NATIONAL ACADEMIES Sciences Engineering Medicine

Publishing in the Age of Open Science: Workshop for Researchers, Institutions, and Publishers | Day I

Location: 500 5th Street, NW, Washington, DC 20001–Keck 100 | Virtual: Zoom Link

WEDNESDAY, FEBRUARY 21ST, 2024 (ALL TIMES ARE US EASTERN STANDARD)

08:30 AM	**Optional: Early Morning Coffee** Provided by the Academies for registered in-person attendees
09:00 AM	**Welcome and Workshop Overview for Day I** Linda Nhon, Program Officer on the Board on Chemical Sciences and Technology, *National Academies of Sciences, Engineering, and Medicine*
09:10 AM	**Keynote | Advancing Public Access and Open Science Across the U.S. Government** Maryam Zaringhalam, Assistant Director for Public Access and Research Policy, *Office of Science Technology Policy*
09:40 AM	**Q&A with Maryam Zaringhalam** Moderated by Rob Maleczka, Professor of Chemistry, *Michigan State University*
09:55 AM	**Break**
10:00 AM	**Session One | Publishing Research Articles** Session Introduction and Overview Moderated by Rob Maleczka
10:05 AM	**Increasing Public Access to U.S. Department of Energy R&D Results – Publications and Open Science** Brian Hitson, Director, *Department of Energy - Office of Scientific and Technical Information*
10:20 AM	**Partnering the Chemical Science Community on the Road to Open Science** Emma Wilson, Director of Publishing, *Royal Society of Chemistry* (virtual)
10:35 AM	**Supporting the Chemical Community During the Transition to Open Science** Sarah Tegen, Senior Vice President and Chief Publishing Officer, *American Chemical Society*
10:50 AM	**Research Libraries: Facilitator and Funder** Elaine Westbrooks, Carl A. Kroch University Librarian and Vice Provost, *Cornell University*
11:05 AM	**Break**
11:20 AM	**Panel Discussion for Session I**

12:20 PM **Lunch**

01:30 PM **Session Two | Sustaining Data Repositories**
Session Introduction and Overview
Moderated by Jake Yeston, Editor at Science, *American Association for the Advancement of Science*

01:35 PM **FAIR Data Repositories: Expectations, Obligations, and Expenses**
Robert Hanisch, Director of Data and Informatics, *National Institute of Standards and Technology*

01:50 PM **FAIRies, Ghosts and Trolls: Data Challenges in the Age of AI**
Olaf Wiest**,** Grace-Rupley Professor of Chemistry and Biochemistry, *University of Notre Dame*; Director, *NSF Center for Computer Assisted Synthesis (C-CAS)*

02:05 PM **Protein Data Bank: From Two Epidemics to the Global Pandemic to mRNA Vaccines and Paxlovid**
Stephen K. Burley, Director, *RCSB Protein Data Bank*; University Professor and Henry Rutgers Chair, *Rutgers, The State University of New Jersey*; Research Scientist, Step IX, *University of California San Diego-San Diego Supercomputer Center*

02:20 PM **Making Data and Data Workflows FAIR and Pervasive with nanoHUB: Present and Future**
Alejandro (Ale) Strachan, Co-Director, *nanoHUB;* Reilly Professor of Materials Engineering, *Purdue University*

02:35 PM **Break**

02:50 PM **Panel Discussion Session II**

03:50 PM **Reflection on Sessions I and II**
Moderated by Robert Maleczka and Jake Yeston

04:30 PM **Adjourn | End of Day I**

NATIONAL ACADEMIES Sciences Engineering Medicine

Publishing in the Age of Open Science: Workshop for Researchers, Institutions, and Publishers | Day II

Location: 500 5th Street, NW, Washington, DC 20001–Keck 100 | Virtual: Zoom Link

THURSDAY, FEBRUARY 22ND, 2024 (ALL TIMES ARE US EASTERN STANDARD)

08:30 AM	**Optional: Early Morning Coffee** Provided by the Academies for registered in-person attendees
09:00 AM	**Welcome** Linda Nhon
09:05 AM	**Session Three | Supporting the Research Community** Session Introduction and Overview Moderated by Leah McEwen, Chemistry Librarian, *Cornell University*
09:15 AM	**What Did We Learn from the Sci-Hub Story?** Luis Sanchez, Associate Professor, *Niagara University*
09:30 AM	**An Industrial Perspective on Navigating OA and FAIR Data Practices** Danielle Schultz, Director of Discovery Process Chemistry, *Merck*
09:45 AM	**Translating Policies into Open Science and FAIR Data Practices in Chemistry: Incentivizing Researchers and Reducing Barriers** Ye Li, Librarian for Chemistry, Chemical Engineering, Materials Science and Engineering, *Massachusetts Institute of Technology*
10:00 AM	**Research Data Management and Sharing Support for Chemistry Researchers** Shannon Farrell, Research Data Services Lead and Director of Data Repository, *University of Minnesota*
10:15 AM	**Panel Discussion for Session III** Moderated by Leah McEwen
11:15 AM	**Break**
11:25 AM	**Reflection Session III and Closing Remarks** Leah McEwen, Rob Malezcka, and Jake Yeston
12:00 PM	**Adjourn | End of Day II**

Panelists' Biographies

Stephen K. Burley is an expert in data science and bioinformatics, structural biology, and structure-guided drug discovery for oncology. He is the Director of the RCSB Protein Data Bank (RCSB.org). Within Rutgers, The State University of New Jersey he serves as University Professor and Henry Rutgers Chair, Founding Director of the Institute for Quantitative Biomedicine, and Cancer Pharmacology Research Program Co-Leader within the Rutgers Cancer Institute of New Jersey. Burley's previous roles were Distinguished Lilly Research Scholar, Eli Lilly and Co.; Chief Scientific Officer and Senior Vice President for Research, SGX Pharmaceuticals, Inc.; Richard M. and Isabel P. Furlaud Professor, The Rockefeller University; and Investigator, Howard Hughes Medical Institute. His degrees include M.D. - Harvard Medical School; D.Phil. - Oxford University; and B.Sc. (physics) and Doctor of Science (Honoris causa) - Western University. Burley has published extensively in data science and bioinformatics, artificial intelligence/machine learning, structural biology, and clinical oncology.

Shannon Farrell is the Research Data Services Lead, Director of DRUM and Co-Director of AgEcon Search. Shannon received her BS in Ecology, Evolution and Behavior from the University of Minnesota, a MS in Environmental Studies from the University of Strathclyde in Glasgow, Scotland and her MS in Library and Information Science from the University of Illinois at Urbana-Champaign. Before becoming a librarian, she spent over a decade working on large-scale, data intensive research projects related to ecology, animal behavior, molecular biology, systematics, entomology, sustainable agriculture, genetically modified organisms, wildlife conservation, water quality, and invasive species. Shannon leads Minnesota University Libraries' Research Data Services Team, which is a group that is focused on campus-wide education and consultation around data management and data sharing. She is the Director of the Data Repository for the University of Minnesota (DRUM) (where she also serves as a scientific data curator) and Co-Director of AgEcon Search, an open access repository that houses scholarly research in the fields of agricultural and applied economics.

Robert J. Hanisch is the Director of the Office of Data and Informatics, Material Measurement Laboratory, at the National Institute of Standards and Technology in Gaithersburg, Maryland. He is responsible for improving data management and analysis practices and helping to assure compliance with national directives on open data access. Prior to coming to NIST in 2014, Dr. Hanisch was a Senior Scientist at the Space Telescope Science Institute, Baltimore, Maryland, and was the Director of the US Virtual Astronomical Observatory. For more than twenty-five years Dr. Hanisch led efforts in the astronomy community to improve the accessibility and interoperability of data archives and catalogs.

Brian A. Hitson is Director of the U.S. Department of Energy (DOE) Office of Scientific and Technical Information (OSTI) in the Office of Science. OSTI fulfills agency-wide responsibilities to collect, preserve, and disseminate scientific and technical information emanating from DOE research and development (R&D) activities. In support of guidance from the Office of Science and Technology Policy (OSTP), Brian co-authored both the 2014 and 2023 DOE Public Access Plans, developing strategies for increasing public access to scholarly publications and scientific data resulting from DOE's R&D efforts. Along with DOE, Office of Science, and OSTI colleagues, he led the launch of the DOE Public Access Gateway for Energy and Science, or DOE PAGES. He also formed partnerships with the National Science Foundation and the Department of Defense in increasing public access to those agencies' R&D results. As Director, he has led strategic efforts to improve discoverability and linkages between diverse, related research objects, including publications, datasets, and scientific software available at OSTI.GOV. He also represents DOE in

interagency and international open science initiatives and organizations. Brian holds a B.A. in Economics and an MBA, both from the University of Tennessee

Ye Li (https://orcid.org/0000-0001-8361-6916) is the Librarian for Chemistry, Chemical Engineering, Materials Science and Engineering at Massachusetts Institute of Technology (MIT) Libraries. Additionally, she services as a Trustee on the Board of Cambridge Crystallographic Data Centre. Ye earned her Ph.D. in Chemistry and M.A. in Library and Information Science from the University of Iowa in 2009. She then started her career as a chemistry librarian at the University of Michigan, and later led scholarly communications initiatives at the Colorado School of Mines from 2016 to 2018. Ye served as Chair 2023 and Program Chair of the Division of Chemical Information, American Chemical Society. Ye has also been active in the community of Force11 Scholarly Communications Institute, where she served as a member in the Program Committee and the Executive Committee as well as the Chair of the Archive Committee. Through research, teaching, and partnership with researchers and students, Ye has established her expertise in chemical information literacy, data management and sharing, as well as open and reproducible research and publishing. She is currently exploring librarians' roles in fostering computational access and use of data and texts, cultivating good data practices among chemists, as well as enabling FAIR data sharing. Ye has also been an active Carpentries instructor and supported local learning communities for basic programming skills and reproducible research workflows.

Luis Sanchez is an Associate Professor of Chemistry at Niagara University. His research interests focus on complex molecule synthesis and development of synthetic methods toward peptide and terpene-based natural products and analogues. Among other positions, he has served as Chair of the Western New York section of the American Chemical Society (2016), Member-at-large of the Executive Committee of the ACS Division of Organic Chemistry (2017–2019), and president of the National Chemistry Honors Society ΓΣΕ (2017–2019). Dr. Sanchez received his Bachelor of Science and Licentiate in Chemistry from the Pontifical Catholic University of Peru (2005) and his Ph.D. in Organic Chemistry from Michigan State University (2010). He was an NSF GOALI postdoctoral fellow at Merck/Michigan State University (2011) as well as a postdoctoral research associate at the University of Pennsylvania (2011-2013), before becoming a chemistry professor.

Danielle Schultz received her PhD from the University of Michigan, working with Professor John Wolfe, and was an NIH postdoctoral fellow at the University of Wisconsin-Madison with Professor Tehshik Yoon. Since joining Merck in 2014, Dani has been a member of Process Chemistry and Enabling Technologies in Rahway, NJ, and as of 2021, became the Director of the Discovery Process Chemistry group in Rahway, NJ, where she leads a group of process chemists in support of the Merck small molecule and peptide portfolio. Beyond her role at Merck, Dani has gained recognition worldwide for her scientific achievements and leadership, delivering over 20 invited lectureships. Her commitment to collaboration is evident through partnerships with seven esteemed academic groups including MacMillan, Schindler, and Sarlah. These collaborations have resulted in notable publications, with her expertise in academic-industrial collaborations being exemplified by a recent Nature Chemistry perspective on the topic. Recognizing the significance of industrial publications in advancing scientific knowledge, Dani actively advocates for their importance and has participated in several OA panels to share valuable insights on how industry might approach an OA future.

Alejandro Strachan is the Reilly Professor of Materials Engineering at Purdue University and the Co-Director of NSF's nanoHUB. Before joining Purdue, he was a Staff Member in the Theoretical Division of Los Alamos National Laboratory and worked as a Postdoctoral Scholar and Scientist at Caltech. He

received a Ph.D. in Physics from the University of Buenos Aires, Argentina. Prof. Strachan's research focuses on the development of predictive atomistic and multiscale models to describe materials from first principles and their combination with data science to address problems of technological or scientific importance. Areas of interest include high-energy density and active materials, metallic alloys for high-temperature applications, materials and devices for nanoelectronics and energy, as well as polymers and their composites. In addition, Strachan's scholarly work includes cyberinfrastructure to make simulations, models, and data widely accessible and useful for research and education. Prof. Strachan has published over 200 peer-reviewed scientific papers and his contributions to research and education have been recognized by several awards, including the Early Career Faculty Fellow Award from TMS in 2009, an R&D 100 award in the category of software and services for nanoHUB (2020), and the Reilly Chair Professorship in 2023.

Sarah Tegen is Senior Vice President and Chief Publishing Officer, ACS Publications, at the American Chemical Society. She leads the development of ACS's pre-eminent portfolio of eighty-five hybrid and open-access journals, oversees the award-winning news magazine Chemical & Engineering News, manages the development of the ACS books program, and heads strategic planning and relationship development in Asia. Dr. Tegen serves as Chair of the STM Association, STM's Research Integrity Governance Board, the American Heart Association's Scientific Publishing Committee, and the American Society for Microbiology's Publishing Committee. Tegen previously served as Co-Chair of STM's Society Day, Co-Chair of PSP's Journals Committee, and CSE President. Tegen has developed several PROSE-Award winning journals and platforms, is a frequent speaker on a range of industry topics. Dr. Tegen is an alumna of the University of California, Berkeley (PhD, Molecular and Cell Biology) and the Massachusetts Institute of Technology (BS, Biology).

Elaine L. Westbrooks received her bachelor's degree from the University of Pittsburgh in 1998, and her master's degree in library and information science from the University of Pittsburgh in 1999. She first came to Cornell University Library in 2000 as a metadata librarian at Albert R. Mann Library. In 2005, she became a senior metadata librarian, and she was selected to be the head of metadata services in 2006. In this position, Westbrooks developed and strengthened strategic alliances with other library units to position metadata services as an integral component of Cornell University Library's digital collections. In 2008, Westbrooks became the associate dean for library operations at the University of Nebraska-Lincoln. In this role, she provided strategic, managerial, staff development, and operational leadership for technical services, research and instruction services, the branch libraries, and the data curation program. In 2012, Westbrooks was selected to be the associate university librarian for research at the University of Michigan, where she coordinated and administered support for the university's research enterprise while providing operational leadership to its copyright office, Asia library, area studies, and most of its subject specialists. From 2017 to 2022, Westbrooks served as the university librarian and vice provost for University Libraries at the University of North Carolina at Chapel Hill where she was responsible for the University Libraries' annual budget of $45 million. An expert who has been interviewed by the Chronicle of Higher Education, Inside Higher Ed, and Vox, Westbrooks coauthored Metadata in Practice with Diane Hillmann in 2004 and coedited Academic Library Management: Case Studies with Tammy Nickelson Dearie and Michael Meth in 2017. She is also a much sought-out speaker on diversity, equity, and inclusion in libraries and other topics related to scholarly communications and leadership. In recognition of her acumen and her efforts to build strategic partnerships across borders, she was honored with the Foreign Expert Award by Fudan University in Shanghai, China, in 2015 and 2016. Westbrooks serves on the board of numerous organizations, including Sage Publishing, the Center

for Open Science, and the Digital Public Library of America. As the chief administrative officer of Cornell University Library, Westbrooks leads one of the world's most innovative research libraries and university presses, empowering scholarship at Cornell and beyond.

Olaf Wiest is the Grace-Rupley Professor of Chemistry and Biochemistry at the University of Notre Dame. He also directs the NSF Center for Computer Assisted Synthesis and was an Associate Editor for *J. Org, Chem*. until 2021. His research focuses on the elucidation of mechanisms using electronic structure theory, data chemistry and machine learning, and the development of synergistic predictive methods in synthetic organic chemistry that combine the Q2MM method he co-developed with experimental studies. The mechanistic insights and design principles derived from these studies are then applied to projects in computational biophysics and drug design for the treatment of infectious and rare diseases, specifically Niemann-Pick Type C, where his work formed the basis of a clinical trial. In 1993, he received his Dr. rer. nat. from the University of Bonn (Germany) and was Feodor Lynen Fellow at UCLA until 1995, after which he joined the faculty at Notre Dame. He is a Camille Dreyfus Teacher-Scholar, a Fellow of the AAAS and held visiting appointments at UCSF, DFCI, HKUST and PKUSZ.

Emma Wilson is the Director of Publishing at the Royal Society of Chemistry. She has strategic responsibility for the RSC's portfolio of journals, books, and chemistry databases, ensuring that the publishing portfolio delivers against the Society's mission to help the chemical sciences community make the world a better place. A particular focus is working with the chemistry community and other stakeholders to accelerate the transition to Open Access publishing models and supporting the development and adoption of open science practices. She has previously served on both the Open Access Scholarly Publishing Association (OASPA) and the Association of Learned and Professional Society Publishers (ALPSP) Boards. Prior to joining the RSC in 2006 she held a variety of editorial and business development roles at Elsevier and before that trained and worked as a biochemist (Sheffield University, Cambridge University and Sapienza University of Rome).

Maryam Zaringhalam is the Assistant Director for Public Access and Research Policy at the White House Office of Science and Technology Policy. There, she works to coordinate programs and policies advancing the Biden-Harris Administration's commitment to providing public access to data, publications, and the other important products of the nation's taxpayer-supported research and innovation enterprise. She comes to OSTP on detail from the National Library of Medicine (NLM) at the National Institutes of Health, where she is the Data Science and Open Science Officer in the Office of Strategic Initiatives. Dr. Zaringhalam joined NLM in 2017, where she works to enhance capacity across the biomedical research community for data science and open science, as well as promote diversity, equity, and inclusion among the research workforce. Dr. Zaringhalam received her Ph.D. in molecular biology from the Rockefeller University in 2017.

Moderators' Biographies

Michael Forster is a recently retired publishing industry executive with long experience leading and formulating strategy for both commercial and membership society organizations. In senior roles with Elsevier, Wiley, and IEEE he was responsible for journals, books, lab protocols, A&I services, databases, publishing technology, platform, and software development activities, living and working in the UK, the US, and Germany. Accomplishments include the deployment of searchable chemistry in online journal articles, automation of chemical taxonomy use, integration of software preservation in article publication, and launch of AI-based synthetic pathway planning. He is focused on ensuring that publishers understand the research workflow and provide tools and products that integrate with it successfully; for more than 10 years this has meant providing leadership around research reproducibility, data preservation, and open access. He has served as chair of CHORUS, with the board of CrossRef and GlobalSpec, and on committees of the STM Association and AAP-PSP, as well as engaging directly with the White House / OSTP on Open Access issues, and numerous organizing or presenting roles for NSF, NASEM, and other workshops and conferences.

Robert E. Maleczka Jr.is a Professor of Chemistry at Michigan State University (MSU). In 1995, he began his independent career at MSU as an assistant professor. He served as Department Chair from 2010 to 2021. The Maleczka group's research interests include the invention of "green" reactions for organic synthesis and the total synthesis of natural products. Honors bestowed on Dr. Maleczka include being named a Fellow of the American Chemical Society (ACS) and the American Association for the Advancement of Science (AAAS), the Merck Technology Collaboration Award, the Environmental Protection Agency's (EPA) 2008 Presidential Green Chemistry Challenge Award, the Astellas USA Foundation Faculty Award, the Novartis Lecturer title at Yale University, and Mentor of the Year award from the MSU Chapter of the National Society for the Professional Advancement of Black Chemists and Chemical Engineers. Among his service and outreach roles, he has served as a Diversity Scholar in the Center for the Integration of Research, Teaching, and Learning, as an invited participant ACS/EPA Green Chemistry Market Roundtable and White House Forum, as a member and Chair of the ACS Award for Affordable Green Chemistry Selection Committee, as an elected member of the Executive Committee of the ACS Division of Organic Chemistry. He currently Chairs the ACS Publications Committee and the C&EN Editorial Board. He earned his Ph.D. in 1992 under the direction of Leo Paquette.

Leah McEwen is the Chemistry Librarian at Cornell University where she has supported information discovery and scholarly communication in the chemical sciences since 1999. She is an active contributor to national and international chemical information initiatives, organizing dozens of thematic programs on research documentation and dissemination. She has served as both Secretary and Program Chair for the ACS Division of Chemical Information, and on the ACS Joint Board-Council Committees on Publishing, Ethics and Chemical Safety. She is a founding chair of the Research Data Alliance (RDA) Chemistry Research Data Interest Group (DIG Chemistry) and a board member of the International Chemical Identifier - InChI Trust. She currently serves as chair of the IUPAC Committee on Publications and Cheminformatics Data Standards to facilitate design and implementation of digital standards. She is lead on the Chemistry case study for the CODATA-RDA WorldFAIR Initiative to advance FAIR data practices and is an international advisor for the NFDI4Chem research data infrastructure consortium in Germany. She holds master's degrees in Nutritional Biochemistry from Cornell University and Library and Information Science from Emporia State University and was the first Paul Otlet - Eugene Garfield Fellow at the Science History Institute.

Jake Yeston joined the staff at Science Magazine in 2004, with responsibilities to edit and coordinate review for original research submissions in chemistry and overlapping segments of biochemistry and applied physics. In his current role as Editor at the journal, he now oversees the editorial group handling all research content in the physical sciences. Previously, he conducted postdoctoral research in ultrafast spectroscopy at the Max Planck Institute for Quantum Optics in Garching, Germany on a Humboldt fellowship (under Karl Kompa), and then worked as a National Research Council fellow at the National Institute of Standards and Technology in Gaithersburg, Maryland (under Edwin Heilweil). Dr. Yeston earned his Ph.D. in chemistry from the University of California, Berkeley in 2001, with a focus at the interface of synthetic organometallic chemistry and vibrational spectroscopy (working jointly under Robert Bergman and C. Bradley Moore).

Panelists

Stephen Burley

Shannon Farrell

Brian Hanisch

Brian Hitson

Ye Li

Luis Sanchez

Dani Schultz

Ale Strachan

Sarah Tegen

Elaine Westbrooks

Olaf Wiest

Emma Wilson

Maryam Zaringhalam

NATIONAL ACADEMIES Sciences Engineering Medicine

Moderators

Michael Forster

Leah McEwen

Rob Maleczka

Jake Yeston