategies for noncommercial scientific databases, including the following tasks (not fully vetted yet by the board members, or approved by the GBEC):

1. Identify and broadly characterize the types of publicly-funded, noncommercial scientific data that would be of broad interest and use by other scientists over time. Develop the criteria or value proposition by which different types of scientific data would be selected for retention.
2. Broadly characterize the status of publicly-funded, noncommercial scientific data retention practices and models, focusing in particular on those of the sponsors. Discuss the different institutions and their roles and responsibilities in a strategic network of short- and long-term data repositories. Describe their functions individually and collectively.
3. Examine the existing sustainability business models and strategies of the different types of data repositories and evaluate the direct and indirect costs and benefits of each major model from an economic, financial, and science policy perspective. Identify and analyze the incentives, principal barriers, and other key factors that lead to success or failure.

4. Provide conclusions and recommendations to the sponsors regarding an overall sustainability strategy, including the optimal models for maintaining and making available different types of scientific data over the long term.

The study will be performed in 20 months and the resulting report will be published in accordance with NRC procedures.

This project was first raised in discussions with the Board's liaisons from the National Science Foundation and discussed with other sponsors at the last Board meeting. Funding of approximately \$600K will be sought from the existing sponsors and other sources over the next year.

## ***6. The Impact of Data Management Mandates***

The BRDI will establish an as hoc committee to undertake a consensus study to review the impact of federal data access and management mandates on research grantees, to judge whether such mandates are working most effectively, how ambiguities might be resolved, and what changes might be suggested.

The study would be performed according to the following statement of task (not fully vetted yet by the board members, or approved by the GBEC):

1. Assess the community experience with data availability and management requirements. Analyze the initial evidence of the benefits of more open data, and how much of a burden the requirement is placing upon research projects.
2. Describe the different choices facing each discipline in terms of confidentiality, embargoes, retention, and curation.
3. Provide conclusions and recommendations to the sponsor(s) regarding their data management mandates to grantees, indicating what aspects of those mandates are working well and which ones may need to be adjusted further.

The study will be performed in 20 months and the resulting report will be published in accordance with NRC procedures.

This project was discussed with other sponsors at the last Board meeting, and is expected to begin in 2013. Funding of approximately \$600K will be sought from the existing sponsors and other sources.

## SECTION IV: BOARD/COMMITTEE AND STAFF PROFILES

### A) BRDI Staff Biographies

**PAUL F. UHLIR** is Director of the Board on Research Data and Information at the U.S. National Academies in Washington, DC. Paul's area of emphasis is on issues at the interface of science, technology, and law, with primary focus on digital data and information policy and management. He also directs the U.S. Committee on Data for Science and Technology, and the InterAcademy Panel's Program on Digital Knowledge Resources and Infrastructure in Developing Countries. From 1999 to 2008, Paul was director of the Office on International S&T Information Programs at the National Academies; from 1991 to 1999, he was the Associate Executive Director of the Commission on Physical Sciences, Mathematics, and Applications; and from 1985 to 1991 he was senior staff officer for the Academies' Space Studies Board, where he directed projects about solar system exploration and environmental remote sensing programs for NASA.

Before joining the National Academies, he worked at the Office of the General Counsel and was a foreign affairs officer at the National Oceanic and Atmospheric Administration, where he worked on remote sensing law and policy and on intergovernmental agreements for cooperation in meteorological satellite programs. Paul is the author or editor of 23 books, and over 60 technical articles. He has been involved in numerous consulting and pro bono activities, and speaks worldwide on a broad range of information policy and management issues.

He has served as private-sector adviser to the US Department of State for the U.N. World Summit on the Information Society and for the Organization for Economic Co-operation and Development, and a co-chair of the Group on Earth Observations' Data Sharing Task Force, among other organizations and activities. In 1997 he received the National Research Council's Special Achievement award and in 2010 the international CODATA Prize, both for his work on international data policy. Paul has a J.D. and an M.A. degree in international relations from the University of San Diego, and a B.A in history from the University of Oregon.

**SUBHASH KUVELKER** is a part-time Senior Program Officer of the Board on Research Data and Information (BRDI). He has been a professional consultant in the fields of technology, economics and law for more than 10 years. Prior to that he was a manager at the Advanced Technology Program of the US Department of Commerce. He is an active member of the Bar in Maryland and in the District of Columbia, and is a Registered Patent Attorney. He holds a Bachelor of Technology in Chemical Engineering, a Master of Business Administration in Finance and Management Science, a Master of Economics in International Finance and Trade, and Juris Doctor in International Law and Finance.

**DANIEL COHEN** is working at the National Academies as a Program Officer at BRDI while on detail from the Library of Congress, where he has been employed since 2001 as a systems planning analyst. From 2000 to 2001, he managed the documentation department of CMSI, Inc., a software company producing risk management software for the lending industry. From 1990 to 2000, he worked for The Bureau of National Affairs, Inc., a publishing company, first as a legal editor, and subsequently as a technical writer and then as manager of the documentation department. From 1984 to 2000, Dan worked as a solo practitioner attorney, primarily representing juveniles in delinquency and abuse/neglect proceedings. He holds a BA in

philosophy from Beloit College, and a JD from the Columbus School of Law at the Catholic University of America. His professional interests include Information Architecture, Human-Computer Interaction (Usability), Project Management, and Process Improvement. His personal interests include photography, cooking, Pilates, and home improvement, (not necessarily in that order).

**MATT HAGGERTY** is a Financial Associate supporting the Board on Research Data and Information (BRDI). Matt has worked at the National Academies since 2008. In addition to BRDI, he also works with the Committee on Science, Technology and Law (CSTL), the Science and Technology for Sustainability Program (STS), the Committee on Human Rights (CHR), the Committee on Women in Science, Engineering and Medicine (CWSEM), and the Christine Mirzayan Science and Technology Policy Fellowship Program. Prior to joining the Academies, he worked in various financial positions at Whirlpool Corporation in Benton Harbor, Michigan. Matt graduated from the University of Notre Dame in 2007 with a Bachelor of Business Administration in Finance and Economics and is currently pursing a Master of Public Policy at the George Washington University in Washington, DC. His interests include baseball/softball, running, and playing the guitar.

**CHERYL WILLIAMS LEVEY** is a Senior Program Associate at the National Academies and provides support services to the Board on Research Data and Information (BRDI). Cheryl has over twenty years of experience working in the federal and international sectors. She has worked for the Committee on Data for Science and Technology (CODATA) as well as for the Department of Commerce (National Institute of Standards and Technology, National Institutes of Health) and for the Department of the Interior (U.S. Geological Survey). Cheryl received her Bachelor of Science Degree in Communication Studies from the University of Maryland University College.

## **B) Board Roster and Biographies**

### Michael Lesk, (NAE) Chair

Professor and Chair, Department of Library and Information Science  
Rutgers University

### Roberta Balstad, Vice Chair

Senior Fellow, CIESIN  
Special Research Scientist, Earth Institute, Columbia University

### Maureen Baginski

Vice President of the intelligence business and National Security Advisor at SERCO

### Francine Berman

Vice President of Research  
Rensselaer Polytechnic Institute

### R. Steven Berry (NAS)

James Franck Distinguished Service Professor Emeritus  
Gordon Center for Integrative Studies  
James Franck Institute  
University of Chicago

### Christine Borgman

Professor and Presidential Chair  
University of California, Los Angeles

### Norman Bradburn

Emeritus Professor  
University of Chicago

### Bonnie Carroll

President, Information International Associates

Michael Carroll

Professor  
Washington College of Law  
American University,

Paul David

Senior Fellow  
Stanford Institute for Economic Policy  
Research, Professional Fellow of Merit

Barbara Entwistle

Interim Vice Chancellor for Research and  
Economic Development  
University of North Carolina

Michael Goodchild (NAS)

Professor of Geography  
University of California-Santa Barbara

Alyssa Goodman

Professor  
Department of Astronomy  
Harvard University

Margaret Hedstrom

Associate Dean for Academic Programs and  
Professor, School of Information  
University of Michigan

Michael Keller

Professor and Director of Academic  
Information Resources  
University Libraries and Academic  
Information Resources  
Stanford University

Alexa McCray (IOM)

Associate Professor of Medicine  
Harvard Medical School

Michael R. Nelson

Visiting Professor, Internet Studies,  
Georgetown University

Daniel Reed

Vice President, Microsoft Research,  
Microsoft Inc.

Cathy H. Wu

Edward G. Jefferson Professor of  
Bioinformatics & Computational Biology,  
University of Delaware  
Professor of Biochemistry and Molecular &  
Cellular Biology, Georgetown University  
Medical Center

**EX-OFFICIO MEMBERS**

Robert Chen (CODATA Secretary General)  
Director CIESIN, Columbia University

Michael Clegg (NAS Foreign Secretary)  
Department of Ecology & Evolution

Sara Graves (CODATA Executive  
Committee)

Director, Information Technology and  
Computing Center  
Professor of Computer Science  
The University of Alabama- Huntsville

Steve Kelling (Co-Chair, Biodiversity Task  
Group)  
Cornell Lab of Ornithology  
Cornell University

Chris Lenhardt

(Co-Chair, Preservation of and Access to  
Scientific and Technical Data in Developing  
Countries)  
Informatics Scientist, ORNL DAAC  
Deputy Manager, Environmental Sciences  
Division  
Oak Ridge National Laboratory

Mark A. Parsons (Co-Chair, International  
Polar Year Data and Information System  
Task Group)

National Snow and Ice Data Center/World  
Data Center for Glaciology  
University of Colorado

## **FEDERAL GOVERNMENT LIAISON MEMBERS**

Elaine S. Collier

Assistant Director for Clinical Research  
Division for Clinical Research Resources  
National Center for Research Resources  
National Institutes of Health

Wo Chang

Group Leader, Digital Media Group  
National Institute of Standards and  
Technology

Leslie A. Collica

Deputy Chief, Information Access Division  
National Institute of Standards and  
Technology

Scott Hausman

Acting Director, NOAA NCDC  
National Climatic Data Center

Deirdre Jones

Director, Systems Engineering Center  
NOAA, National Weather Service  
Office of Science and Technology

Suzette M. Kimball

Deputy Director  
U.S. Geological Survey

Mimi McClure

National Science Foundation  
Office of Cyberinfrastructure

Lewis McCulloch

Contractor  
NOAA, National Environmental Satellite  
Data and Information Service (NESDIS)

Rob Pennington  
Office of Cyberinfrastructure  
National Science Foundation

Joyce Ray

Associate Deputy Director for Library  
Services  
Institute of Museum and Library Services  
(IMLS)

Kamie Roberts

Associate Director for Federal and Industrial  
Relations -Information Technology  
Laboratory  
National Institute of Standards and  
Technology

Jerry Sheehan

Assistant Director for Policy Development  
National Library of Medicine  
National Institutes of Health

Sylvia Spengler

Program Director, Bioinformatics  
National Science Foundation

George Strawn

Director, National Coordination Office for  
Networking and Information Technology  
Research and Development (NITRD)

Charles Thomas

Senior Library Program Officer  
Institute of Museum and Library Services  
(IMLS)

Peter Young

Consultant  
Library of Congress

## **BRDI BOARD MEMBER BIOGRAPHIES**

*As of September 7, 2011*

### **Michael Lesk, *Chair***

Michael Lesk (NAE) is the Chair of the Board on Research Data and Information. He is Professor and the Chair of the Department of Library and Information Science in the School of Communication, Information, and Library Studies at Rutgers University. Prior to that, Prof. Lesk worked in a research group at Bellcore, and on the CORE project, a large chemical information system with Cornell, OCLC, and the Chemical Abstracts Service. From 1998-2002 he headed the Division of Information and Intelligent Systems at the National Science Foundation. Professor Lesk has also worked on the SMART retrieval code and other software projects. His current research interests are digital libraries and digitization of data and scientific information. He has written numerous books and professional articles. Professor Lesk received the “Flame” award for lifetime achievement from Usenix in 1994 and is a Fellow of the Association of Computing Machinery (ACM). His service to the NRC includes Committee on Applied and Theoretical Statistics (Member; 10/4/2005 -- 6/30/2008), and Computer Science and Telecommunications Board (Member - Resigned from committee; 7/1/1997 -- 12/22/1997). He received his PhD in Chemical Physics from Harvard University in 1969, and B.A. in physics and chemistry from Harvard College in 1964.

### **Roberta Balstad, *Vice Chair***

Roberta Balstad is the Vice Chair of the Board on Research Data and Information. In her capacity as Vice Chair she will be primarily responsible for oversight of the projects and activities of the US National Committee for CODATA. Dr. Balstad is Senior Research Scientist at Columbia University and until recently Director of the University's Center for International Earth Science Information Network (CIESIN). CIESIN is a center within the Earth Institute at Columbia University that works at the intersection of the social, natural, and information sciences, and specializes in on-line data and information management, spatial data integration and training, and interdisciplinary research related to human interactions in the environment. Dr. Balstad has published extensively on science policy, information technology and scientific research, and the role of the social sciences in understanding global environmental change. Dr. Balstad has chaired or served on several NRC committees over the years, including the Committee on Climate Data Records from Operational Satellites: Development of a NOAA Satellite Data Utilization Plan (2003-2005); Panel on Earth Science Applications and Societal Needs (2005); Committee on the Geographic Foundation for Agenda 21 (2001-2003); Steering Committee on Space Applications and Commercialization (1999-2003); Committee on Global Change Research (1999-2002); Space Studies Board (1995-2000); Climate Research Committee (1997-1999); Panel on Climate Observing Systems Status (1998-1999); Task Group on Research and Analysis Programs (1996-1998); and Committee on Buildings & Community Systems Energy Conservation (1982-1983). She is currently the Chair of the US National Committee for CODATA at the NRC, as well as a member of the U.S. National Committee for the International Institute for Applied Systems Analysis, the Committee on Earth System Science for Decisions about Human Welfare: Contributions of Remote Sensing, and the Panel on Earth Science Applications and Societal Needs. Dr. Balstad received her Ph.D. in history in 1974, her MA in history in 1970, and her BA in history in 1964, all from the University of Minnesota. See her full CV at the end of this memo for additional details.

## **Maureen Baginski**

Maureen Baginski is President of National Security Systems Sector of Sparta, Inc. From November 2005 to October 2006 she was a Director in the Intelligence Sector of BearingPoint. From 2003 to 2005, Ms. Baginski served as the FBI's Executive Assistant Director for Intelligence (EADI) where she was responsible for establishing and managing the FBI's first-ever intelligence program, including technology acquisition and workforce development. From 1979-2005, she was employed at the National Security Agency, where she held a variety of positions, including lead analyst for the Soviet Union, Executive Assistant to the Director, Chief Policy and Customer Support, Chief Office of the Director, Assistant Deputy Director for Technology, and Director of SIGINT. Ms. Baginski is the recipient of two Presidential Rank Awards, two Director of Central Intelligence National Achievement Medals, the Director of Military Intelligence's Leadership Award, and NSA's Exceptional Civilian Service Award. In addition, Ms Baginski was the first-ever recipient of NSA's Outstanding Leadership Award. In December 2005 she received an honorary Doctorate of Humane Letters from the University of Albany for her service to the nation. Ms. Baginski holds a BA and MA in Slavic Languages and Linguistics from the University of Albany.

## **Francine Berman**

Francine Berman is Vice President for Research at Rensselaer Polytechnic Institute. Prior to 2010, she was Professor and High Performance Computing Endowed Chair in the Computer Science and Engineering Department at University of California at San Diego (UCSD). She is currently serving as Director of the San Diego Supercomputer Center, a national cyberinfrastructure facility. Prof. Berman was on the Purdue University faculty for five years prior joining UCSD in 1984. Her academic research during the past two decades has focused on high-performance and grid computing, in particular in the areas of programming environments, adaptive middleware, scheduling, and performance prediction. Most recently, she has led or co-led the AppLeS (Application-Level Scheduling) Project, the design and development of adaptive middleware for Grid environments, and the large NSF "Virtual Instrument/MCell" Information Technology Research project. From 2001 to 2004, Prof. Berman served as Director of the National Partnership for Advanced Computational Infrastructure, a consortium of 41 research groups, institutions, and universities. She is one of two founding principal investigators of the NSF-supported TeraGrid, the largest coordinated grid deployment project to date. Prof. Berman is a Fellow of the ACM and serves on numerous technical, advisory, conference, and scientific committees including the NSF Engineering Advisory Committee and the NIH NIGMS Advisory Council. She also is a founding member of the Computing Research Association's Committee on the Status of Women. Prof. Berman received her Ph.D. in mathematics from the University of Washington in 1979, and B.A. in mathematics from UCLA in 1973.

## **R. Stephen Berry**

(NAS) is James Franck Distinguished Service Professor Emeritus at Gordon Center for Integrative Studies, James Franck Institute, University of Chicago. His research interests include the structure and dynamics of simple atoms and molecules; intramolecular energy transfer; electron correlation; few-body dynamics; melting, freezing, and dynamics of clusters. multi-photon excitation and ionization; and energy and resource management. His service to the NRC includes: Report Review Committee (Ex Officio Member; 7/1/2004 -- 6/30/2010), Committee on Nuclear Forensics (Member; 1/25/2008 -- 7/20/2009), Panel on Energy Efficiency (Member; 6/30/2007 -- 12/31/2008), Committee on Review of the DOE Nuclear Energy R&D Program (Member; 6/5/2006 -- 12/31/2007), Report Review Committee (Chair; 7/1/2000 -- 6/30/2004), 2004 NAS Public Welfare Medal Selection Committee (Chair; 3/1/2003 -- 4/30/2004), U.S.-Iran Joint Workshop on Issues Concerning Higher Education (Member; 10/8/2002 -- 3/31/2004), Council of the National Academy of Sciences (Home Secretary; 7/1/1999 -- 6/30/2003), Executive Committee of the NAS Council (Member; 7/1/1999 -- 6/30/2003), and NAS Council Committee on Membership Affairs (Chair; 7/1/1999 -- 6/30/2003). He received his Ph.D. in physical chemistry in 1956, A.M. in chemistry in 1953, and B.A. in chemistry in 1952, all from Harvard University.

## **Christine Borgman**

Christine L. Borgman is Professor and Presidential Chair in Information Studies at the University of California, Los Angeles (UCLA). She is a co-principal investigator for the Center for Embedded Networked Sensing (CENS) funded by the National Science Foundation. She is the author of more than 150 publications in the fields of information studies, computer science, and communication. Her NRC committee service includes membership on the Study Committee on Internet Navigation and the Domain Name System (Computer Science and Telecommunications Board), and currently on the US National Committee for CODATA. She also is serving on the Advisory Board to the Electronic Privacy Information Center, and the Association for Computing Machinery Public Policy Committee. Prof. Borgman holds a PhD in Communication from Stanford University (1984), an MLIS from the University of Pittsburgh (1974), and a BA in mathematics from Michigan State University (1973).

## **Norman Bradburn**

Norman M. Bradburn, is the Tiffany and Margaret Blake Distinguished Service Professor Emeritus, and serves on the faculties of the Harris School, the Department of Psychology, the Graduate School of Business, and the College, at the University of Chicago. He is a former provost of the University (1984-1989), chairman of the Department of Behavioral Sciences (1973-1979), and associate dean of the Division of the Social Sciences (1971-1973). From 2000-2004, he was the assistant director for Social, Behavioral and Economic Sciences at the National Science Foundation. Prof. Bradburn is currently a senior fellow at the National Opinion Research Center (NORC). He has been director of NORC and president of its Board of Trustees. Prof. Bradburn, a social psychologist, has been at the forefront in developing theory and practice in the field of sample survey research. He has written numerous books and articles on the subject of social sciences research. His NRC active service includes Data Panel (Member; 10/24/2005 -- 12/31/2008), Committee on the State of the USA Health Indicators (Member; 5/30/2008 -- 11/30/2008), and his inactive service includes Committee on Standards of Evidence and the Quality of Behavioral Social Science Research (Chair; 11/1/2004 -- 4/30/2005), Roundtable on

Social and Behavioral Sciences and Terrorism (Ex Officio Member; 3/13/2002 -- 6/30/2004), Panel to Review the Statistical Procedures for the Decennial Census (Member; 8/3/1998 -- 3/31/2004), Committee on the Evaluation of National and State Assessments of Educational Progress (Member; 2/1/1996 -- 6/30/1999), Committee on National Statistics (Chair; 7/1/1993 -- 6/30/1998), and Commission on Behavioral and Social Sciences and Education (Member; 4/7/1982 -- 6/30/1987). He received his PhD in Social Psychology (1960) and MA Clinical Psychology (1958), both from Harvard University, BA in Philosophy, Politics, and Economics from University of Oxford (1955), and BA from The University of Chicago (1952).

### **Bonnie Carroll**

Bonnie Carroll is founder and President of Information International Associates, Inc. . She is also the Secretariat Director of CENDI (U.S. Departments of Commerce; Energy; EPA, NASA; National Libraries of Agriculture, Education, and Medicine; Defense; and Interior Information Managers Group), an interagency cooperative effort of the scientific and technical information (STI) managers of ten major science and engineering agencies. In addition, Ms. Carroll serves as a Senior Consultant to the Biological Resources Division of the U.S. Geological Survey and the Department of Energy's (DOE) Office of Scientific and Technical Information (OSTI) and is the planning lead for the Southern Appalachian Information Node (SAIN) of the National Biological Information Infrastructure. She participated in a major assessment of STI in the U.S. for the National Science Foundation, the development of a topography of science and technology information systems for the Library of Congress, and the development of a system to provide input to the DOE's Energy Science and Technology Database. She has been a senior advisor to the NASA STI program, was a consultant to Conservation International on knowledge management, and did a survey for the Department of Commerce to examine U.S. industry requirements for foreign scientific and technical information. She has been a consultant to foreign and international organizations including the International Atomic Energy Agency (IAEA), the Kingdom of Jordan, UNESCO, and the International Council for Scientific and Technical Information (ICSTI). Her service to the NRC includes U.S. National Committee for CODATA (Member; 7/14/2003 -- 6/30/2009), and U.S. National Committee for CODATA (Ex Officio Member; 7/1/2000 -- 6/30/2005). She received her MS in library and Information Science from Columbia University and BA from Cornell University.

### **Michael Carroll**

Michael W. Carroll is Professor of Intellectual Property Law at the Washington College of Law of American University in Washington, D.C. Prior to 2010, he was a member of the faculty of Villanova University School of Law. He teaches and writes about intellectual property law and cyberlaw. Professor Carroll's research focuses on the search for balance in intellectual property law over time in the face of challenges posed by new technologies. Prior to entering the academia, he served as a law clerk to Judge Judith W. Rogers, U.S. Court of Appeals for the D.C. Circuit, and Judge Joyce Hens Green, U.S. District Court for the District of Columbia. He practiced law at Wilmer, Cutler & Pickering in Washington, D.C. He is a founding member of Creative Commons, Inc., a global organization that provides free, standardized copyright licenses to enable and to encourage legal sharing of creative and other copyrighted works. He also is on the sub-group of board members of Creative Commons who advise the organization's Science Commons division and its education division, ccLearn. He has a JD. (magna cum laude)

from the Georgetown University Law Center (1996), and a B.A. in Anthropology from the University of Chicago (1986).

### **Paul David**

Paul Alan David is Professor Titulaire and Chaire, Innovation et regulation des services numeriques – Telecom-Paris Tech, and L'école Polytechnique Research Chair in the Digital Economy, Paris, France. He is Professor of Economics (Emeritus) and Senior Fellow of the Institute for Economic Policy Research at Stanford University. He was until recently Professor Emeritus of Economics and Economic History in the University of Oxford, Emeritus Fellow of All Souls College, Oxford., as well as a Senior Fellow of the Oxford Internet Institute. He is nominated as a member of the Board because of his expertise in economics.

### **Barbara Entwistle**

Barbara Entwistle is Director of the Carolina Population Center at the University of North Carolina. From 1994-2007, she was Professor of Sociology and from 1985-1994 she was Associate Professor of Sociology at the University of North Carolina. She is currently active on the NAS Committee on Population (Member; 4/10/2006 -- 3/31/2012) and has served on the following committees in the past: U.S. National Committee for the International Institute for Applied Systems Analysis (Member - Resigned from committee; 6/1/2005 -- 1/1/2009); Committee on the Human Dimensions of Global Change (Member; 2/3/2000 -- 7/31/2005); Panel on New Research on Population and the Environment (Chair; 1/3/2003 -- 6/30/2005). Dr. Entwistle holds a Ph.D. in Sociology from Brown University (1980). She is nominated because of her expertise in social science data.

### **Michael Goodchild**

Michael F. Goodchild (NAS) is Professor of Geography at the University of California Santa Barbara (UCSB); Chair of the Executive Committee, National Center for Geographic Information and Analysis (NCGIA); Associate Director of the Alexandria Digital Library Project; and Director of NCGIA's Center for Spatially Integrated Social Science. Prior to joining the faculty at UCSB, Prof. Goodchild taught at the University of Western Ontario. He is a Foreign Fellow of the Royal Society of Canada. He has received numerous awards including the Canadian Association of Geographers Award for Scholarly Distinction, the Association of American Geographers Award for Outstanding Scholarship, the Canadian Cartographic Association's Award of Distinction for Exceptional Contributions to Cartography, the Educator of the Year Award from the University Consortium for Geographic Information Science, and the Lifetime Achievement Award from Environmental Systems Research Institute, Inc. He has published numerous articles and books in the field of geography, remote sensing, and spatial databases. His service to NRC includes Committee on Strategic Directions for the Geographical Sciences in the Next Decade (Member; 2/15/2008 -- 7/31/2009), Committee on Applied and Theoretical Statistics (Member; 11/10/2006 -- 6/30/2009), Panel on Earth Science Applications and Societal Needs (Member; 2/18/2005 -- 1/31/2007), Committee on Planning for Catastrophe: A Blueprint for Improving Geospatial Data, Tools, and Infrastructure (Chair; 10/19/2004 -- 12/31/2006), Geographical Sciences Committee (Member; 1/19/2001 -- 12/31/2006), Committee

on Beyond Mapping: The Challenges of New Technologies in the Geographic Information Sciences (Member; 4/16/2002 -- 6/30/2006), Committee for a Workshop on Visualization of Uncertain Data (Member; 11/5/2004 -- 3/31/2005), Commission on Physical Sciences, Mathematics, and Applications (Member; 7/22/1999 -- 12/31/2000), Mapping Science Committee (Chair; 1/1/1997 -- 12/31/1999), Panel on Distributed Geolibraries: Spatial Information Resources (Chair; 10/30/1997 -- 6/30/1999), and Mapping Science Committee (Member; 6/1/1992 -- 6/30/1996). He has received honorary doctorates from Laval University (1991), and from Keele University (2001). He received his Ph.D. in Geography from McMaster University in 1969, and his B.A. in physics from Cambridge University in 1965.

### **Alyssa Goodman**

Alyssa Goodman is Professor of Astronomy at Harvard University, and a Research Associate of the Smithsonian Astrophysical Observatory of the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts. Her research interests are in the fields of mapping of space and scientific computing. She co-founded the Initiative in Innovative Computing (IIC) at Harvard and served as its Director from 2005-2008. Prof. Goodman has been a visiting professor at Yale University, Post-doctoral Fellow at University of California at Berkeley, and Summer Fellow at NASA-Goddard Institute for Space studies. She has published numerous scholarly articles in the fields of astronomy. She was awarded the Newton Lacy Pierce Prize from the American Astronomical Society for her work on interstellar matter and has received other numerous awards and fellowships. She currently serves as Chair of the Astronomy Section of the American Association for the Advancement of Science and is a member of many professional societies. Prof. Goodman received her Ph.D. in physics from Harvard in 1989, and B.S. in physics from MIT in 1984.

### **Margaret Hedstrom**

Margaret Hedstrom is an associate professor in the School of Information. Before joining the U-M faculty in 1995, Hedstrom was chief of state records advisory services and director of the Center for Electronic Records at the New York State Archives and Records Administration (1985-95). She has served on the following NAS Committees: Committee to Study Digital Archiving and the National Archives and Records Administration (Member; 6/1/2002 -- 6/30/2005) and Committee on Information Technology Strategy for the Library of Congress (Member; 1/18/1999 -- 6/30/2001). She holds a Ph.D. in History and a MLS in Library Science and History from the University of Wisconsin-Madison. Dr. Hedstrom is nominated because of her expertise in information management.

### **Michael Keller**

Michael A. Keller is the Ida M. Green University Librarian, and Director of Academic Information Resources at Stanford University. He is also the Publisher of the Highwire Press, which he established to provide online co-publishing services to scholarly journals, and of the Stanford University Press. He is a co-founder of the Stanford-California State Library Institute on 21st Century Librarianship. His research interests are online publishing, library science, and musicology. Prior to his appointment at Stanford, he was Associate University Librarian and

Director of Collection Development at Yale University. He also taught musicology at Stanford University, University of California at Berkeley, and Cornell University. He has served as advisor, consultant, and committee member to the American Association for the Advancement of Science and other scholarly societies. He received an M.L.S. from SUNY in 1971, a M.A in musicology from SUNY Buffalo in 1970, and B.A. in musicology from Hamilton College in 1967.

### **Alexa McCray**

Alexa T. McCray (IOM) is Associate Professor of Medicine at Harvard Medical School and the Department of Medicine, Beth Israel Deaconess Medical Center. She is the former Director of the Lister Hill National Center for Biomedical Communications, a research division of the National Library of Medicine (NLM) at the NIH. Before joining NLM in 1986, she was on the research staff of IBM's T.J. Watson Research Center. Prior to that Dr. McCray was on the faculty of Georgetown University where she earned her Ph. D. in linguistics in 1981. She conducted her pre-doctoral research at MIT. Dr. McCray's primary research area is biomedical informatics with a focus on health care information storage, dissemination and use. Her publications are in the areas of medical language processing, digital libraries, and consumer health informatics. Dr. McCray was elected to the IOM in 2001. She is a fellow of the AAAS and a fellow of the American College of Medical Informatics. She is a past member of the board of both the American Medical Informatics Association and the International Medical Informatics Association. Dr. McCray is nominated because of her expertise in health informatics.

### **Michael R. Nelson**

Michael R. Nelson is currently Visiting Professor, Internet Studies in Georgetown University's Communication, Culture, and Technology Program. Prior to his appointment at Georgetown University Dr. Nelson was Director of Internet Technology and Strategy at IBM. He has also held several positions in the public sector as Director for Technology Policy at the Federal Communications Commission, Special Assistant for Information Technology at the White House Office of Science and Technology Policy, and the senior staff member for the Senate's Subcommittee on Science, Technology, and Space. He was the lead Senate staffer for the High-Performance Computing Act. He was Chairman of the Board of Directors for the Telecommunications Policy Research Conference and serves on the Industry Strategy Council of the Internet2 research consortium. He participated in the World Summit on Information Society and Internet Governance Forum. He is serves on the Strategy Council of the UN's Global Alliance for ICT and Development. He serves as the chair the Technology Section of the American Association for the Advancement of Science (AAAS). Dr. Nelson received his Ph.D. in geophysics from MIT and B.S. in geology from Caltech.

### **Daniel Reed**

Daniel A. Reed is Director of Scalable Computing and Multicore at Microsoft Research. He is a computer and computational scientist, known for his contributions to high-performance computing and science policy. He founded the Renaissance Computing Institute in 2004 and served as its director until December 2007. Dr. Reed also was Chancellor's Eminent Professor,

and served as senior adviser for strategy and innovation to UNC-Chapel Hill Chancellor James Moeser. He served as CIO and Vice Chancellor for Information Technology Services at UNC-Chapel Hill from January 2004 through April 2007. Before coming to UNC-Chapel Hill, Dr. Reed spent 19 years at the University of Illinois at Urbana-Champaign, where he led the National Center for Supercomputing Applications (NCSA) from 2000 – 2003 and chaired the University of Illinois Computer Science Department. In 2001, Dr. Reed led the effort to launch the National Science Foundation's TeraGrid, and served as TeraGrid's chief architect through 2003. He was appointed to the President's Council of Advisors on Science and Technology (PCAST), by President Bush in 2006 and served on the President's Information Technology Advisory Committee (PITAC) from 2003 – 2005. He also has been chair of the Board of Directors of the Computing Research Association (CRA) since 2005, and a member of the board since 1998. Dr. Reed received his Ph.D. and M.S. from Purdue University, and B.S. from the University of Missouri, Rolla, all in computer science.

### **Cathy H. Wu**

Cathy Wu is the Director of Protein Information Resource (PIR), and Professor of Biochemistry and Molecular Biology at Georgetown University. She has conducted bioinformatics research since 1990, taught computer science, and developed several protein classification systems and databases, including one with a U.S. patent. She is a member of the Protein Structure Initiative Advisory Committee, NIGMS, NIH, and the Board of Directors of the International Society for Computational Biology. Additionally, Prof. Wu has served on many bioinformatics review panels for NIH, NSF, and DOE, and on numerous program committees for international bioinformatics conferences. She has published over 80 papers and three books, and is a frequent invited speaker for lectures and tutorials at universities, companies and conferences. Prof. Wu has the following degrees: B.S., Plant Pathology, National Taiwan University, Taiwan (1978); M.S., Plant Pathology (1982) and Ph.D., Molecular Plant Pathology (1984), both from Purdue University; and an M.S. in Computer Science from the University of Texas (1989).