NATIONAL ACADEMIES Sciences Engineering Medicine

Accelerating and Deepening Approaches to FAIR Data Sharing: A Workshop
AGENDA

April 20, 2023 (Thursday)
All times are US Eastern Time
The public is invited to join virtually.

9:00 AM **Chair's Opening Remarks**
Sarah Nusser, Iowa State University and BRDI Chair

9:05 AM **After the Nelson Memo: New Priorities and Opportunities for Research Data Sharing**
Jerry Sheehan, Deputy Director for Policy and External Affairs, National Library of Medicine
Patricia Knezak, Program Scientist, National Aeronautics and Space Administration
Alan Tomkins, Deputy Director, Division of Social and Economic Sciences, National Science Foundation

9:30 AM **Redesigning Research Processes to Capitalize on FAIR**

Moderator: Alexa McCray, Professor of Medicine, Harvard University
Panelists: Daniel Atkins, Professor Emeritus of Information, University of Michigan
Carl Kesselman, Professor of Industrial and Systems Engineering, University of Southern California

10:15 AM Break

10:30 AM **Stakeholder Perspectives: The Current State of Research Data Sharing and FAIR**

Moderator: Ramanathan Guha, Google
Panelists: Vint Cerf, Vice President/Chief Internet Evangelist, Google
Erik Schultes, International Science Coordinator, GO FAIR
David Hayes, Lecturer, Stanford Law School

11:30 AM **Recent Initiatives and New Approaches to Promoting FAIR Research Practices**

Moderator: Martin Halbert, Special Advisor for Public Access, National Science Foundation
Panelists: Matthew Mayernik, Project Scientist, National Center for Atmospheric Research
Christine Kirkpatrick, Director, Research Data Services, San Diego Supercomputer Center
Ian Foster, Senior Scientist and Distinguished Fellow, Argonne National Laboratory

12:30 PM **Concluding Thoughts and Next Steps**

12:45 PM **Adjourn to Lunch**

NATIONAL ACADEMIES Sciences Engineering Medicine

Accelerating and Deepening Approaches to FAIR Data Sharing: A Workshop
April 20, 2023

Speaker Biographical Information

CHAIR'S OPENING REMARKS

SARAH NUSSER (Chair) is a professor emerita of statistics at Iowa State University. Her current research focuses on improving the reusability and impact of publicly accessible research data. Dr. Nusser is actively involved in national efforts to promote open science, transparency, and public access to research data. She is currently a senior fellow with the Association of American Universities (AAU)-and a research professor at University of Virginia. Dr. Nusser has held leadership roles in the AAU-Association of Public and Land-grant Universities Accelerating Public Access to Research Data Initiative since its inception in 2017 and is currently appointed to the NIST Research Data Framework Steering Committee. Dr. Nusser previously served as vice president for research (VPR) at Iowa State University, where she led efforts to advance the research mission and provided leadership for campus-wide programs for faculty research success, interdisciplinary initiatives, and ethical research practice. Prior to serving as VPR, Dr. Nusser served as the director of the Center for Survey Statistics and Methodology (CSSM) at Iowa State University, where she conducted research in survey statistics and methodology for land- and population-based surveys, including sampling and estimation for longitudinal natural resource surveys, measurement error models for dietary intake and physical activity data, and geospatial methods for address listing and for data collection in natural resource surveys. During her time with CSSM, Dr. Nusser directed statistical research and development for the annual USDA National Resources Inventory (NRI) survey program; the NRI integrates longitudinal geospatial survey data with administrative data and other geospatial resources to provide a trending database for agri-environmental issues on non-federal lands in the US. Dr. Nusser has served on numerous scientific panels, advisory committees and governing boards with the National Academies, federal agencies, the UN and statistical societies. She is a fellow of the American Statistical Association (ASA) and an elected member of the International Statistical Institute, and has received awards for excellence and service from Iowa State and the ASA. Dr. Nusser received a B.S. in botany from the University of Wisconsin-Madison, M.S. in botany from North Carolina State University, and Ph.D. in statistics from Iowa State University.

AFTER THE NELSON MEMO: NEW PRIORITIES AND OPPORTUNITIES FOR RESEARCH DATA SHARING

JERRY SHEEHAN is the Deputy Director of the National Library of Medicine (NLM), which is part of the National Institutes of Health, U.S. Department of Health and Human Services. Mr. Sheehan joined the NLM as Assistant Director for Policy Development in September 2006. From September 2015 to January 2017, Mr. Sheehan was on detail to the White House Office of Science and Technology Policy where he served as Assistant Director for Scientific Data and Information and helped advance open science policies across the Federal Government. Mr. Sheehan came to NLM from the Organization for Economic Cooperation and Development (OECD) in Paris, where he served as Principal Administrator and Senior Economist in the Science & Technology Policy Division from 2000 to 2006. Prior to joining the OECD, Mr. Sheehan held positions as a Senior Program Officer with the Computer Science and Telecommunications Board of the National Research Council and as an Analyst in the Congressional Office of Technology Assessment. Mr. Sheehan has been actively involved in international government discussions about

intellectual property rights, public access to the results of government-funded research and industry-science relationships. He serves as Chairman of the OECD Working Party on Innovation and Technology Policy, Co-chair of the Interagency Working Group on Open Science, and Vice President of the International Council for Scientific and Technical Information. He previously served as Chair of CENDI (the Federal STI Managers group). He holds B.S. (Electrical Engineering) and M.S. (Technology and Policy) degrees from the Massachusetts Institute of Technology.

PATRICIA KNEZEK is a Senior Advisor in the Mathematical and Physical Sciences Directorate at the National Science Foundation (NSF). She is currently on a three-year detail as a Program Scientist to the Astrophysics Division at National Aeronautics and Space Administration (NASA) Headquarters. She joined the NSF in March 2013. Prior to that she had been with the National Optical Astronomy Observatory as a staff scientist. Her research has focused on the impact of star formation on the evolution of galaxies. She has actively served the astronomical community in a number of ways, including being elected as a Councilor for the American Astronomical Society. Dr. Kzenek has been active in issues of diversity and inclusion for her entire career. She previously served on CSWA, where some of her activities have included leading the development of "Equity Now! The Pasadena Recommendations for Gender Equality in Astronomy" and launching (with Rachel Ivie of the American Institute of Physics) the ad hoc group that developed the Longitudinal Study of Astronomy Graduate Students. Currently, she is co-chair of the CSWA and also a member of the AAS Ethics Task Force.

ALAN TOMKINS is Deputy Division Director for Social and Economic Sciences (SES), Directorate of Social, Behavioral, and Economic Sciences (SBE), with the National Science Foundation (NSF). Previously, Dr. Tomkins was the founding director of the Public Policy Center (appointed in July 1998) and served as professor of psychology and law at the University of Nebraska-Lincoln. He lead projects that worked to advance methods and interdisciplinary theories for understanding trust in government; examining effective public participation and science communication; and understanding distrust and unauthorized online activities, such as hacking. In addition, he consulted with local, state, and federal government entities on participatory budgeting, strategic planning, performance measure options, workplace climate, resident input on policy decisions, and so on. Since 2007 he has served as co-editor of Court Review, journal of the American Judges Association, and prior to Court Review he served 12 years as Co-Editor and Editor for Behavioral Sciences & the Law. Dr. Tomkins received a B. A. degree from Boston University (1975) with a joint major in Psychology and Philosophy. He earned a J. D. and Ph.D. in Social Psychology from Washington University in St. Louis in 1984.

REDESIGNING RESEARCH PROCESSES TO CAPITALIZE ON FAIR

ALEXA MCCRAY (NAM) (Moderator) 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. Dr. McCray joined Harvard Medical School in 2005, where she co-founded the Center for Biomedical Informatics, now the Department of Biomedical Informatics. She currently serves as a Principal Investigator of the US-wide Undiagnosed Diseases Network, a National Institutes of Health (NIH) research study that seeks to provide answers for patients and families affected by undiagnosed conditions. Dr. McCray is the former director of the Lister Hill National Center for Biomedical Communications, an intramural research division of the National Library of Medicine at NIH. 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 conducted pre-doctoral research at the Massachusetts Institute of Technology. Dr. McCray was elected to the National Academy of Medicine in 2001. She is a fellow of the American Association for the Advancement of Science, the American College of Medical Informatics,

and the International Academy of Health Sciences Informatics. She is the immediate past chair of the National Academies of Sciences, Engineering, and Medicine's (NASEM) Board on Research Data and Information, and she chaired a 2018 NASEM consensus study entitled Open Science by Design: Realizing a Vision for 21st Century Research.

DANIEL E. ATKINS (NAE) is Emeritus W.K. Kellogg Professor of Information and Emeritus Professor of Electrical Engineering and Computer Science at the University of Michigan (UM), Ann Arbor. The first phase of his career, beginning at the University of Illinois, focused on computer architecture including high-speed arithmetic methods now widely used in modern computers, as well as the design and construction of experimental machines for image processing including pioneering work on computer-assisted tomography. The second phase of his career has focused on interdisciplinary research on cyberenabled distributed knowledge communities including collaboratories and digital libraries applied to both scientific research and education. He has served as UM Dean of Engineering, Founding Dean of the School of Information, and Associate VP for Research at UM, as well as the inaugural director of the Office of Cyberinfrastructure at the National Science Foundation (NSF). He chaired the Blue Ribbon Panel on Research Cyberinfrastructure for the NSF that became an international roadmap for initiatives on cyber-enabled research in the digital age. He has chaired or served on many advisory boards for government, academia, philanthropy, and industry. He recently chaired a study for the National Academies of Science, Engineering , and Medicine on the future of research workflow in the emerging age of AI and automation. Professor Atkins is a member of the National Academy of Engineering, a Fellow of the American Association for the Advanced of Science, and the recipient of numerous other honors and awards.

CARL KESSELMAN is a William H. Keck Professor of Engineering in the University of Southern California (USC) Viterbi School of Engineering Daniel J. Epstein Department of Industrial and Systems Engineering and a professor of Computer Science and Preventative Medicine at Keck School of Medicine. He is a USC Information Sciences Institute (ISI) Fellow, where he directs the Informatics Systems Research Division, and the Director of the Center of Excellence for Discovery Informatics in the Michelson Center for Convergent Biosciences. One of the fathers of grid computing and the GLOBUS open-source toolbox, now the de facto grid computing standard, Dr. Kesselman has received numerous honors for his pioneering research including the Lovelace Medal from the British Computing Society and the Goode Memorial Award from the IEEE Computing Society. He is a Fellow of the British Computing Society and the Association for Computing Machinary. Dr. Kesselman joined ISI in 1997 as a USC Computer Science Department research associate professor. He received his Ph.D. in Computer Science from the University of California at Los Angles, a Masters in Electrical Engineering from the University of Southern California and a Bachelors of Electrical Engineering from the State University of New York at Buffalo.

STAKEHOLDER PERSPECTIVES: THE CURRENT STATE OF RESEARCH DATA SHARING AND FAIR

RAMANATHAN GUHA (Moderator) is the founder and lead for DataCommons.org, a platform which synthesis a wide range of data sets into a single knowledge graph, for use by students and researchers. He is the creator of widely used web standards such as RSS, RDF and Schema.org, and products such as Google Custom Search. Co-founder of Epinions.com and Alpiri, he is currently a Google Fellow and Vice President at Google. He has a Ph.D.in Computer Science from Stanford University, M.S. from University of California, Berkeley and B.Tech.in Mechanical Engineering from IIT Madras.

VINTON G. CERF (NAS/NAE) is vice president and Chief Internet Evangelist for Google. He contributes to global policy development and continued spread of the Internet. Widely known as one of the "Fathers of the Internet," Dr. Cerf is the co-designer of the TCP/IP protocols and the architecture of the Internet. He has served in executive positions at MCI, the Corporation for National Research Initiatives and the

Defense Advanced Research Projects Agency and on the faculty of Stanford University. Dr. Cerf served as chairman of the board of the Internet Corporation for Assigned Names and Numbers (ICANN) from 2000-2007 and has been a Visiting Scientist at the Jet Propulsion Laboratory since 1998. Dr. Cerf is a Foreign Member of the British Royal Society and Swedish Academy of Engineering, and Fellow of IEEE, ACM, the American Association for the Advancement of Science, the American Academy of Arts and Sciences, the International Engineering Consortium, the Computer History Museum, the British Computer Society, the Worshipful Company of Information Technologists, and the Worshipful Company of Stationers. He currently serves as Past President of the Association for Computing Machinery, chairman of the American Registry for Internet Numbers (ARIN) and completed a term as Chairman of the Visiting Committee on Advanced Technology for the U.S. National Institute of Standards and Technology. President Obama appointed him to the National Science Board in 2012. Dr. Cerf is a recipient of numerous awards and commendations in connection with his work on the Internet, including the U.S. Presidential Medal of Freedom, U.S. National Medal of Technology, the Queen Elizabeth Prize for Engineering, the Prince of Asturias Award, the Tunisian National Medal of Science, the Japan Prize, the Charles Stark Draper award, the ACM Turing Award, Officer of the Legion d'Honneur and 29 honorary degrees.

ERIK SCHULTES is the scientific director at partners in FAIR, as well as the FAIR implementation lead at the GO FAIR Foundation. Before this, Dr. Schultes was international science coordinator at the GO FAIR International Support and Coordination Office. Dr. Schultes is co-author on the original publication of the FAIR Guiding Principles and has, since 2016, been working with a diverse, international community of stakeholders to develop FAIR data and services. This work included the development of the first accredited FAIR awareness training programs for executives and practical introductions to FAIR data stewardship for front-line data producers. Dr. Schultes is also co-developer of numerous FAIR maturity evaluation services and is the architect of the now widely adopted Three-Point FAIRification Framework (M4Ms, FIPs, and FDPs) that is systematically deployed by partners in FAIR. Along with Barbara Magagna, Dr. Schultes is currently co-chair of the FIP & Practice Working Group of the FAIR Digital Object Forum. Dr. Schultes is an evolutionary biologist with a data-intensive research focus and has held previous academic appointments at the University of California Los Angeles, The Whitehead Institute for Biomedical Research at the Massachusetts Institute of Technology, and the Santa Fe Institute. He currently holds an appointment at DTL Projects.

DAVID HAYES is a lecturer at Stanford Law School. Mr. Hayes has focused his career on energy, environmental and natural resources matters. He most recently served in the White House as Special Assistant to the President for Climate Policy. Prior to working for President Biden, Mr. Hayes was Executive Director of the State Energy & Environmental Impact Center at the NYU School of Law, where he worked with state attorneys general on climate, environment and clean energy initiatives. Mr. Hayes is a former Distinguished Visiting Lecturer at the Stanford Law School; a Fellow at Stanford University's Precourt Institute for Energy and Woods Institute for the Environment; the Senate-confirmed Deputy Secretary and Chief Operating Officer at the U.S. Department of the Interior for Presidents Barack Obama and Bill Clinton; and Chairman of the Board of the Environmental Law Institute. Between his stints in government, he was a partner and Global Chair of the Environment, Land and Resources Department at Latham & Watkins. Mr. Hayes is a graduate of the University of Notre Dame and Stanford Law School.

RECENT INITIATIVES AND NEW APPROACHES TO PROMOTING FAIR RESEARCH PRACTICES

MARTIN HALBERT (Moderator) is a Special Advisor for Public Access at the National Science Foundation (NSF). In this role he leads the programmatic activities of the National Science Foundation aimed at advancing the understanding and adoption of open science practices utilizing public access mechanisms,

and agency efforts to ensure that research products arising from NSF-funded projects are publicly accessible. Dr. Halbert has an interdisciplinary Ph.D. from Emory University. His research and primary areas of expertise include digital scholarship in the humanities, research library services, and building inter-institutional collaborative alliances for new research functions. Dr. Halbert has previously worked for Emory University, Rice University, The University of Texas at Austin, University of North Carolina at Greensboro, the University of North Texas, and the IBM Corporation. He has been principal investigator for grants and contracts totaling more than $7 million, funding more than a dozen large-scale collaborative projects among many educational institutions, with a special focus on information systems in digital humanities. He served as Dean of Libraries at both the University of North Carolina at Greensboro and the University of North Texas. He founded the Educopia Institute, an educational nonprofit that advances the well-being of libraries by fostering the advancement of shared information systems and infrastructures. Dr. Halbert led one of the founding projects of the U.S. National Digital Preservation Program beginning in the first years of the 21st Century.

MATTHEW MAYERNIK is a Project Scientist and Research Data Services Specialist in the NCAR (National Center for Atmospheric Research)/UCAR (University Corporation for Atmospheric Research) Library. Dr. Mayernik is also the Deputy Director of the NCAR Library. His work is focused on research and service development related to research data curation. His research interests include metadata practices and standards, data curation education, data citation and identity, and social and institutional aspects of research data. He is a member of the Committee on Open Environmental Information Services (COEIS) within the American Meteorological Society, and the Joint Editor-in-Chief of the Data Science Journal.

CHRISTINE KIRKPATRICK leads the San Diego Supercomputer Center's (SDSC) Research Data Services division, which manages large-scale infrastructure, networking, and services for research projects of regional and national scope. Her research is in data-centric AI, working at the intersection of ML and FAIR, with a focus on making AI more efficient to save on power consumption and 'time to science.' In addition to being PI of the EarthCube Office (ECO), Kirkpatrick founded the GO FAIR US Office, is PI of the West Big Data Innovation Hub, is on the Executive Committee for the Open Storage Network, and Co-PI of the NSF-funded Transboundary Groundwater Resiliency Research (TGRR) network. She serves as the Secretary General of the International Science Council's Committee on Data (CODATA), co-Chairs the FAIR Digital Object Forum, is on the Advisory Board for the Helmholtz Federated IT Services (HIFIS), and served on the National Academies of Sciences' U.S. National Committee for the Committee on Data.

IAN FOSTER is the Director of Argonne National Laboratory's Data Science and Learning Division, Argonne Senior Scientist and Distinguished Fellow and the Arthur Holly Compton Distinguished Service Professor of Computer Science at the University of Chicago. He was the Director of Argonne's Computation Institute from 2006 to 2016. Dr. Foster's research contributions span high-performance computing, distributed systems, and data-driven discovery. He has published hundreds of scientific papers and eight books on these and other topics. Methods and software developed under his leadership underpin many large national and international cyberinfrastructures. Dr. Foster received a BSc (Hons I) degree from the University of Canterbury, New Zealand, and a Ph.D. from Imperial College, United Kingdom, both in computer science. His awards include the Global Information Infrastructure (GII) Next Generation award, the British Computer Society's Lovelace Medal, R&D Magazine's Innovator of the Year, the IEEE Tsutomu Kanai award, and honorary doctorates from the University of Canterbury, New Zealand and CINVESTAV, Mexico. He is an elected Fellow of the American Association for the Advancement of Science, the Association for Computing Machinery, and British Computer Society.