Committee on

Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data

Committee Meeting #3
OPEN SESSION MATERIALS

May 6, 2019
Washington, DC

Keck Center
500 5th Street, NW
Washington, DC 20001
Room 206

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

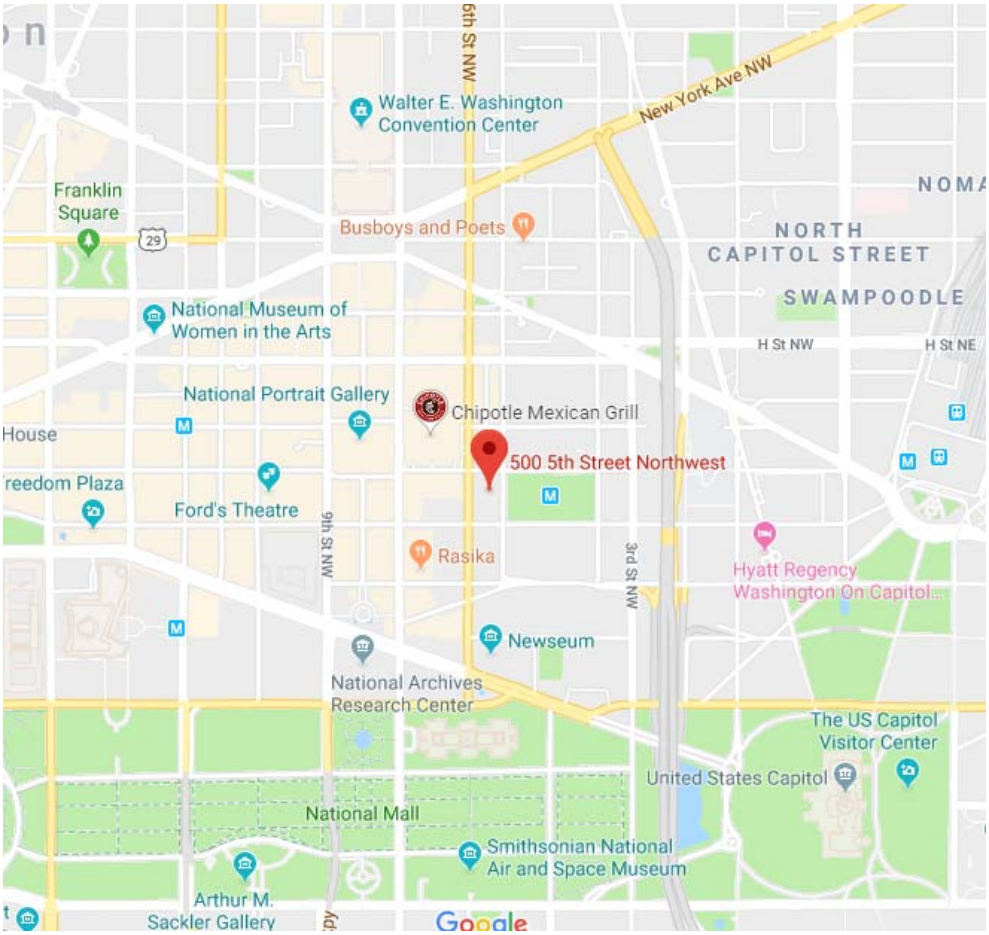
Directions to the National Academies Keck Center

Keck Center
500 5th Street NW
Washington, DC 20001

Directions

Parking: there is free parking in the Keck Center's underground lot. Access at the F street end of 6th street.

Public Transportation: The Keck Center is easily accessible by DC's Metro. The nearest Metro stations are Gallery Place/China Town (Red, Green, and Yellow lines) 2 blocks west and Judiciary Square (Green and Yellow lines) one block east.



Committee on Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data
Board on Mathematical Sciences and Analytics
National Academies of Sciences, Engineering, and Medicine

May 6, 2019
Keck Center of the National Academy of Sciences, Engineering, and Medicine
500 5th Street NW, Washington, DC 20001
Room 206

Open Agenda

Day 1 **May 6, 2019**

8:30 am –10:00 am
CLOSED SESSION—Committee and NAS Staff Only

10:00 am – 1:40 pm
OPEN SESSION DISCUSSION—DIGITAL DATA ARCHIVING DISRUPTORS

Open session remote login: https://nasem.zoom.us/j/471850063

10:00 Welcome, introductions, and statement of meeting objectives
David Chu, Committee Chair

10:05 Disruptors in Digital Archiving: Presentation from the U.S. National Archives
Leslie Johnston, Director of Digital Preservation, U.S. National Archives

Prompting Questions:
1) What models do you use to budget for data preservation?
2) How do you factor in unexpected cost or budget allocation fluctuations related to data preservation?
3) What disruptors have affected appraisal/reappraisal and redaction decisions, how?
4) How have those disruptions affected decisions regarding preservation of existing data? Planning for future data?
5) If you employ a cloud-based strategy, what happens if a cloud vendor's services are no longer available?
6) How do you think about format obsolescence?

11:05 Disruptors in the Cloud
Vamshidhar Kommineni, Principal Project Manager, Azure Blob Storage, Microsoft

Prompting questions:
1) What changes in technologies, data volumes & types, and data uses might appear in the next 5-10-25 years that would be disruptive to cost models and risk assessment for data preservation, archiving and access?
2) How do you forecast total cost of ownership of a cloud-based archive based archive over a 5-year life span? Over 10 years?
3) What specific steps does your organization take to prepare for any of these eventualities?

12:05 Lunch—available for purchase in the refectory

1:00 Indicators of data management costs at CERN
Simone Campana, Deputy Project Leader of the Worldwide Computing Grid

Prompting questions:

1) How does CERN determine what the lifespan of data saved?
2) CERN has long time lines and the data generating rate is reported to be 25 PetaB/year. How does CERN plan for storage costs? What is CERN's idea of a planning tool?
3) Zenodo - a general-purpose open-access repository - is run "as a marginal activity" - What does that imply for cost forecasting (e.g., how can CERN assume that it remains marginal)?
4) How has the archival infrastructure evolved at CERN? How do they expect it to evolve? How open is CERN about its forecasting assumptions?

1:40 Open session adjourns

2:00 pm – 8:00 pm **CLOSED SESSION—Committee and NAS Staff Only**

COMMITTEE INFORMATION

Committee on Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data

Statement of Task

A National Academies of Sciences, Engineering, and Medicine-appointed ad hoc committee will develop and demonstrate a framework for forecasting long-term costs for preserving, archiving, and accessing various types of biomedical data and estimating potential future benefits to research. In so doing, the committee will examine and evaluate the following considerations:

- Economic factors to be considered when examining the life-cycle cost for data sets (e.g., data acquisition, preservation, and dissemination);
- Cost consequences for various practices in accessioning and de-accessioning data sets;
- Economic factors to be considered in designating data sets as high value;
- Assumptions built in to the data collection and/or modeling processes;
- Anticipated technological disruptors and future developments in data science in a 5- to 10-year horizon; and
- Critical factors for successful adoption of data forecasting approaches by research and program management staff.

The committee will provide two case studies illustrating application of the framework to different biomedical contexts relevant to the National Library of Medicine's data resources. Relevant life-cycle costs will be delineated, as well as the assumptions underlying the models. To the extent practicable, the committee will identify strategies to communicate results and gain acceptance of the applicability of these models.

As part of its information gathering, the committee will plan and organize a 2-day workshop to gather input on the following topics:

- Tools and practices that NLM could use to better integrate risk management practices and considerations into data preservation, archiving, and accessing decisions;
- Methods to encourage NLM-funded researchers to consider, update, and track lifetime data costs (e.g., through data management plans and project renewals, or other interactions with the NIH); and
- Burdens on the academic researchers and industry staff to implement these tools, methods, and practices.

Speaker Biographies

Leslie Johnston is the Director of Digital Preservation for the National Archives and Records Administration (NARA), with responsibility to develop and execute a digital preservation strategy for the agency. Ms. Johnston has over 30 years of experience in information and data management in the cultural heritage, higher education, and federal communities including the Getty Museum, Stanford and Harvard University libraries, and the Library of Congress. Her expertise includes the design and implementation of digital content management and delivery systems and services, setting and applying content and metadata standards, and the preservation of heterogeneous born-digital and digitized collections. She is a frequent presenter internationally on these topics, and has published in several books and journals, as well as serving as a peer reviewer for conferences and as a grant reviewer for NSF, NEH, and IMLS. She has a B.A. in Anthropology and a M.A. in Archaeology, both from UCLA.

Vamshidhar Kommineni is a Principal PM Manager at Microsoft where he oversees a group of product managers focused on Azure Blob Storage, Azure's object storage service. In this role, he and his team are responsible for Azure's Object (Blob) Storage service including product strategy and growth of the business, working closely with customers, and delivering new features. Mr. Kommineni has been at Microsoft since 2003 in a variety of roles ranging from Software Engineer to Engineering Manager and Program Manager. He has been in the Azure team for nearly ten years, beginning in the "Project Red Dog" days when it was a small startup endeavor with less than 25 people. Prior to Azure, he was a developer in the Windows team working on reliability features in the core operating systems division.

Simone Campana is an applied physicist and senior staff member at CERN and currently the deputy project leader of the Worldwide LHC Computing Grid (WLCG). In this role ha is focusing on the long-term evolution of the software and the infrastructure in preparation for the High Luminosity LHC upgrade, with focus on Data Organisation, Management and Access. He obtained his PhD in Particle Physics at the University of California in 2003 and worked software and distributed computing projects since then. He was project leader of the Data Management system of the ATLAS experiment at CERN, the responsible for ATLAS distributed computing, the WLCG Service and Operations coordinator and the ATLAS Software and Computing Coordinator. He represented the ATLAS experiment and/or the WLCG project at the Resources Review Board, the LHC scientific Committee and the CERN Scientific Policy Committee, reporting directly to the CERN council.

Roster for the Committee on Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data

Chair:

Dr. David S.C. Chu, Chair
President
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 2231
Phone: (703) 845-2300
Email: dchu@ida.org

Members:

Dr. Ilkay Atlintas de Callafon
Chief Data Science Officer
San Diego Supercomputer Center, UC San Diego
9500 Gilman Drive
MS: 0505
La Jolla, CA 92093
Phone: (858) 822-5453
Email: ialtintas@ucsd.edu

Mr. Golam Sayeed Choudhury
Associate Dean for Research Data Management and Hodson Director of the Digital Research and Curation Center
John Hopkins
MSE Library
3400 N. Charles Street
Baltimore, MD 1218
Phone: (410) 516-4930
Email: sayeed@jhu.edu

Dr. Margaret Levenstein
Director
Inter-University Consortium for Political and Social Research
3478 Institute for Social Research
University of Michigan
Ann Arbor, MI 48106-1248
Phone: (734) 615-8400
Email: maggiel@umich.edu

Dr. Clifford Lynch
Executive Director
Coalition for Networked Information
21 Dupont Circle, NW
Suite 800
Washington, DC 20036
Phone: (202)296-0884
Email: clifford@cni.org

Dr. David Maier
Maseeh Professor of Emerging Technologies
Portland State University
PO Box 751
Portland, OR 97207-0751
Phone: (503) 725-2406
Email: maier@cs.pdx.edu

Dr. Charles Manski
Board of Trustees/ Professor in Economics
Northwestern University
2211 Campus Drive
Evanston, IL 60208
Phone: (847) 491- 8223
Email: cfmanski@northwestern.edu

Roster for the Committee on Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data

Dr. Maryann Martone
Professor Emerita
University of California, San Diego
9500 Gilman Drive #0608
La Jolla, CA 92093
Phone: (858) 822-0745
Email: mmartone@ucsd.edu

Dr. Alexa McCray
Professor of Medicine
Harvard Medical School
10 Shattuck Street
4th Floor
Boston, MA 02115
Phone: (617) 432-2144
Email: alexa_mccray@hms.harvard.edu

Dr. Michelle Meyer
Assistant Professor and Associate Director
Center for Translational Bioethics and
Health Care Policy, Geisinger
100 North Academy Av.
M-C 30-57
Danville, PA 17822
Phone: (571) 214-3380
Email: mmeyer@geisinger.edu

Dr. William Stead
Chief Strategy Officer and McKesson
Foundation Professor
Vanderbilt University Medical Center
1161 21st Avenue S.
D3300 Medical Center North
Nashville, TN 37232-2104
Phone: (615) 936-1424
Email: bill.stead@vumc.org

Dr. Lars Vilhuber
Executive Director
Labor Dynamics Institute, Cornell University
352 East Ives Hall
Ithaca, NY 14853
Phone: (607) 330-5743
Email: lars.vilhuber@cornll.edu

COMMITTEE MEMBERS ON FORECASTING COSTS FOR PRESERVING, ARCHIVING, AND PROMOTING ACCESS TO BIOMEDCIAL DATA

Chair

David S.C. Chu serves as president of the Institute for Defense Analyses. IDA is a non-profit corporation operating in the public interest. Its three federally funded research and development centers provide objective analyses of national security issues and related national challenges, particularly those requiring extraordinary scientific and technical expertise. As president, Dr. Chu directs the activities of more than 1,000 scientists and technologists. Together, they conduct and support research requested by federal agencies involved in advancing national security and advising on science and technology issues. Dr. Chu served in the Department of Defense as Under Secretary of Defense for Personnel and Readiness from 2001-2009, and earlier as Assistant Secretary of Defense and Director for Program Analysis and Evaluation from 1981-1993. From 1978-1981 he was the Assistant Director of the Congressional Budget Office for National Security and International Affairs. Dr. Chu served in the U. S. Army from 1968-1970. He was an economist with the RAND Corporation from 1970-1978, director of RAND's Washington Office from 1994-1998, and vice president for its Army Research Division from 1998-2001. He earned his doctorate in economics, as well as a bachelor of arts in economics and mathematics, from Yale University. Dr. Chu is a member of the Defense Science Board and a fellow of the National Academy of Public Administration. He is a recipient of the Department of Defense Medal for Distinguished Public Service with Gold Palm, the Department of Veterans Affairs Meritorious Service Award, the Department of the Army Distinguished Civilian Service Award, the Department of the Navy Distinguished Public Service Award, and the National Academy of Public Administration's National Public Service Award.

Members

Ilkay Altintaş de Callafon is the Chief Data Science Officer at the San Diego Supercomputer Center (SDSC), UC San Diego, where she is also the founder and director for the Workflows for Data Science Center of Excellence, and a fellow of the Halicioglu Data Science Institute (HDSI). In her various roles and projects, she leads collaborative multidisciplinary teams with a research objective to deliver impactful results through making computational data science work more reusable, programmable, scalable and reproducible. Since joining SDSC in 2001, she has been a principal investigator and a technical leader in a wide range of cross-disciplinary projects. Her work has been applied to many scientific and societal domains including bioinformatics, geoinformatics, high-energy physics, multi-scale biomedical science, smart cities, and smart manufacturing. She is a co-initiator of the popular open-source Kepler Scientific Workflow System and the co-author of publications related to computational data science at the intersection of workflows, provenance, distributed computing, big data, reproducibility, and software modeling in many different application areas. Among the awards she has received are the 2015 IEEE TCSC Award for Excellence in Scalable Computing for Early Career Researchers and the 2017 ACM SIGHPC Emerging Woman Leader in Technical Computing Award.

Golam Sayeed Choudhury is the associate dean for research data management and Hodson Director of the Digital Research and Curation Center at the Sheridan Libraries of Johns Hopkins University. Choudhury is also a member of the executive committee for the Institute of Data Intensive Engineering and Science (IDIES) based at Johns Hopkins. Choudhury is a President Obama appointee to the National Museum and Library Services Board. He was a member of the National Academies' Board on Research Data and Information and the Blue Ribbon Task Force on Sustainable Digital Preservation and Access. He has testified for the Research Subcommittee of the Congressional Committee on Science, Space, and Technology. He was a member of the board of the National Information Standards Organization, OpenAIRE2020, DuraSpace, the ICPSR Council, Digital Library Federation advisory committee, Library of Congress' National Digital Stewardship Alliance Coordinating Committee, Federation of Earth Scientists Information Partnership (ESIP) Executive Committee, and the Project MUSE Advisory Board. Choudhury was a member of the ECAR Data Curation Working Group. He has been a Senior Presidential Fellow with the Council on Library and Information Resources, a lecturer in the Department of Computer Science at

Johns Hopkins and a research fellow at the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign. He is the recipient of the 2012 OCLC/LITA Kilgour Award. Choudhury has served as principal investigator for projects funded through the National Science Foundation, Institute of Museum and Library Services, Library of Congress' NDIIPP, Alfred P. Sloan Foundation, Andrew W. Mellon Foundation, Microsoft Research, and a Maryland-based venture capital group. He is the product owner for the Data Conservancy, which focuses on the development of data curation infrastructure, and the Public Access Submission System, which supports simultaneous submission of articles to PubMedCentral and institutional repositories. He has oversight for data curation research and development and data archive implementation at the Sheridan Libraries at Johns Hopkins University. Choudhury has published articles in journals such as *the International Journal of Digital Curation*, *D-Lib*, the *Journal of Digital Information*, *First Monday*, and *Library Trends*. He has served on committees for the Digital Curation Conference, Open Repositories, Joint Conference on Digital Libraries, and Web-Wise. He has presented at various conferences including Educause, CNI, JISC-CNI, DLF, ALA, ACRL, and international venues including IFLA, the Kanazawa Information Technology Roundtable, eResearch Australasia, the North America-China Conference, eResearch New Zealand, and the Arabian-Gulf Chapter of the Special Libraries Conference.

Margaret Levenstein is director of ICPSR, the Inter-university Consortium for Political and Social Research; research professor at the Institute for Social Research and the School of Information; and adjunct professor of business economics and public policy at the Stephen M. Ross School of Business. She has taught economics at the University of Michigan since 1990. She serves as co-executive director of the Michigan Federal Statistical Research Data Center (FSRDC) and co-chair of the Executive Committee of the FSRDC national network. She is the associate chair of the American Economic Association's Committee on the Status of Women in the Economics Profession and past president of the Business History Conference. She is PI of CenHRS, a Sloan Foundation-funded project building an enhancement to the Health and Retirement Study based on linkages to administrative and survey data on HRS employers and co-workers. She is PI of an NSF-funded project to establish a repository of linked data and data linkage algorithms at ICPSR; a Sloan and NSF-funded effort to establish a Researcher Passport using open badges for credentialed, trusted researchers to access restricted data; and an NSF-funded project conducting experiments to encourage citizen-scientists to improve research metadata. She received a Ph.D. in economics from Yale University and a B.A. from Barnard College, Columbia University. She is the author of numerous studies on competition and collusion, the development of information systems, and using "organic" data to improve social and economic measurement. Her project using Tweets to predict unemployment is updated weekly at http://econprediction.eecs.umich.edu/ study. You can see her discuss her research on the impact of the 1930s Great Depression on innovative firms in the Midwest at http://www.youtube.com/watch?v=g8Ms7s-tPM4.

Clifford Lynch has been the executive director of the Coalition for Networked Information (CNI) since 1997. CNI, jointly sponsored by the Association of Research Libraries and EDUCAUSE, includes about 200 member organizations concerned with the intelligent uses of information technology and networked information to enhance scholarship and intellectual life. CNI's wide-ranging agenda includes work in digital preservation, data intensive scholarship, teaching, learning and technology, and infrastructure and standards development. Prior to joining CNI, Lynch spent 18 years at the University of California Office of the President, the last 10 as Director of Library Automation. Lynch, who holds a Ph.D. in computer science from the University of California, Berkeley, is an adjunct professor at Berkeley's School of Information. He is both a past president and recipient of the Award of Merit of the American Society for Information Science, and a fellow of the American Association for the Advancement of Science, the Association for Computing Machinery, and the National Information Standards Organization. He served as co-chair of the National Academies' Board on Research Data and Information from 2011-2016; he is active on numerous advisory boards and visiting committees. His work has been recognized by the American Library Association's Lippincott Award, the EDUCAUSE Leadership Award in Public Policy and Practice, and the American Society for Engineering Education's Homer Bernhardt Award.

COMMITTEE MEMBERS ON FORECASTING COSTS FOR PRESERVING, ARCHIVING, AND PROMOTING ACCESS TO BIOMEDCIAL DATA

David Maier is Maseeh Professor of Emerging Technologies at Portland State University. Prior to his current position, he was on the faculty at SUNY-Stony Brook and Oregon Graduate Institute. He has spent extended visits with INRIA, University of Wisconsin–Madison, Microsoft Research, and the National University of Singapore. He is the author of books on relational databases, logic programming, and object-oriented databases, as well as papers in database theory, object-oriented technology, scientific databases, and dataspace management. He is a recognized expert on the challenges of large-scale data in the sciences. He received an NSF Young Investigator Award in 1984 and was awarded the 1997 SIGMOD Innovations Award for his contributions in objects and databases. He is also an ACM Fellow and IEEE Senior Member. He holds a dual B.A. in mathematics and in computer science from the University of Oregon (Honors College, 1974) and a Ph.D. in electrical engineering and computer science from Princeton University (1978).

Charles Manski has been Board of Trustees Professor in Economics at Northwestern University since 1997. He previously was a faculty member at the University of Wisconsin-Madison (1983-1998), the Hebrew University of Jerusalem (1979-1983), and Carnegie Mellon University (1973-1980). He received his B.S. and Ph.D. in economics from M. I. T. in 1970 and 1973. He has received honorary doctorates from the University of Rome 'Tor Vergata' (2006) and the Hebrew University of Jerusalem (2018). Manski's research spans econometrics, judgment and decision, and analysis of public policy. He is author of *Public Policy in an Uncertain World* (Harvard 2013), *Identification for Prediction and Decision* (Harvard 2007), *Social Choice with Partial Knowledge of Treatment Response* (Princeton 2005), *Partial Identification of Probability Distributions* (Springer, 2003), *Identification Problems in the Social Sciences* (Harvard 1995), and *Analog Estimation Methods in Econometrics* (Chapman & Hall, 1988), co-author of *College Choice in America* (Harvard 1983), and co-editor of *Evaluating Welfare and Training Programs* (Harvard 1992) and *Structural Analysis of Discrete Data with Econometric Applications* (MIT 1981). He has served as director of the Institute for Research on Poverty (1988-1991), chair of the Board of Overseers of the Panel Study of Income Dynamics (1994-1998), and chair of the National Research Council Committee on Data and Research for Policy on Illegal Drugs (1998-2001). Editorial service includes terms as editor of the *Journal of Human Resources* (1991-1994), co-editor of the *Econometric Society Monograph Series* (1983-1988), member of the editorial board of the *Annual Review of Economics* (2007-2013), member of the Report Review Committee of the National Research Council (2010-2018), and associate editor of the *Annals of Applied Statistics* (2006-2010), *Econometrica*, (1980-1988), *Journal of Economic Perspectives* (1986-1989), *Journal of the American Statistical Association* (1983-1985, 2002-2004), and *Transportation Science* (1978-84). Manski is an elected member of the National Academy of Sciences. He is an elected fellow of the American Academy of Arts and Sciences, the Econometric Society, the American Statistical Association, and the American Association for the Advancement of Science, distinguished fellow of the American Economic Association, and corresponding fellow of the British Academy.

Maryann Martone is a professor emerita at UCSD, but still maintains an active laboratory and currently serves as the chair of the University of California Academic Senate Committee on Academic Computing and Communications. She received her B.A. from Wellesley College in biological psychology and ancient Greek and her Ph.D. in neuroscience from the University of California, San Diego. She started her career as a neuroanatomist, specializing in light and electron microscopy, but her main research for the past 15 years focused on informatics for neuroscience, i.e., neuroinformatics. She led the Neuroscience Information Framework (NIF), a national project to establish a uniform resource description framework for neuroscience, and the NIDDK Information Network (dknet), a portal for connecting researchers in digestive, kidney, and metabolic disease to data, tools, and materials. She just completed five years as editor-in-chief of *Brain and Behavior*, an open access journal, and has just launched a new journal as editor-in-chief, *NeuroCommons*, with BMC. Dr. Martone is past president of FORCE11, an organization dedicated to advancing scholarly communication and e-scholarship. She completed two years as the chair of the Council on Training, Science, and Infrastructure for the International Neuroinformatics

Coordinating Facility and is now the chair of the Governing Board. Since retiring, she served as the director of biological sciences for Hypothesis, a technology non-profit developing an open annotation layer for the web (2015-2018) and founded SciCrunch, a technology start up based on technologies developed by NIF and dkNET.

Alexa McCray is professor of medicine at Harvard Medical School and the Department of Medicine, Beth Israel Deaconess Medical Center. She conducts research on knowledge representation and discovery, with a special focus on the significant problems that persist in the curation, dissemination, and exchange of scientific and clinical information in biomedicine and health. McCray is the former director of the Lister Hill National Center for Biomedical Communications, a research division of the National Library of Medicine at the National Institutes of Health. While at the NIH, she directed the design and development of a number of national information resources, including ClinicalTrials.gov. Before joining the NIH, she was on the research staff of IBM's T.J. Watson Research Center. She received the Ph.D. from Georgetown University, and for three years was on the faculty there. She conducted pre-doctoral research at the Massachusetts Institute of Technology. McCray joined Harvard Medical School in 2005, where she was founding co-director of the Center for Biomedical Informatics and associate director of the Francis A. Countway Library of Medicine. McCray was elected to the National Academy of Medicine in 2001. She is chair of the National Research Council's Board on Research Data and Information. She is a fellow of the American Association for the Advancement of Science, a fellow of the American College of Medical Informatics (ACMI), an honorary fellow of the International Medical Informatics Association, and a founding fellow of the International Academy of Health Sciences Informatics. She is a past president of ACMI and a past member of the board of both the American Medical Informatics Association and the International Medical Informatics Association. She is a former editor-in-chief of *Methods of Information in Medicine*, and she is a past member of the editorial board of the *Journal of the American Medical Informatics Association*. She chaired the 2018 National Academies of Sciences, Engineering, and Medicine consensus study entitled Open Science by Design: Realizing a Vision for 21st Century Research.

Michelle Meyer is an assistant professor and associate director, research ethics, in the Center for Translational Bioethics and Health Care Policy at Geisinger, a large, integrated health system in Pennsylvania and New Jersey, where she chairs the IRB Leadership Committee and directs the Research Ethics Advice and Consultation Service. She is also faculty co-director of Geisinger's Applied Behavioral Insights Team (a.k.a. "nudge unit") in Geisinger's Steele Institute for Health Innovation. Her empirical and normative research focuses on judgment and decision making by patients, clinicians, research participants, and IRBs that has implications for law, ethics, or policy. She has served on the advisory board of the Social Science Genetic Association Consortium; the board of directors of the Open Humans Foundation (formerly PersonalGenomes.org); the Ethics & Compliance Advisory Board of PatientsLikeMe; the American Psychological Association's Commission on Ethics Processes; the ClinGen Working Group on Complex Diseases; an NAM/PCORI working group on generating stakeholder support and demand for health data sharing, linkage, and use; and a DARPA-funded technical exchange on complex social systems (TECSS). She developed a commissioned white paper addressing ethical issues raised by plans for developing a new data sharing institute. In most of those roles, she has focused on consent; data privacy; and data access and use, especially with respect to genomic data. Immediately before joining the faculty at Geisinger, Michelle was an assistant professor and director of bioethics policy in the Clarkson University–Icahn School of Medicine at Mount Sinai School of Medicine Bioethics Program and adjunct faculty at Albany Law School. Previously, she was an academic fellow at the Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics at Harvard Law School, a Greenwall Fellow in Bioethics and Health Policy at The Johns Hopkins and Georgetown Universities, and a research fellow at the John F. Kennedy School of Government at Harvard. She earned a Ph.D. in religious studies, with a focus on practical ethics, from the University of Virginia under the supervision of James F. Childress and a J.D. from Harvard Law School, where she was an editor of the *Harvard Law Review*.

Following law school, she clerked for Judge Stanley Marcus of the U.S. Court of Appeals for the Eleventh Circuit. She graduated summa cum laude from Dartmouth College.

William Stead is chief strategy officer for Vanderbilt University Medical Center (VUMC). In this capacity, he facilitates structured decision making to achieve strategic goals and concept development to nurture system innovation. Dr. Stead received his B.A., M.D., and residency training in internal medicine and nephrology from Duke University. He remained on Duke's faculty in nephrology as the physician in the physician-engineer partnership that developed The Medical Record (TMR), one of the first practical electronic medical record systems. He also helped Duke build one of the first patient-centered hospital information systems (IBM's PCS/ADS). He came to VUMC in 1991 and holds appointments as the McKesson Foundation Professor of Biomedical Informatics and Professor of Medicine. For two decades, he guided development of the Department of Biomedical Informatics and operational units providing information infrastructure to support health care, education, research programs of the Medical Center. He aligned organizational structure, informatics architecture, and change management to bring cutting-edge research in decision support, visualization, natural language processing, data mining, and data privacy into clinical practice. His current focus is on system-based care, learning and research leading toward personalized medicine, and population health management. Dr. Stead is a founding fellow of both the American College of Medical Informatics and the American Institute for Engineering in Biology and Medicine. He served as founding editor-in-chief of the *Journal of the American Medical Informatics Association*. His awards include the Collen Award for Excellence in Medical Informatics and the Lindberg Award for Innovation in Informatics. Most recently, the American Medical Informatics Association named the Award for Thought Leadership in Informatics in his honor. He served as president of the American College of Medical Informatics, chairman of the Board of Regents of the National Library of Medicine, presidential appointee to the Commission on Systemic Interoperability, chair of the National Research Council Committee on Engaging the Computer Science Research Community in Health Care Informatics, and co-chair of the Institute of Medicine Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records. He chairs the National Committee for Vital and Health Statistics (NCVHS) of the Department of Health and Human Services and the Technical Advisory Committee of the Center for Medical Interoperability. He is a member of the Council of the National Academy of Medicine, and the American Medical Association's Journal Oversight Committee.
In addition to his academic and advisory responsibilities, Dr. Stead is a director of HealthStream.

Lars Vilhuber is presently on the faculty of the Department of Economics at Cornell University, executive director of ILR's Labor Dynamics Institute, a senior research associate at the ILR School at Cornell University, Ithaca, and affiliated with the U.S. Census Bureau (Center for Economic Studies, CES). He holds a Ph.D. in economics from Université de Montréal, Montreal, Canada, having previously studied economics at the Universität Bonn, Germany, and Fernuniversität Hagen, Germany. He has worked in both academic and government research positions and continues to consult and collaborate with government and statistical agencies in Canada, the United States, and Europe. His research interests lie in the dynamics of the labor market: working with highly detailed longitudinally linked data, he has analyzed the effects and causes of mass layoffs, worker mobility, and the interaction between housing and the local labor market. Over the years, he has also gained extensive expertise on the data needs of economists and other social scientists, having been involved in the creation and maintenance of several data systems designed with analysis, publication, replicability, and maintenance of large-scale code bases in mind. His research in statistical disclosure limitation issues is a direct consequence of his profound interest in making data available in a multitude of formats to the broadest possible audience. His knowledge about various data enclave systems comes from both personal experience and the desire to improve the experience of others. He is data editor of the American Economic Association and managing editor of the *Journal of Privacy and Confidentiality*; chair of the Scientific Advisory Committee of the Centre d'accès sécurisé aux données (CASD) in France, senior advisor of the New York Federal Statistical Research Data Centers (NYRDC) in the United States. Dr. Vilhuber speaks English, German, and French fluently and can communicate effectively in Portuguese and Spanish.

He aligned organizational structure, informatics architecture, and change management to bring cutting-edge research in decision support, visualization, natural language processing, data mining, and data privacy into clinical practice. His current focus is on system-based care, learning and research leading toward personalized medicine, and population health management. Dr. Stead is a founding fellow of both the American College of Medical Informatics and the American Institute for Engineering in Biology and Medicine. He served as founding editor-in-chief of the *Journal of the American Medical Informatics Association*. His awards include the Collen Award for Excellence in Medical Informatics and the Lindberg Award for Innovation in Informatics. Most recently, the American Medical Informatics Association named the Award for Thought Leadership in Informatics in his honor. He served as president of the American College of Medical Informatics, chairman of the Board of Regents of the National Library of Medicine, presidential appointee to the Commission on Systemic Interoperability, chair of the National Research Council Committee on Engaging the Computer Science Research Community in Health Care Informatics, and co-chair of the Institute of Medicine Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records. He chairs the National Committee for Vital and Health Statistics (NCVHS) of the Department of Health and Human Services and the Technical Advisory Committee of the Center for Medical Interoperability. He is a member of the Council of the National Academy of Medicine, and the American Medical Association's Journal Oversight Committee. In addition to his academic and advisory responsibilities, Dr. Stead is a director of HealthStream.

Lars Vilhuber is presently on the faculty of the Department of Economics at Cornell University, executive director of ILR's Labor Dynamics Institute, a senior research associate at the ILR School at Cornell University, Ithaca, and affiliated with the U.S. Census Bureau (Center for Economic Studies, CES). He holds a Ph.D. in economics from Université de Montréal, Montreal, Canada, having previously studied economics at the Universität Bonn, Germany, and Fernuniversität Hagen, Germany. He has worked in both academic and government research positions and continues to consult and collaborate with government and statistical agencies in Canada, the United States, and Europe. His research interests lie in the dynamics of the labor market: working with highly detailed longitudinally linked data, he has analyzed the effects and causes of mass layoffs, worker mobility, and the interaction between housing and the local labor market. Over the years, he has also gained extensive expertise on the data needs of economists and other social scientists, having been involved in the creation and maintenance of several data systems designed with analysis, publication, replicability, and maintenance of large-scale code bases in mind. His research in statistical disclosure limitation issues is a direct consequence of his profound interest in making data available in a multitude of formats to the broadest possible audience. His knowledge about various data enclave systems comes from both personal experience and the desire to improve the experience of others. He is data editor of the American Economic Association and managing editor of the *Journal of Privacy and Confidentiality*; chair of the Scientific Advisory Committee of the Centre d'accès sécurisé aux données (CASD) in France, senior advisor of the New York Federal Statistical Research Data Centers (NYRDC) in the United States. Dr. Vilhuber speaks English, German, and French fluently and can communicate effectively in Portuguese and Spanish.