NATIONAL ACADEMIES Sciences Engineering Medicine

Workshops to Support EPA's Development of Human Health Assessments: Artificial Intelligence and Open Data Practices in Chemical Hazard Assessment

May 25th and 26th, 2022

Participant Biosketches

Session I. AI and Data Science Applications: Promises and Prospects

Daniel E. Ho is the William Benjamin Scott and Luna M. Scott Professor of Law at Stanford Law School, Professor of Political Science, and Senior Fellow at the Stanford Institute for Economic Policy Research. He is also Associate Director of the Stanford Institute for Human-Centered Artificial Intelligence, Faculty Fellow at the Center for Advanced Study in the Behavioral Sciences, and is Director of the Regulation, Evaluation, and Governance Lab (RegLab). Ho serves on the National Artificial Intelligence Advisory Commission (NAIAC), advising the White House on artificial intelligence, and as a Public Member of the Administrative Conference of the United States (ACUS). He received his J.D. from Yale Law School and Ph.D. from Harvard University and clerked for Judge Stephen F. Williams on the U.S. Court of Appeals, District of Columbia Circuit.

Christopher Mungall is a Staff Scientist in the Environmental Genomics and Systems Biology Division at LBNL, where he heads the Biosystems Data Science department. Chris's research interests center around the capture, computational integration, and dissemination of biological research data, and the development of methods for using this data to elucidate biological mechanisms underpinning the health of humans and of the planet.
*bio via https://biosciences.lbl.gov/profiles/chris-mungall/

Nicole Kleinstreuer, PhD., is the acting director of the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), the US federal resource for alternatives to animal testing. At NICEATM, she leads domestic and international efforts to develop novel testing and analysis strategies that provide more rapid, mechanistic, and human-relevant predictions of potential environmental chemical hazards. Kleinstreuer's research focuses on mathematical and computational modeling of biological systems and their susceptibility to perturbations that result in adverse health outcomes. She has a secondary appointment in the NIEHS Division of Intramural Research Biostatistics and Computational Biology Branch, and adjunct faculty positions in the Yale University School of Public Health and the Eshelman

School of Pharmacy at the University of North Carolina at Chapel Hill. She has published over 100 peer-reviewed publications and won numerous awards, including the 2019 Society of Toxicology Achievement Award, the 2016 Young Researcher Americas Award from the Lush Prize Foundation, the 2012 US EPA Office of Research and Development Impact Award, and the 2008 B.H. Neumann Prize from the Australian Mathematical Society.

Session II. Addressing Challenges for Applying Systematic Review Methods Using AI

Malcolm Macleod is Professor of Neurology and Translational Neurosciences and Academic Lead for Research Improvement and Research Integrity at the University of Edinburgh. With Howells he co-founded the CAMARADES collaboration in 2005. He was Academic Coordinator of the European Quality in Preclinical Data IMI consortium and now leads the Guarantors of EQIPD. He was a lead author in the Lancet Series on Research Waste in 2014, and is a member of the UK Reproducibility Network steering committee. His current research interests include providing evidence for the effectiveness (or not) of strategies which might be adopted by funders, journals and institutions to improve the quality of their research; and developments in the automation of systematic review processes.

Rens van de Schoot works as full professor 'Statistics for Small Data Sets' at Utrecht University in the Netherlands and as extra-ordinary professor North-West University in South-Africa. He is also program director of the research master 'Methodology and Statistics for the Behavioural, Biomedical and Social Sciences' and coordinator of the post-graduate program at the department of Methods and Statistics. His main research projects are the open-source ASReview project (Active learning for systematic text reviewing) and solutions for small data sets (S4) in the field of structural equation modeling with solutions in the areas of expert elicitation, Bayesian statistics and constrained statistical inference. Rens obtained his PhD cum laude on the topic of applying Bayesian statistics to real-life data at the Methodology and Statistics department at Utrecht University, The Netherlands.
*bio via https://www.rensvandeschoot.com/about-rens

Karen A. Robinson, PhD is a Professor in the Department of Medicine at the Johns Hopkins University School of Medicine with joint appointments in the Departments of Epidemiology and Health Policy & Management at the university's Bloomberg School of Public Health, and with an affiliate appointment with the Western Norway University of Applied Sciences.

Dr. Robinson is Director of the Johns Hopkins University Evidence-based Practice Center and, within the EPC Program, serves as an Associate Editor and on the Methods Steering Committee. Dr. Robinson conducts systematic reviews and research on the use of evidence in making decisions. For over 20 years she has been an active member of Cochrane and is currently an editor for two review groups (including the

methodology review group).

Dr. Robinson was a founding member of the International Collaboration for the Automation of Systematic Reviews (ICASR) and served as chair of the governance board for the Systematic Review Data Repository (SRDR). Within the Guidelines International Network (G-I-N) she serves on the steering committees for two groups (Tech; North America). She serves as a founding steering group member of the Evidence-Based Research Network. Dr. Robinson has been active in making evidence computable, including working with the COVID-9 Knowledge Accelerator (COKA) and serving on the board of the Scientific Knowledge Accelerator Foundation (SKAF).

In addition to being an invited participant in NASEM workshops, Dr. Robinson has served on 5 National Academies' Committees assessing the use of systematic review methodologies by EPA and others (Endocrine- Related Low-Dose Toxicity; the Committee to Review Advances Made to the IRIS Process; the Committee to Review DOD's Approach to Deriving an Occupational Exposure Level for Trichloroethylene; the Committee to Review EPA's IRIS Assessment Plan for Inorganic Arsenic; and the Committee to Review the Use of Systematic Review in EPA's Toxic Substances Control Act Risk Evaluations).

Olwenn Martin, UCL
Building on an academic interdisciplinary background in both Natural Sciences (Chemistry, Environmental Sciences) and Social Sciences (Environmental Decision Making, Development Management, Environmental Policy), Olwenn's expertise lies in the integration and translation of both fundamental and observational scientific evidence into policy. Her research focuses on mismanaged emissions of chemicals from the technosphere (processes, materials and products) to the geosphere, hydrosphere, atmosphere and ultimately biosphere (our living planet). She was one of the earliest adopters of applying the principles of evidence-based medicine to environmental health and toxicology, including systematic review and meta-analysis and socio-economic analyses. Issues of specific interest include endocrine disruption and mixture effects. Olwenn provides scientific expertise to the UK Health and Safety Executive, the UK Food Standards Agency, is a member of the OECD Issue Team on Sustainable Chemistry and represents the European Parliament on the European Chemical Agency's management board. She is also an active member of the Planetary Health Alliance and French-speaking Alliance Santé Planétaire.

Session III. Optimizing Data Extraction for Evidence Synthesis and High Level Decision-Making

Weida Tong is Director of the Division of Bioinformatics and Biostatistics at FDA's National Center for Toxicological Research (NCTR). His role is to apply bioinformatics, Artificial Intelligence (AI), molecular modeling and data analytics for biomarker

discovery, drug safety and repurposing, pharmacogenomics, toxicogenomics, and precision medicine. He has been:

1. Developing machine learning and AI for digital health and drug repositioning;

2. Supervising and leading an FDA-led community wide consortium to analyze technical performance and utility of emerging genomics technologies with an emphasis on regulatory science and precision medicine;

3. Developing Liver Toxicity Knowledge Base to address drug safety concerns related to drug-induced liver injury;

4. Designing and developing computer-based technology to support FDA's effort on bioinformatics and scientific computing; and

5. Conducting molecular modeling and structure-activity relationships on various toxicological endpoints, such as endocrine disruption and carcinogenicity.

Dr. Tong has published over 300 peer-reviewed papers and book chapters.

Jason Fries is a scientist at Stanford University's Center for Biomedical Informatics Research. His research focuses on methods that enable experts to rapidly build machine learning models in domains such as medicine, where obtaining large-scale, expert-labeled data is a significant challenge. His interests include weak supervision, zero and few-shot learning, and multimodal representation learning. He previously completed his postdoc in computer science at Stanford where he collaborated on development of the weakly supervised framework Snorkel and its applications in healthcare.

Dan Sanders received a Bachelor's degree in Polymer Science and Engineering in 1996 and Master's degree in Macromolecular Science in 1999, both from Case Western Reserve University (Cleveland, OH). He went on to receive his Ph.D. in Chemistry from the California Institute of Technology (Pasadena, CA) working in the lab of Prof. Robert H. Grubbs. In 2004, he joined IBM's Almaden Research Center as a Research Staff Member and is currently, a Principal Research Staff Member and Senior Manager in charge of the Materials Discovery department at IBM Research – Almaden. As part of IBM Research's Accelerated Discovery strategy, Dan's organization seeks to apply informatics, simulation, Artificial Intelligence, and advanced lab automation to speed the discovery of new sustainable materials and apply these techniques to pertinent use cases of relevance for IBM and the world. Currently, his department is working on advanced battery materials, CO2 upcycling into novel monomers and polymers, and sustainable materials for semiconductor fabrication.

Marianthi-Anna Kioumourtzoglou is an environmental engineer and epidemiologist. She holds a Master of Science in Public Health (MSPH) from the Environmental Sciences and Engineering Department at the University of North Carolina at Chapel Hill and a Doctor of Science (ScD) in Environmental Health from the Harvard TH Chan School of Public Health. She is currently an Assistant Professor at the Department of Environmental Health Sciences at Columbia University's Mailman School of Public Health. Her research focuses on applied statistical issues related to environmental epidemiology, including quantifying and correcting for exposure measurement error, exposure prediction uncertainty propagation, and assessment of high-dimensional and complex exposures in health analyses. Her studies focus on air pollution and climate-related exposures and, additionally, on identifying vulnerable sub-populations and characterizing how risks may vary across neighborhood-level and other urban characteristics.

Session IV. Poster Presentations

Zach Calhoun is a first year PhD student at Duke University, in the department of Civil and Environmental Engineering, where he is co-advised by Michael Bergin and David Carlson. Prior to starting his PhD, he received a B.S. in Systems and Information Engineering from the University of Virginia, then spent a few years working at several start-up companies as a technology analyst. His research mainly focuses on the application of machine learning to monitoring air pollution and its effect on human health, with other research interests including chemical risk assessment, remote sensing, and deep learning.

Mireya Diaz received her doctorate degree in Biostatistics from Case Western Reserve University. She is a Professor and Chief of Epidemiology and Biostatistics at Homer Stryker M.D. School of Medicine, Western Michigan University. She offers 19.5 years of statistical consulting and research experience. Her research interests include the development, application, and evaluation of statistical methods for correlated data and effectiveness research with large observational data. Among her contributions to health technology assessment and these methods are evaluating:

(1) The comparative effectiveness of robotic prostatectomy from the Vattikuti Urology Institute, one of the largest and pioneering centers of the technique;

(2) Through meta-analyses the evidence used in the guidelines developed by the American Urological Association related to the management of vesicoureteral reflux in children, cryptorchidism, and follow-up of the renal mass;

(3) The evidence about infiltrates incidence with extended wear lenses, later used by Vistakon for the FDA approval of their corresponding product;

(4) The bivariate random effects model for diagnostic accuracy; and

(5) The methodological state of the art regarding indirect treatment comparisons and network meta-analysis through the pertinent ISPOR Task Force.

Ramkiran Gouripeddi is an assistant professor in the department of biomedical informatics, University of Utah. His interests are in developing novel informatics methods for performing exposure health and translational research.

Shawn "Froggi" Jackson is a Scientist at Gryphon Scientific, with over 25 years of experience in biological and chemical sciences, including over 15 years at the bench. Dr. Jackson holds a M.A. in Marine Biology from Boston University and a Ph.D. in Virology from Harvard University. She has broad experience across the chemical, biological, and associated technology development areas. In the biological sciences, she has in-depth knowledge of both human and marine/aquatic systems, with time spent at four major U.S. academic marine science centers. Additionally, Dr. Jackson has led infectious disease-related studies for a number of federal agencies (including DoD, DHS, HHS, and State). For chemical exposures, she has worked extensively to model health outcomes following human exposure to a variety of toxic industrial chemicals and toxins. With FEMA, she has worked to compile and present information that assists U.S. federal, state, and local officials with chemical incident planning and preparedness for pre- and post-chemical release event actions and activities. Over the past several years, Dr. Jackson has led projects at Gryphon identifying and assessing advances in chemical and biological technologies, including a study that was published in Nature Biotechnology (Nat. Biotech. 2019, 37: 1403.) and a study that developed a mathematical algorithm for chemical hazard prediction.

Natassja Lewinski is an Associate Professor of Chemical and Life Science Engineering at Virginia Commonwealth University. She holds a Ph.D. in Bioengineering and a B.S. in Chemical Engineering from Rice University. Dr. Lewinski has focused her career on integrating biological and environmental compatibility into the design process of engineered nanomaterials. Her research interests include nanotoxicology, nanoinformatics, sustainable nanotechnology, and comparative in vitro/in vivo analyses. Dr. Lewinski has authored and co-authored 38 peer-reviewed journal articles, 4 book chapters, and 1 provisional patent. Graduates from her group include 1 Ph.D. student and 3 M.S. students.

Kristin Magnuson is a Health Science Manager at ICF. She has 6 years of consulting experience supporting toxicological and epidemiological systematic reviews, human health risk assessment, and technical report development, as both a project manager and scientist with expertise in data extraction, study quality evaluation, evidence synthesis, tool development, workflow processes and efficiencies. She has played a key role in several systematic literature reviews, and has contributed to white papers for

the U.S. Environmental Protection Agency (U.S. EPA) and monographs for the National Institute of Environmental Health Sciences (NIEHS) National Toxicology Program (NTP).

Julia Varshavsky is an Assistant Professor of Environmental Health at Northeastern University, with a joint appointment in the Department of Health Sciences, Bouvé College of Health Sciences, and the Department of Civil and Environmental Engineering, College of Engineering. She focuses on environmental exposures and maternal-child health outcomes, as well as conducts biomonitoring studies related to advancing risk assessment in vulnerable communities.
*bio via https://coe.northeastern.edu/people/varshavsky-julia/

Session VA. Using AI Tools and Resources in Systematic Review

Andrew Rooney is acting director of the Office of Health Assessment and Translation in the National Toxicology Program at NIEHS. Rooney has been actively involved in developing risk assessment methods and guidance throughout his professional career and is a principal author of the 2012 WHO/IPCS Guidance for Immunotoxicity Risk Assessment for Chemicals. He has 20 years of experience in toxicology and risk assessment for the protection of public health and has authored over 40 peer-reviewed documents in the field including manuscripts and government assessments. For the last several years, Rooney has been working on emerging issues in toxicology and environmental health including methods to address study quality in terms of risk of bias for human, animal, and mechanistic studies and adaptation of systematic review methods for addressing environmental health questions.
*bio via https://www.niehs.nih.gov/research/atniehs/labs/iha/ohat/staff/rooney/index.cfm

Vickie R. Walker is a health scientist in the Division of the National Toxicology Program's Integrative Health Assessment Branch at National Institute of Environmental Health Sciences. She has been actively involved in the methods development and harmonization of literature-based environmental health science assessments for over 12 years. Vickie leads the evidence informatics activities for the Integrative Health Assessment branch investigating, developing, and applying semi-automated technologies (e.g., machine learning, natural language processing, and artificial intelligence approaches) and interactive data visualization in the systematic review and systematic evidence mapping workflow. The integration of these methods has improved efficiency in the time and resources required and improved communication and exploration of the gathered data. Her expertise and areas of interest are focused on the early phases of the systematic review process- problem formulation, search and selection and data extraction. Vickie's accomplishments have led to numerous invited speaker roles to present systematic review evaluation methods

at the local and international level.

Brian Howard is a senior data scientist at Sciome, LLC, a research and technology consulting company located in RTP, North Carolina. During his past 10 years at Sciome, Brian has focused primarily on the development of novel machine learning and natural language processing methods and software for the automation of systematic review. In this capacity, and working with several talented colleagues at Sciome, Brian has conducted novel research to develop methods for document prioritization and information extraction. Results of these efforts include the SWIFT-Review, SWIFT-Active Screener and FIDDLE software applications, all of which are currently used at a variety of government, academic and commercial organizations, including EPA. Dr. Howard received his PhD from North Carolina State University, his master's degree from Johns Hopkins University, and his bachelor's degrees from the University of Maryland, Baltimore County.

Nancy Baker is a literature informatics researcher specializing in finding ways to make scientific literature tasks more effective. She has a Ph.D. from the University of North Carolina at Chapel Hill through the School of Information and Library Science and the Program for Bioinformatics and Computational Biology. As an employee of Leidos, she has worked as a contractor at the EPA's Center for Computational Toxicology and Exposure for over ten years. Before becoming a scientist, she worked many years in information technology at Glaxo Wellcome. She is also an adjunct Research Assistant Professor at the UNC Eshelman School of Pharmacy.

Karen E. Ross is a Senior Bioinformatics Scientist at the Protein Information Resource and an Associate Professor in the Department of Biochemistry and Molecular & Cellular Biology at Georgetown University Medical Center. Her research interests include biomedical text mining, ontologies and data integration. She has contributed to the development of the Protein Ontology, a reference ontology in the Open Biological and Biomedical Ontologies (OBO) Foundry for proteins and protein forms, and iPTMnet, an integrated resource for post-translational modifications.

Mark Musen is Professor of Biomedical Informatics at Stanford University, where he is Director of the Stanford Center for Biomedical Informatics Research. Dr. Musen conducts research related to open science, data stewardship, intelligent systems, and biomedical decision support. His group developed Protégé, the world's most widely used technology for building and managing terminologies and ontologies. He is principal investigator of the National Center for Biomedical Ontology, one of the original National Centers for Biomedical Computing created by the U.S. National Institutes of Heath (NIH). He is principal investigator of the Center for Expanded Data Annotation and Retrieval (CEDAR). CEDAR is a center of excellence supported by the NIH Big Data to Knowledge Initiative, with the goal of developing new technology to

ease the authoring and management of biomedical experimental metadata. Dr. Musen chaired the Health Informatics and Modeling Topic Advisory Group for the World Health Organization's revision of the International Classification of Diseases (ICD-11) and he currently directs the WHO Collaborating Center for Classification, Terminology, and Standards at Stanford University.
*bio via https://med.stanford.edu/profiles/mark-musen

Julie McMurry, Associate Research Professor, University of Colorado, Anschutz Medical Campus

Session VB. Systematic Review Tools

Ryan Jones started as a Classics major, and got his Bachelor's degree in Classical Greek, but changed track when he realized the strong chance of his children starving. He moved from Utah to North Carolina and graduated from UNC SILS in 2009 with a Masters in Information Science, and has been working with HERO (Health and Environmental Research Online) since 2007, making him the only person to be involved in the project for its entire lifetime, around 15 years. He has been focused on bringing better tools and approaches to EPA science assessment, allowing human experts to cover more ground quickly.

Sean Watford is an environmental health data scientist at US Environmental Protection Agency (EPA) Office of Research and Development (ORD) Center for Public Health and Environmental Assessment (CPHEA). Sean has an MSPH and PhD in environmental health from University of North Carolina Chapel Hill. He worked with EPA/ORD through graduate school leading the redevelopment of Toxicity Reference Database (ToxRefDB) and investigated methods to assess the biological space of the Toxicity Forecaster (ToxCast) program. Starting in 2019, Sean worked as a contractor supporting data science projects across clinical domains at other federal agencies including Department of Defense (DoD), Food and Drug Administration (FDA) and Veterans Administration (VA). He rejoined EPA in June 2021 to support and develop efficiencies for EPA/ORD assessment programs.

Derek Lord
With over 10 years in the technology industry, Derek Lord applies his problem-solving skills and logical mind to help clients achieve their organizational and workflow goals. In his role as team lead for solution engineers on the Professional Services team at Evidence Partners, Derek has been directly involved in implementing solutions and best practices for dozens of clients across multiple fields. These include, but are not limited to, medical device clinical evaluation reports (CER), health economics and

outcomes research (HEOR), and pharmacovigilance (PV) solutions.

Eitan Agai, Founder of PICO Portal, has extensive expertise in machine learning (ML), natural language processing (NLP) and text mining, with a focus on applications in health relevant data and financial data. A recognized expert in the field of ML, Mr. Agai was a speaker at the 2020 North America Systematic Review Methods conference, the Evidence-Based Research (EBR) conference and the Guideline International Network in 2021. In response to the pandemic in 2020, Mr. Agai applied his extensive knowledge in the complex user interface & automation of the financial world to develop PICO Portal, an online systematic review platform and it is a perfect for other evidence synthesis projects of any size or scale. PICO Portal utilizes ML & NPL algorithms combined with an intuitive user interface.

Artur Nowak is a co-founder and the CTO of Evidence Prime. He helps the brightest minds answer the most challenging questions in healthcare through his work in the area of artificial intelligence, especially in the context of systematic review automation.

Iain Marshall, Clinical Senior Lecturer in Population Health Sciences, King's College London

Session VI. Ensuring Rigor and Reproducibility in AI Applications

Marzyeh Ghassemi is an Assistant Professor at MIT in EECS and IMES, and a Vector Institute faculty member holding a Canadian CIFAR AI Chair and Canada Research Chair. She holds MIT affiliations with the Jameel Clinic and CSAIL. Professor Ghassemi holds a Herman L. F. von Helmholtz Career Development Professorship, and was named a CIFAR Azrieli Global Scholar and one of MIT Tech Review's 35 Innovators Under 35. Previously, she was a Visiting Researcher with Alphabet's Verily and an Assistant Professor at University of Toronto. Prior to her PhD in Computer Science at MIT, she received an MSc. degree in biomedical engineering from Oxford University as a Marshall Scholar, and B.S. degrees in computer science and electrical engineering as a Goldwater Scholar at New Mexico State University.

Aarti Singh is an associate professor in the Machine Learning Department at Carnegie Mellon University (CMU). Her research lies at the intersection of machine learning, statistics, and signal processing, and focuses on designing interactive algorithms and their application to scientific domains. She received her Ph.D. in electrical engineering from the University of Wisconsin-Madison and was a postdoctoral research associate at the Program in Applied and Computational Mathematics at Princeton University. Her work is recognized by a National Science Foundation (NSF) Career Award, a U.S. Air Force Young Investigator Award, A. Nico

Habermann Junior Faculty Chair Award, Harold A. Peterson Best Dissertation Award, and four best student paper awards. Her service honors include serving on National Academies of Sciences, Engineering, and Medicine's (NASEM) Committee on Applied and Theoretical Statistics, program chair for the International Conference on Machine Learning (ICML) 2020; program chair for Artificial Intelligence and Statistics (AISTATS) 2017 conference; associate editor for Journal of Machine Learning Research and IEEE Transactions on Information Theory; expert team member for The Minerals, Metals & Materials Society's science and technology study on artificial intelligence for materials and manufacturing innovation, sponsored by the Office of Naval Research/National Institute of Standards and Technology; steering committee for NSF innovation laboratory on data-driven chemistry; and NASEM advisory committee for evaluating NSF DMREF.

John Absher, is an Associate Manager on the Product Data Science team at Squarespace (NYSE: SQSP), one of the biggest website builders. He has years of experience using AI, Data, and Analytics to build better products, improve operations, and drive business value. Currently, he is focused on projects using data and AI to improve search, personalization, and pricing. Prior to Squarespace, his roles included Data Science at Walmart, Analytics at Jet.com (acquired by Walmart for $3.3 billion), and Advanced Analytics at IBM. He graduated from Georgetown University with a B.S. in Physics and Math.