Appendix A: Workshop Background

> It is exceedingly rare that fundamentally new approaches to research and education arise. Information technology has ushered in such a fundamental change. Digital data collections are at the heart of this change. They enable analysis at unprecedented levels of accuracy and sophistication and provide novel insights through innovative information integration. Through their very size and complexity, such digital collections provide new phenomena for study.
>
> —NSB Report: "Long-lived Data Collections: Enabling Research and Education in the 21st Century," September 2005.

Workshop Motivation

Digital technology offers the potential for fundamental change in the conduct of science for two reasons:

- Data access—Digital technology could make data openly accessible to scientists, reducing data-management burdens, formalizing generalizable and replicable science, and enabling new kinds of data-driven science.
- Knowledge access—Digital technology could facilitate the dissemination and transmission of knowledge by making information widely available electronically.

These technologies can transform science by enabling a broad range of scientists to create and transmit knowledge, educators to impart that knowledge to future generations, and decision-makers to make well-informed policy based on sound and reproducible research. But institutional and social barriers that exist limit the acceptance of these transformational technologies by the scientific community. It is likely that these barriers could be reduced by concerted and informed efforts by scientific funding agencies and international organizations.

An informed approach to digital technology can be fostered by advancing understanding of the technical, behavioral, and social factors conducive to its widespread adoption. These factors might include the role of economic incentives, human capital, social networks, and (most obviously) scientific attribution. Lessons can also be learned from practical experience. Fields as disparate as biotechnology, geosciences, and astronomy have been transformed by both data and knowledge access. Several broad-based initiatives, such as ORCID, the Brazilian Lattes Platform, and the VIVO project, have promoted widespread access to knowledge. Both the research and practical experiences should help identify new approaches that are neither nation specific nor domain

specific—indeed, that can be used in a cooperative international effort to help foster the adoption and use of digital technologies.

Objective

The goal of this workshop is to combine the expertise of computer and information scientists with those of behavioral and social scientists to identify guiding principles and approaches that can help inform organizations that fund research, scientific research organizations, and publishing houses. Specific questions that could be answered include the following:

> *Technical constructs*—What are the most important digital technologies that could be used to facilitate data and knowledge access? To what extent is progress being made already, and how can progress be accelerated? What role might the private sector play in bringing about change?
>
> *Social constructs*—What incentives are necessary to engage scientists in making data accessible and shared with the broader community? What are the appropriate business models necessary to promote connecting publications to data? How might private-sector participants be engaged in the effort? What are the social barriers to adopting and using unique researcher numbers?
>
> *The Pragmatic Experience*—What lessons have been learned from ORCID and the Brazilian experience? What can we learn from data preservation and libraries and other coordinated data and publication efforts? What can we learn from domain-specific successes?

Workshop Structure

The workshop is intended to be small and informal in nature, and the discussion will be focused on addressing the motivating questions. To achieve that goal, the workshop will have four separate sessions. Selected participants will be asked to start the dialogue by providing a short opening commentary on their experience with data and knowledge access, focusing on describing the technical and social challenges and how what was learned might help inform international efforts. All workshop participants are asked to participate in a discussion of opportunities and constraints. The discussions should focus on relating enabling aspects of the technologies to incentives for changing conduct, with the goal of identifying approaches that are related to the technological solutions. It is particularly important that the discussion address social, technical, and institutional challenges, using specific examples from previous experience. The moderator will provide a synthesis of the discussion and identify useful approaches for funding agencies.

To facilitate planning, all workshop participants for the sessions are asked to produce, in advance, a brief document summarizing their views about new approaches that could be used to foster the adoption and use of digital technologies. Although these views might be informed by either theory or practice, providing a list of relevant literature and examples would be extremely helpful. All participants will receive the briefing documents in advance, and it is expected that these documents will inform and lead the discussion at the workshop.

Appendix B: Workshop Structure

A diverse group of international research scientists, computer and information scientists, and behavioral and social scientists was invited to participate in the workshop. Invitations were also extended to representatives from funding agencies worldwide and publishers likely to participate as partners in open-access data initiatives. This provided a broad base of expertise encompassing social, professional, and technical issues related to data and knowledge access.

Workshop Wiki

NSF created a wiki to facilitate sharing information and ideas both before and after the workshop.[6] Before the workshop, background readings and participant biographies were provided, and each participant was asked to draft a white paper. The white paper was to be a brief document summarizing the participant's views on what approaches could be used to encourage adopting information and communication technologies to enable open access to data and data sharing. All workshop participants received briefing documents in advance of the workshop, with the expectation that these documents would inform the workshop discussions. In the spirit of openness, following the workshop, the presentations were posted and comments were solicited on each.

Sessions

By design, the workshop was small and informal in nature to focus discussion on the workshop's motivating questions. To achieve this goal, the workshop had four separate sessions:

1. Data Access—Digital technology and scientific communities.
2. Data Access—Digital technology and multiple scientific communities.
3. Knowledge Access—The role of scientific attribution.
4. Knowledge Access—The role of funding agencies.

Selected participants were asked to initiate and moderate the sessions by providing brief commentary on their experience with data and knowledge access, particularly technical and professional challenges and how the lessons learned from these might help inform international efforts to address open access to data more broadly. All workshop participants were encouraged to participate in the ensuing discussion of

[6] A "wiki" (Hawaiian for "fast") is a page or collection of interlinked Web pages designed to enable collaborative websites. Anyone who accesses the wiki is able to contribute or modify content using a simplified markup language.

opportunities and constraints, with the goal to develop “ideas, exemplars and concrete recommendations” (Seidel 2010). The discussions focused on identifying approaches that addressed both the potential of ICTs to enable data sharing and open access to data and the incentives for researchers, agencies, institutions, and publishing houses to facilitate this. Participants also addressed the technical and professional challenges to data sharing, drawing on specific examples from prior experience. At the end of each session, the moderator was asked to provide a synthesis of the discussion and identify useful approaches for funding agencies. For the purposes of this report, we have separated the workshop into two categories: (1) data access and (2) knowledge access and attribution.

Appendix C: Participant and Observer Biographies

Workshop Participants

Dr. Carlos Aragão, Professor of Physics, Universidade Federal do Rio de Janeiro. Dr. Aragão obtained a B.S. in Physics in 1973 and an M.S. in Physics in 1976 at the Pontificia Universidade Católica do Rio de Janeiro (PUC/RJ). He obtained a Ph.D., also in Physics, from Princeton University in 1980. Dr. Aragão is a Professor of Physics at the Universidade Federal do Rio de Janeiro (UFRJ). He is a Member of the Brazilian Academy of Sciences and was awarded the Brazilian National Order of Scientific Merit Medal in 1998.

Dr. Shenda Baker, Professor of Chemistry, Harvey Mudd College. Dr. Baker received a B.S. in Chemistry and French from Grinnell College in 1985. In 1991, she received a Ph.D. in Physical Chemistry from the California Institute of Technology. Dr. Baker is a Professor of Chemistry at Harvey Mudd College in California, where she became the Clare Boothe Luce Assistant Professor of Chemistry in 1993. In 1996, Dr. Baker received an NSF CAREER Award, the DOE Young Scientists and Engineers Award, and the Presidential Early Career Award for Scientists and Engineers. She is on the Executive Committee for the National Neutron Scattering Society of America and on the Advisory Board for the Office of Cyberinfrastructure at NSF.

Dr. Katy Börner, Victor H. Yngve Professor of Information Science, School of Library and Information Science, Indiana University. Dr. Börner received an M.S. in Electrical Engineering from the University of Technology in Leipzig in 1991 and a Ph.D. in Computer Science from the University of Kaiserslautern in 1997. She is the Victor H. Yngve Professor of Information Science at the School of Library and Information Science, Adjunct Professor in the School of Informatics, Core Faculty of Cognitive Science, Research Affiliate of the Biocomplexity Institute, Fellow of the Center for Research on Learning and Technology, Member of the Advanced Visualization Laboratory, and Founding Director of the Cyberinfrastructure for Network Science Center at Indiana University. She also serves as a curator of the Places & Spaces: Mapping Science exhibit.

Dr. Sayeed Choudhury, Associate Dean for Library Digital Programs and Hodson Director of the Digital Research and Curation Center, Sheridan Libraries, Johns Hopkins University. Dr. Choudhury holds graduate degrees in Civil Engineering and Systems Analysis and Economics from Johns Hopkins University. He is the Associate Dean for Library Digital Programs and Hodson Director of the Digital Research and Curation Center at the Sheridan Libraries of Johns Hopkins University. Dr. Choudhury is also the Director of Operations for the Institute of Data Intensive Engineering and Science (IDIES) at Johns Hopkins, a Lecturer in the Department of Computer Science at Johns Hopkins, a Research Fellow at the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign, and a Senior Presidential Fellow with the Council on Library and Information Resources.

Dr. Elaine Collier, Assistant Director for Clinical Research, Division for Clinical Research Resources, National Center for Research Resources, U.S National Institutes of Health.

Dr. Michael Conlon, Director of Data Infrastructure, University of Florida. Dr. Conlon received a Bachelor's degree from Bucknell University and a Ph.D. in Statistics from the University of Florida. He is Director of Data Infrastructure for the University of Florida. Dr. Conlon is also Research

Associate Professor of Biostatistics at the University of Florida and co-principal investigator of INVEST—the International Verapamil/Trandolapril Study, a randomized trial of 22,000 patients at 860 sites in 14 countries. The trial is conducted entirely online using Internet-based software designed and developed by a team under Dr. Conlon's direction. The software approach has been patented and licensed to MarCon Global Data Solutions, a company he co-founded along with Dr. Ronald Marks.

Dr. Peter Elias, Professor, Institute for Employment Research, University of Warwick. Dr. Elias received a Bachelor's degree in Chemistry before undertaking his doctoral studies in applied labor economics at the University of California, Berkeley. He is a Professor at the Institute for Employment Research, University of Warwick. He is a Fellow of the Royal Statistical Society, and has been the Strategic Advisor for Data Resources to the UK Economic and Social Research Council (ESRC) since 2004.

Dr. James Evans, Assistant Professor of Sociology, University of Chicago. Dr. Evans received a B.A. from Brigham Young University in 1994 and an M.A. and Ph.D. from Stanford University in 1999 and 2004, respectively. He is Assistant Professor of Sociology at the University of Chicago and a member of the Committee on the Conceptual and Historical Studies of Science. He is also a Fellow at the Computation Institute.

Dr. Martin Fenner, Hannover Medical School Cancer Center and ORCID Board of Directors. Dr. Fenner studied medicine in Berlin, and he did further training in basic and clinical oncology, including a postdoctoral fellowship in Boston. He now works at the Hannover Medical School Cancer Center. He writes the weblog Gobbledygook (http://blogs.plos.org/mfenner), is on the board of directors of ORCID (Open Researcher & Contributor ID; http://www.orcid.org/), and helps organize the Science Online London (http://www.scienceonlinelondon.org/) conference.

Dr. Chris Greer, Assistant Director for Information Technology Research and Development, U.S. Office of Science and Technology Policy. Dr. Greer received a Ph.D. in Biochemistry from the University of California, Berkeley, and did his postdoctoral work at the California Institute of Technology before teaching at the University of California, Irvine, in the Department of Biological Chemistry. Dr. Greer is Assistant Director for Information Technology Research and Development in the White House Office of Science and Technology Policy (OSTP). He also serves as co-chair of the Interagency Working Group on Digital Data, which has been charged by the Committee on Science of the National Science and Technology Council with developing and promoting implementation of strategic frameworks for digital scientific data preservation and access.

Dr. Tony Hey, Corporate Vice President, External Research Division, Microsoft Research. Dr. Hey received a Bachelor's degree in Physics and a Ph.D. in Theoretical Physics from Oxford University. He is corporate vice president of the External Research Division of Microsoft Research, where he is responsible for worldwide external research (ER) collaboration in Microsoft Research. Dr. Hey is a fellow of the U.K. Royal Academy of Engineering and has served on several national committees in the U.K., including committees of the U.K. Department of Trade and Industry and the Office of Science and Technology. He is a fellow of the British Computer Society, the Institute of Engineering and Technology, the Institute of Physics, and the U.S. American Association for the Advancement of Science (AAAS).

Dr. Haym Hirsh, Professor of Computer Science, Rutgers University. Dr. Hirsh received his B.S. in Mathematics-Computer Science from UCLA and his M.S. and Ph.D. in Computer Science from Stanford University. He is Professor and past-Chair of Computer Science at Rutgers University. From 2006 to 2010 he served as Director of the Division of Information and Intelligent Systems at the U.S. National Science Foundation, managing an organization responsible for over $500 million of research grants in Computer and Information Science and Engineering. He has also held visiting positions at Bar-Ilan University, CMU, MIT, and the University of Zurich.

Mr. Patrick Lambe, Adjunct Professor in Knowledge Management, Hong Kong Polytechnic University, and Principal Consultant, Straits Knowledge. Mr. Lambe studied at Oxford University. He is now based in Singapore and is an expert in knowledge management, knowledge organisation systems, and taxonomies. He has weblogs at http://www.greenchameleon.com and http://www.organisingknowledge.com. Patrick is currently working with the National Science Foundation Division of Science Resources Statistics on taxonomy management and development projects to support that division's mission.

Dr. Gerhard Lauer, Professor of German Literature, Georg-August-Universität Göttingen. Dr. Lauer received his Ph.D. in 1992. He is professor of German Literature at Georg-August-Universität Göttingen in Germany. Dr. Lauer is also co-editor of the *Journal of Literary Theory,* a member of the German Research Council commission for electronic publishing, an advisory board member of the Open Access Publishing in European Networks (OAPEN; http://www.oapen.org), a board member for the European Science Foundation Bibliometric Database Scoping Project (2008-2010), coordinator for German-Israeli academic exchange, and head of the TransCoop-programme committee of the Alexander von Humboldt-Stiftung.

Ms. Ruth Lee, Director, Research Councils UK Office in the United States. Ms. Lee received a B.A. from the University of Sheffield and a Master of Education from the University of Manchester. She is the Director of the U.S. office for the Research Councils UK, the primary public body in the UK charged with funding research and supporting the next generation of researchers.

Dr. David Lipman, Director, National Center for Biotechnology Information, National Library of Medicine, U.S. National Institutes of Health. Dr. Lipman earned a B.A. in Biology from Brown University in 1976 and an M.D. from the State University of New York, Buffalo, in 1980. He is the Director of the National Center for Biotechnology Information (NCBI), a division of the National Library of Medicine within the National Institutes of Health (NIH). He was appointed as NCBI's first Director in 1989, shortly after Congress created the Center in 1988, and has overseen its growth into one of the most heavily used resources in the world for the search and retrieval of biomedical information, with about 2 million users each day. He is a member of the National Academy of Sciences, the Institute of Medicine, and the American Academy of Arts and Sciences.

Mrs. Lucia Carvalho Pinto de Melo, President, Center for Strategic Studies and Management. Mrs. Carvalho received a Bachelor's degree in Chemical Engineering from Universidade Federal de Pernambuco in 1973; a Masters degree in Physics from the Universidade Federal de Pernambuco in 1976; a Masters degree in Energy and Environment at the University of California, Santa Barbara, in 1980; and a Masters degree from the Technology and Policy Program at the Massachusetts Institute of Technology in 1987. From 1990 to 1991 she led the Department of

Science and Technology of Pernambuco State, and she was President of the Foundation for Research of the State of Pernambuco (FACEPE) from 1995 to 1998. Mrs. Carvalho is currently the President of the Center for Strategic Studies and Management (CGEE).

Dr. August Muench, Astronomer, Harvard-Smithsonian Center for Astrophysics. Dr. Muench received a B.S. in Physics from the Georgia Institute of Technology in 1995 and a Ph.D. in Astronomy from the University of Florida in 2002. He is an astronomer at the Harvard-Smithsonian Center for Astrophysics and a member of the Seamless Astronomy team at Harvard, which is a collective of projects funded by NASA, NSF, and Microsoft Research working to better integrate astronomy's numerous open-access data repositories, comprehensive literature catalogues, and diverse software tools.

Dr. Theodore Papazoglou, Policy Analyst, European Research Council. Dr. Papazoglou received a B.S. in Physics from the University of Crete in 1985 and a Ph.D. in Biomedical Engineering from the University of Southern California in 1989. After completing his postdoctoral training at the Laser Research Center at Cedars Sinai Medical Center in Los Angeles, he attained tenure at the Institute of Electronic Structure and Lasers of the Foundation for Research and Technology Hellas (FO.R.T.H.-IESL). In 2001, Dr. Papazoglou was detached to the European Commission and worked as a scientific officer at the *Marie Curie Fellowships* programme. At the end of 2003 he was recruited as temporary agent in the Directorate General for Research of the European Commission. He is now working in Unit A1 (Support to the ERC Scientific Council) of the ERC Executive Agency, and implementation of the ERC's Open Access strategy is among his duties.

Dr. Hans Pfeiffenberger, Leader, IT Infrastructure Department, Alfred Wegener Institute for Polar and Marine Research. Dr. Pfeiffenberger leads the IT infrastructure department of the Alfred Wegener Institute. He is a speaker for the Helmholtz Open Access working group and advises the Knowledge Exchange (DFG, JISC, SURF, and DEFF). In 2008, Dr. Pfeiffenberger was appointed representative of the Helmholtz Association to the Alliance on Permanent Access (APA). He is also chief editor of the journal *Earth System Science Data,* an innovative journal providing quality assurance to published data through peer review.

Dr. Marcio de Miranda Santos, Executive Director, Centre for Strategic Management and Studies in Science, Technology and Innovation. Dr. Santos earned an M.S. in Genetics and Plant Breeding and a Ph.D. in Biochemical Genetics. Dr Santos is the current Executive Director of the Centre for Strategic Management and Studies in Science, Technology and Innovation and Chair of the Board of Trustees of the Center of Reference on Environmental Information (CRIA). He has been a consultant to the FAO (Food and Agriculture Organization of the United Nations), the Interamerican Institute for Cooperation in Agriculture, Bioversity International (formerly the International Plant Genetic Resources Institute), and national governments on policies for the conservation and use of plant genetic resources for food and agriculture. Dr. Santos has represented Brazil in the FAO/Commission on Genetic Resources for Food and Agriculture and in the UNEP/CBD Conference of the Parties (COP).

Dr. Lorenza Saracco, Research Programme Officer, Research Infrastructure Unit, European Commission. Dr. Saracco is a research programme officer within the Research Infrastructure Unit of the European Commission. She joined the EC in 2003, and since then she has followed policies and projects dealing with data and ICT infrastructures. Her background is in computer science

(her degree is from the University of Pisa, Italy), and before joining the Commission she worked for the Italian National Research Council (CNR), first as researcher in the area of conceptual modeling and management of Earth science data and then as policy officer dealing with the European Union research policies. As an EC officer she participated in various task forces and groups on data organisation and management.

Dr. Henry Sauermann, Assistant Professor, College of Management, Georgia Institute of Technology. Dr. Sauermann holds undergraduate degrees in Economics and Business Administration from the University of Potsdam, Germany, and a Ph.D. in Business Administration from Duke University. He is an Assistant Professor at the College of Management, Georgia Institute of Technology. One stream of Dr. Sauermann's research examines scientists' pecuniary and non-pecuniary motives and incentives and their effects on research productivity in industry as well as academia. A second stream of research focuses on scientific labor markets and on the career choices of junior scientists.

Dr. Bernard Schutz, Director, Max Planck Institute for Gravitational Physics, Albert Einstein Institute. Dr. Schutz earned a Ph.D. in physics from the California Institute of Technology. He is a Director of the Max Planck Institute for Gravitational Physics (Albert Einstein Institute) in Potsdam, Germany, and a member of the management team of the Max Planck Digital Library. He is also a Fellow of the American Physical Society and of the German Leopoldina Society, and he is a recipient of the Amaldi Gold Medal of the Italian Society for Gravitation. Dr. Schutz is active in science outreach and recently co-founded the Scienceface project (http://www.scienceface.org/), which is currently releasing a series of highly praised video interviews of leading scientists by a young interviewer with no scientific background; these are designed to make black-hole physics and astronomy accessible to young people.

Dr. Victoria Stodden, Assistant Professor, Department of Statistics, Columbia University. Dr. Stodden completed a Ph.D. in Statistics at Stanford University and a Masters in Legal Studies at Stanford Law School, where she created a new licensing structure for computational research. Previously, she had been a Postdoctoral Associate in Law and a Kauffman Fellow in Law at the Information Society Project at Yale Law School. She is an assistant professor of Statistics at Columbia University. She is currently co-chairing a working group on Communities and Virtual Organizations in the NSF Office of Cyberinfrastructure Task Force on Grand Challenge Communities. She is a Science Commons fellow, a member of the Sigma Xi scientific research society, and a member of the AAAS. Her website, which includes talks and publications, is http://www.stodden.net, and she occasionally blogs at http://blog.stodden.net.

Dr. Alex Szalay, Professor, Department of Physics and Astronomy, Johns Hopkins University. Dr. Szalay is the Alumni Centennial Professor of Astronomy and a professor in the Department of Computer Science at Johns Hopkins University. He is a cosmologist, working on the statistical measures of the spatial distribution of galaxies and galaxy formation. Dr. Szalay was born and educated in Hungary and has written over 450 papers in various scientific journals, covering areas from theoretical cosmology to observational astronomy, spatial statistics, and computer science. He is a Corresponding Member of the Hungarian Academy of Sciences and a Fellow of the American Academy of Arts and Sciences. Dr. Szalary received an Alexander Von Humboldt Award in Physical Sciences in 2004 and a Microsoft Award for Technical Computing in 2008.

Dr. Caitlin Trasande, Resident Scientist and Analyst, Nature Publishing Group. Dr. Trasande received a B.A. in Philosophy from St. John's College and a Ph.D. in Neuroscience from the University of Chicago in 2004. After serving as a postdoctoral scholar at Yale School of Medicine and Mount Sinai School of Medicine, she joined Nature Publishing Group (NPG) as a resident scientist and analyst in its technology unit, working on digital content management, special-content collections, and content search. From 2007 to 2009 she spearheaded development of a science metrics platform. In 2010 she and the science metrics project moved to a new digitally focused business unit within Macmillan (NPG's parent company), codename "Project Babbage."

Dr. Evelyne Viegas, Director, Microsoft Research. Dr. Viegas completed a Ph.D. in France. She is responsible for the Data Intelligence initiative at Microsoft Research in Redmond, WA. Before her present role, Dr. Viegas was a Technical Lead at Microsoft, delivering Natural Language Processing components to projects for MSN, Office, and Windows. Before Microsoft, and after completing her Ph.D. in France, she worked as a Principal Investigator at the Computing Research Laboratory in New Mexico on an ontology-based Machine Translation project.

Dr. Gert Wagner, Professor of Economics, Berlin University of Technology. Dr. Wagner received a Bachelor's degree in Economics in 1978 from the Johann Wolfgang Goethe University in Frankfurt, Germany, and a Ph.D. from the Berlin University of Technology in 1984. He is a professor of Economics at the Berlin University of Technology and a Max Planck Fellow at the MPI for Human Development in Berlin. Dr. Wagner is also Director of the German Socio-Economic Panel Study (SOEP) at DIW Berlin. He is chairman of the German Census Commission and German Council for Social and Economic Data, and he serves on the Advisory Board to Statistics Germany. He is a member of the Working Group on Social Sciences and Humanities of the European Strategy Forum for Research Infrastructures (ESFRI), the Founding Committee of the International Data Forum (IDF), and the Research Resources Board of ESRC/UK. He is research fellow of IZA and a Research Associate of CEPR. In 2007, Dr. Wagner was awarded the "Knight's Cross" of the Order of Merit of the Federal Republic of Germany.

Dr. John Wood, Professor, Imperial College London. Dr. Wood holds Doctoral degrees from Cambridge and Sheffield Universities. He is currently senior international relations adviser at Imperial College London. He will become the Secretary-General of the Association of Commonwealth Universities in July. Dr. Wood is a non-executive director of a number of companies, including Bio-Nano Consulting, and sits on the advisory board of the British Library. He is on the board of the Joint Information Services Committee, which is responsible for the UK academic computing network, and chairs its Support for Research Committee. He also chairs the European Commission's high-level group on the future management of scientific data. Dr. Wood was elected as a fellow of the Royal Academy of Engineering in 1999 and was made a Commander of the British Empire in 2007 for "services to science."

Workshop Participants from the U.S. National Science Foundation

Dr. Philip Bogden, Program Officer, Office of Cyberinfrastructure, U.S. National Science Foundation. Dr. Bogden holds a B.A. in Engineering and Applied Sciences from Harvard and a Ph.D. in Oceanography from Scripps Institution of Oceanography at the University of California, San Diego. He is currently the Program Director for the Office of Cyberinfrastructure at the

National Science Foundation. Dr. Bogden also holds a position as a Research Professor at the Center for Land-Sea Interactions at the University of New England. Before coming to the NSF, he was the CEO of GoMOOS, Inc., a private nonprofit organization with member institutions representing a broad array of stakeholders interested in ocean observations. GoMOOS manages a multi-institutional partnership that collects a wide variety of real-time ocean measurements. From 2003 to 2009, Dr. Bogden was also the Acting Director of the Southeastern Universities Research Association SURA Coastal Ocean Observing and Prediction (SCOOP) Program.

Dr. Myron Gutmann, Assistant Director, Directorate for Social, Behavioral & Economic Sciences, U.S. National Science Foundation. Dr. Gutmann holds a Ph.D. from Princeton University, and he is currently the Assistant Director of the Directorate for Social, Behavioral & Economic Sciences at the National Science Foundation. Before his appointment in 2009, Dr. Gutmann was the director of the Inter-University Consortium for Political and Social Research and a Research Professor at the Population Studies Center at the University of Michigan, where he was also a professor of history. He has a broad range of interests in interdisciplinary historical population studies relating population to agriculture, the environment, and health. He also studies ways that digital materials can be properly preserved and shared and how the confidentiality of research subjects can be protected when data about them is made available for secondary use.

Dr. Julia I. Lane, Program Director, Science of Science & Innovation Policy Program, U.S. National Science Foundation. Dr. Lane holds an undergraduate degree in Economics and Japanese from Massey University in New Zealand and an M.A. in Statistics and a Ph.D. in Economics from the University of Missouri-Columbia. She is the Program Director of the Science of Science & Innovation Policy program at the National Science Foundation. Her previous jobs included Senior Vice President and Director, Economics Department at NORC/University of Chicago; Director of the Employment Dynamics Program at the Urban Institute; Senior Research Fellow at the U.S. Census Bureau; and Assistant, Associate, and Full Professor at American University. Dr. Lane has organized over 30 national and international conferences, received several national awards, given keynote speeches all over the world, and served on a number of national and international advisory boards. She is one of the founders of the LEHD program at the Census Bureau, which is the first large-scale, linked employer-employee data set in the United States. She is also a fellow of the American Statistical Association.

Dr. Cora Marrett, Assistant Director, Directorate for Education and Human Resources, U.S. National Science Foundation. Dr. Marrett received a B.A. from Virginia Union University in 1963, an M.A. in 1965, and a Ph.D. in 1968 from the University of Wisconsin, Madison, all in Sociology. She served as University of Wisconsin's senior vice president for academic affairs for 6 years before coming to NSF. Before her appointment at the UW System, Dr. Marrett served as senior vice chancellor for academic affairs and provost at the University of Massachusetts-Amherst for 4 years. She was a member of the UW-Madison faculty from 1974 to 1997, with appointments in sociology and Afro-American studies. Dr. Marrett advanced from associate professor to full professor and was associate chairperson of the Department of Sociology (1988–1991). She was affiliated with the Energy Analysis and Policy Program and the Wisconsin Center for Education Research. She received an honorary doctorate from Wake Forest University in 1996, and she was elected a fellow of the American Academy of Arts and Sciences in 1998 and the American

Association for the Advancement of Science in 1996. She is widely published in the field of sociology and has held a number of public and professional service positions.

Dr. Edward H. Seidel, Acting Assistant Director, Directorate for Mathematical & Physical Sciences, U.S. National Science Foundation. Dr. Seidel earned a Ph.D. from Yale University in Relativistic Astrophysics. He is Acting Assistant Director of the Mathematical and Physical Sciences Directorate at the National Science Foundation. Dr. Seidel is a physicist recognized for his work on numerical relativity and black holes, as well as in high-performance and grid computing. In June 2008, the National Science Foundation selected Seidel as its director for the Office of Cyberinfrastructure (OCI). On 1 September 2008, he began this position, in which he oversees advances in supercomputing, high-speed networking, data storage and software development on a national level. He has recently assumed the role of Acting Assistant Director for Mathematics and Physical Sciences at NSF.

Workshop Observers

Dr. Stefano Bertuzzi, Office of Science Policy, U.S. National Institutes of Health. Dr. Bertuzzi received a Ph.D. in Molecular Biotechnology at the Catholic University of Milan, Italy, and after postdoctoral training in the Laboratory of Molecular Neurobiology at the Salk Institute in San Diego, became a tenured Associate Professor at the Dulbecco Telethon Institute in Milan, Italy. Dr. Bertuzzi is responsible for return-on-investment analyses in the Office of Science Policy, Office of the NIH Director, U.S. Department of Health and Human Services. In this position, Dr. Bertuzzi advises the NIH Director on a wide range of health-science policy matters. He is the recipient of several NIH Director's awards, along with other national and international awards.

Dr. Rachel Bruce, Innovation Director for Digital Infrastructure, JISC. Dr. Bruce is the Innovation Director for Digital Infrastructure. She oversees innovation programs and activities that are funded by the Support for Research committee and the Infrastructure and Resources committee. These include a number of programs, for example digital preservation, management of research data, and geospatial infrastructure and resources. She is concerned with the updating of infrastructure for the creation, sharing, and managing of digital resources and related shared services, as well as the policy and practices required to improve their reuse and exploitation to enhance education and research.

Ms. Sarah Colon, Research Associate, Japan Science & Technology Agency. Ms. Colon has an undergraduate degree in Biochemistry from Cornell and Master's degrees in Advanced Japanese Studies from Sheffield University and International Economics and Public Policy from the Johns Hopkins University School of Advanced International Studies (SAIS). She is a research associate with the Japan Science and Technology Agency, an independently administered sub-agency of Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT). She works at the liaison office in Washington, D.C., and follows and reports on U.S. science and technology trends for the headquarters in Tokyo.

Dr. Diane DiEuliis, Senior Policy Advisor, U.S. Office of Science and Technology Policy. Dr. DiEuliis received a Ph.D. in Biological Sciences at the University of Delaware. She then completed a research fellowship at the National Institutes of Health intramural research program in cellular neurobiology, focusing on the molecular and morphological features of neuronal cells. Following

her laboratory research, Dr. DiEuliis became a program director at the National Institute of Neurological Disorders and Stroke, where she began managing the Alzheimer's and Parkinson's disease portfolio of research grants and programs. She developed several strategic research plans for Parkinson's disease, coordinating with the Department of Defense and Veteran's Administration programs, which helped to expand and diversify the field of federal research on Parkinson's. She now maintains many of these planning programs annually, and manages the Udall Centers program. Dr. DiEuliis is also working as a senior policy advisor in the President's Office of Science and Technology Policy. Her policy focus is within the life sciences, and she is the staff director for several subcommittees within the Committee on Science, including research business models, human subjects' research, and the science of science policy.

Dr. Amy Friedlander, Senior Advisor, Directorate for Social, Behavioral and Economic Sciences, U.S. National Science Foundation. Dr. Friedlander graduated from Vassar College, where she was elected to Phi Beta Kappa. She holds an M.A. and Ph.D. from Emory University and an M.S.L.I.S. from The Catholic University of America. She works with the Assistant Director for SBE to coordinate a strategic planning exercise to articulate the driving questions in the SBE sciences for the year 2020 and beyond. She also helps the directorate to develop cooperative work within NSF (CISE, Engineering, and OCI) and with other federal agencies (e.g., NEH, NARA, and Library of Congress). Dr. Friedlander is also Editor-in-Chief of the *ACM Journal on Computing and Cultural Heritage.* Before joining NSF in June 2010, she was Director of Programs at the Council on Library and Information Resources.

Dr. Daniel Goroff, Senior Policy Analyst, U.S. Office of Science and Technology Policy. Dr. Goroff earned his B.A. and M.A. degrees in Mathematics from Harvard University as a Borden Scholar, an M.Phil. in Economics at Cambridge University as a Churchill Scholar, a Masters in Mathematical Finance at Boston University, and a Ph.D. in Mathematics at Princeton University as a Danforth Fellow. On loan to the White House Office of Science and Technology Policy (OSTP), Dr. Goroff is a Program Director at the Alfred P. Sloan Foundation, working on science, technology, education, and economics. He is currently on leave from Harvey Mudd College in Claremont, California, where he is Professor of Mathematics and Economics and where he previously served as Vice President for Academic Affairs and Dean of the Faculty.

Dr. Neil Jacobs, Acting Programme Director, Information Environment, JISC. Dr. Jacobs is Acting Programme Director for Digital Infrastructure (Information Environment). He oversees a variety of projects and programs in the areas of access to and management of digital resources, including linked data and digital repositories, scholarly communications, and research information management. These cover issues of technical interoperability, cultural and organizational change, sustainability, and business models.

Mr. Kei Koizumi, Assistant Director for Federal Research and Development, U.S. Office of Science and Technology Policy. Mr. Koizumi received his M.A. from the Center for International Science, Technology, and Public Policy program at George Washington University, and he received his B.A. in Political Science and Economics from Boston University. He is a Fellow of the American Association for the Advancement of Science. He joined OSTP in February 2009 after serving on the Obama transition team as part of the Technology, Innovation and Government Reform Policy Working Group. Before joining OSTP, Koizumi served as the longtime Director of the R&D Budget and Policy Program at the American Association for the Advancement of Science

(AAAS). He is known as a leading authority on federal science and technology funding and budget issues and is a frequent speaker to public groups and to the press.

Dr. Tanu Malik, Research Associate, Computation Institute, University of Chicago. Dr. Malik earned an undergraduate degree from the Department of Civil Engineering at Indian Institute of Technology, Kanpu, and an M.S. and Ph.D. in Computer Science from Johns Hopkins University. Dr. Malik is a Research Associate with the Computation Institute (CI) at the University of Chicago. Her research interests are in issues relating to building large-scale data-management systems such as federating distributed data systems, replicating large databases, data approximation, data provenance, and data quality. A recurrent theme in her research is to reexamine the core principles of database technology in the light of new requirements emerging from scientific data. Her research has resulted in some innovative database technology for handling large amounts of distributed scientific data.

Ms. Jeri Metzger Mulrow, Senior Mathematical Statistician, Division of Science Resources Statistics, U.S. National Science Foundation. Ms. Mulrow holds a B.S. in Mathematics from Montana State University and an M.S. in Statistics from Colorado State University. She is Senior Mathematical Statistician in the Division of Science Resources Statistics at the National Science Foundation. Ms. Mulrow is currently the project leader of the SRS Taxonomy Project, senior advisor to the SRS Early Career Doctorates Survey, and lead author of the State Chapter for Science and Engineering Indictors 2012. She was named a fellow of the American Statistical Association (ASA) in 2010, is a member of the ASA Board, a senior member of the American Society for Quality, and a member of the American Association of Public Opinion Research. As a statistician in the federal statistical system, she is particularly interested in data quality, data usability, data visualization, data access, data sharing, and the role that taxonomy plays in all of it.

Dr. James Onken, Special Assistant to the Acting Deputy Director for Extramural Research, U.S. National Institutes of Health. Dr. Onken received an M.S. and Ph.D. in Psychology from Northwestern University and an M.P.H. with a concentration in Biostatistics from George Washington University. He is responsible for analyzing and presenting data on NIH research programs and research personnel for use in program evaluation and policy studies. He is also program manager for the NIH Research Portfolio Online Reporting Tool (RePORT) website (http://RePORT.nih.gov/), the RePORT Expenditures and Results (RePORTER) system, and the companion ExPORTER site, where users can download databases of NIH-funded projects and publications and patents citing support from NIH.

Dr. Walter Schaffer, Research Training Officer, Extramural Research Training and Career Development Programs, U.S. National Institutes of Health. Dr. Walter Schaffer is the NIH Research Training Officer responsible for the extramural research training and career development programs. He received a Ph.D. in Chemistry from the University of Texas at San Antonio in 1978, with a dissertation on oxidative metabolism in rat brains. He then served as a Staff and Senior Staff Fellow in the Lab of Metabolism at the National Institute of Alcohol Abuse and Alcoholism. In 1986 Dr. Schaffer began a career as a Research Training Officer. He is a Captain in the U.S. Public Health Service Commissioned Corps.

Dr. Mya Sjogren, Performance and Accountability Analyst, Office of Research and Development, U.S. Environmental Protection Agency. Dr. Sjogren works in the Environmental Protection Agency (EPA) Office of Research and Development (ORD). She led the performance and accountability team, which reviews the performance for the agency's research and development programs. She directed the development of ORD's stakeholder surveys and has contributed to internal and external evaluation efforts such as bibliometric analysis and organizational scorecards. She facilitated the EPA-sponsored NAS study on Evaluating Research Efficiency, the Board of Scientific Counselor reviews, and the internal pilot that assesses which EPA research is cited in regulatory decisions, including rules, guidance, and records of decision.

Dr. Michael Stebbins, Assistant Director, Biotechnology, U.S. Office of Science and Technology Policy. Dr. Stebbins received his B.S. at SUNY Stony Brook and Ph.D. in Genetics while working at Cold Spring Harbor Laboratory. He is the Assistant Director for Biotechnology at the White House Office of Science and Technology Policy. Before joining OSTP he was the Director of Biology Policy for the Federation of American Scientists. He is a co-founder of Scientists and Engineers for America and a former Adjunct Professor of Bioethics at University of Pennsylvania. He has worked as a Legislative Fellow for Senator Harry Reid and on policy issues at the National Human Genome Research Institute.

Dr. George Strawn, Chief Information Officer, U.S. National Science Foundation. Dr. Strawn has an undergraduate degree from Cornell College and holds a Ph.D. in Mathematics from Iowa State. Since 1991 he has been at the National Science Foundation, where he is currently the Chief Information Officer (CIO). He was Director of the Directorate for Computer and Information Science and Engineering (CISE) Division of Advanced Networking Infrastructure and Research and the NSFNET Program Director. Before working at NSF, Dr. Strawn was a computer science faculty member at Iowa State University, where he also held several administrative positions. From 1986 to 1995 he served as Director of the ISU Computation Center. Under his leadership, ISU became a charter member of the regional NSFNET network, MIDnet, and ISU created a thousand-workstation academic system based on an extension of the MIT Athena system. From 1983 to 1986 he served as Chair of the ISU Computer Science Department. Dr. Strawn currently serves as co-chair of both the interagency Large Scale Networking Working Group and the international Coordinating Committee for Intercontinental Research Networks. He served as co-chair of the interagency Federal Networking Council from 1995 to 1997. Dr. Strawn also has held several positions in the computer industry and has worked as an information technology consultant in both private industry and government.

Dr. Edmund (Ned) Talley, Program Director for Channels, Synapses and Circuits, National Institute of Neurological Disorders and Stroke, U.S. National Institutes of Health. In 2001, Dr. Talley received his Ph.D. from the University of Virginia (UVA), studying the physiology and pharmacology of motor neurons involved in respiration. After his Ph.D., he remained at UVA as a Research Assistant Professor. He initiated investigations into the CNS functions of two-pore-domain potassium channels, with an emphasis on their modulation by neurotransmitters and clinically important drugs. Dr. Talley joined the NINDS in 2005 as a Program Director for Channels, Synapses and Circuits. His program at the NINDS is focused on basic research in synaptic transmission and neuromodulation.

Appendix D: List of Background Papers

Blue Ribbon Task Force on Sustainable Digital Preservation and Access, "Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information," National Science Foundation, February 2010, https://extwiki.nsf.gov/download/attachments/5799975/BRTF_Final_Report.pdf?version=1&modificationDate=1270471033207.

Choudhury, Sayeed, Benjamin Hobbs, and Mark Lorie, "A Framework for Evaluating Digital Library Services," *D-Lib Magazine* 8, July/August 2002, http://www.dlib.org/dlib/july02/choudhury/07choudhury.html.

European Union, "European Union Observes Data Protection Day," http://eurunion.org/emailcampaigns/preview.php?previewtype=html&nl=21&c=237&m=158&s=6b9351a6f2b8c87a0a5d1e223a2907d4.

European Union, "Riding the Wave: How Europe Can Gain from the Rising Tide of Scientific Data," a final report of the High Level Expert Group on Scientific Data, A submission to the European Commission, October 2010.

German Data Forum (RatSWD), "RatSWD Working Paper Series No. 150: Recommendations for Expanding the Research Infrastructure for the Social, Economic, and Behavioral Sciences," Federal Ministry of Education and Research, July 2010.

Greer, Chris, "NSTC Releases Strategy for Digital Scientific Data," 23 March 2009, http://www.dtic.mil/dtic/pdf/announcements/IWGDDRelease.doc&#124.

Hammond, Tony, Timo Hannay, and Ben Lund, "The Role of RSS in Science Publishing," *D-Lib Magazine* 10, December 2004, http://www.dlib.org/dlib/december04/hammond/12hammond.html.

Hsinchun Chen, Ronald N. Kostoff, Chaomei Chen, Jian Zhang, Michael S. Vogeley, Katy Börner, Nianli Ma, Russell J. Duhon, Angela Zoss, Venkat Srinivasan, Edward A. Fox, Christopher C. Yang, and Chih-Ping Wei, "AI and Global Science and Technology Assessment," *IEEE Intelligent Systems* 24, no. 4: 68–88, July/Aug. 2009, doi:10.1109/MIS.2009.68, http://www.computer.org/portal/web/csdl/doi?doc=doi/10.1109/MIS.2009.68.

Kahn, Robert E., and Jay Allen Sears, "A Brief Overview of the Digital Object Architecture," Corporation for National Initiatives, October 2003, https://extwiki.nsf.gov/download/attachments/5799975/OverviewDigitalObjectArchitecture.pdf?version=1&modificationDate=1270582345245.

Kahn, Robert E., and Robert Wilensky, "A Framework for Distributed Digital Object Services," *International Journal on Digital Libraries* 6 (2): 115–23, DOI 10.1007/s00799-005-0128-x, published online 13 March 2006, https://extwiki.nsf.gov/download/attachments/5799975/fulltext.pdf?version=1&modificationDate=1270582327918.

Lane, Julia, "Let's Make Science Metrics More Scientific," *Nature* 464: 488–89, 25 March 2010, doi: 10.1038/464488a, Published online 24 March 2010, http://www.nature.com/nature/journal/v464/n7288/full/464488a.html.

Lipman, David, "Ten Years of PubMed Central," *Columbia University Libraries Library News,* 26 January 2010, http://library.columbia.edu/news/exhibitions/2010/20100308_pubmed.html.

Mesirov, Jill P, "Accessible Reproducible Research," *Science* 327: 415–16, doi: 10.1126/science.1179653, 22 January 2010, https://extwiki.nsf.gov/download/attachments/5799975/Sciencemag+2009+01+23+Accessible+reproducible+research.pdf?version=1&modificationDate=1270470638792.

National Institutes of Health, "Research Portfolio Online Reporting Tools (RePORT)," U.S. Department of Health and Human Services, http://projectreporter.nih.gov/exporter/.

National Science Board, "Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century," National Science Foundation, September 2005.

Neylon, Cameron, "Science in the Open, An Openwetware Blog on the Challenges of Open and Connected Science: New Year, New Me," 23 January 2010, http://blog.openwetware.org/scienceintheopen.

Neylon, Cameron, and Shirley Wu, "Open Science: Tools, Approaches, and Implications," *Pacific Symposium on Biocomputing* 14: 540–44, 2009, http://psb.stanford.edu/psb-online/proceedings/psb09/workshop-opensci.pdf.

Office of Science and Technology Policy, "Harnessing the Power of Digital Data for Science and Society," Report of the Interagency Working Group on Digital Data to the National Science and Technology Council, January 2009, http://www.nitrd.gov/About/Harnessing_Power_Web.pdf.

Pullinger, John, and Gert G. Wagner, "RatSWD Working Paper Series No. 153: On the Respective Roles of National Libraries, National Archives and Research Data Centers in the Preservation of and Access to Research Data," German Data Forum, August 2010, https://extwiki.nsf.gov/download/attachments/5799975/RatSWD_WP_153.pdf?version=1&modificationDate=1289417297212.

Raddick, Jordan M., and Alexander S. Szalay, The Universe Online," *Science* 329: 1028–29, 27 August 2010, http://www.sciencemag.org/content/329/5995/1028.full.

Sauermann, Henry, and Roach, Michael, "The Price of Silence: Scientists: Trade-Offs between Pay and the Ability to Publish," 24 October 2010, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1696783.

Stodden, Victoria, "The Legal Framework for Reproducible Scientific Research: Licensing and Copyright," http://www.stanford.edu/~vcs/papers/LFRSR12012008.pdf.

UK Data Forum, "A UK National Strategy for Data Resources for Social and Economic Research 2009–2012," http://www2.warwick.ac.uk/fac/soc/ier/publications/2009/nds_publication_sep09.pdf.

Appendix E: List of White Papers from November 12, 2010, Workshop

Author/Title
Baker, Shenda, et al., "Data-Enabled Science in the Mathematical and Physical Sciences: Workshop Report"
Börner, Katy, "Briefing Document for Changing the Conduct of Science in the Information Age"
Conlon, Mike, "The Objects of Science and Their Representation in eScience"
Elias, Peter, "Digital Technology and the Conduct of Scientific Research"
European Union, "Riding the Wave: How Europe Can Gain from the Rising Tide of Scientific Data"
Evans, James, "Identification and the Complex System of Research"
Fenner, Martin, "White Paper for Changing the Conduct of Science in the Information Age"
Fenner, Martin, "Scientific Attribution Principles"
German Data Forum (RatSWD), "RatSWD Working Paper Series No. 150: Recommendations for Expanding the Research Infrastructure for the Social, Economic, and Behavioral Sciences"
Hey, Tony, "Open Access, Open Data, Open Science"
Hirsh, Haym, "How Do You Cite a Crowd?"
Lambe, Patrick, "Changing the Conduct of Science in the Information Age: Discussion Points"
Lauer, Gerhard, "Changing the Conduct of Science in the Information Age: Focusing on Sharing Knowledge and Data"
National Science Board, "Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century"
Office of Science and Technology Policy, "Harnessing the Power of Digital Data for Science and Society"
Papazoglou, Theodore, "IT-Based Approaches in Support of ERC's Mission to Support 'Frontier Research': First Experiences"
Pfeiffenberger, Hans, "Focusing on Social Constructs"
Sauermann, Henry, "Discussion Points for Session 3: Social Constructs; in Particular: Incentives"
Schutz, Bernard, "Data Access: Digital Technology and Scientific Communities"
Trasande, Caitlin, and Timo Hannay, "Changing the Conduct of Science: A Publisher's Perspective"
Viegas, Evelyne, "Data as an Enabler of Open Innovation: Challenges and Opportunities"

To access an extended version of this report that includes the workshop participants' white papers, go to: http://www.nsf.gov/publications/pub_summ.jsp?ods_key=oise11003.