

Human and Organizational Factors in AI Risk Management - A Workshop

[Virtual]

The goal for this meeting is to identify and explore approaches to addressing human and organizational risks in AI systems. The emphasis will be on approaches that could be included in a more detailed guidance document complementing the recently issued "AI Risk Management Framework" from the National Institute of Standards and Technology (NIST). Areas of interest in the context of the NIST framework include human insights about AI-produced output and human oversight of AI systems and their operation in real-world environments.

Join us online on **June 26 from 10:45am-3:30pm ET** to discuss how to set meaningful metrics when designing tools; highlight existing organizational safety frameworks in an array of industries including healthcare, aviation, and transportation; and consider the impacts of organizational processes on the success of designing and deploying AI safety frameworks.

WEDNESDAY, JUNE 26, 2024

10:45-10:55 am ET

Welcome by Workshop Co-Chairs

Mona Sloane (co-chair), Assistant Professor, University of Virginia

Ben Shneiderman (co-chair), NAE, Professor Emeritus, University of Maryland

10:55-11:15 am ET

Opening Remarks

Tara Behrend, John Richard Butler II Endowed Professor, Michigan State University

11:15-12:15 pm ET

I. Maturing AI and Its Consequences

Moderator: **Madeleine Clare Elish**, Head of Responsible AI, Google Cloud

Speakers:

- **Missy Cummings**, Professor and Director of Mason Autonomy and Robotics Center (MARC); First American Bank Chair, George Mason University
- **Aiha Nguyen**, Program Director of Labor Futures Initiative, Data & Society
- **Hatim Rahman**, Assistant Professor of Management and Organizations, Northwestern University

12:15-12:45 pm ET

Lunch

12:45 pm-1:45 pm ET

II. AI Integration and Safety

Moderator: **Ravi Parikh**, Assistant Professor of Medical Ethics and Health Policy, and Medicine; Associate Director, Penn Center for Cancer Care Innovation; Director of Human-Algorithm Collaboration Lab, University of Pennsylvania

Speakers:

- **David Bates**, Chief of the Division of General Medicine, Brigham and Women's Hospital; Professor of Health Policy and Management, Harvard University
- **Cristina Banks**, Associate Director of the California Labor Laboratory, University of California San Francisco; Founder and Director of the Interdisciplinary Center for Healthy Workplaces, University of California Berkeley; Senior Lecturer, University of California Berkeley
- **Mark Peters**, Executive Vice President for National Laboratory Management & Operations, Battelle

1:45-2:15 pm ET

Break

2:15-3:15 pm ET

III. Organizational Dynamics and AI

Moderator: **Tara Behrend**, John Richard Butler II Endowed Professor, Michigan State University

Speakers:

- **Melissa Valentine**, Associate Professor of Management Science and Engineering, Stanford University
- **Amy Edmonson**, Novartis Professor of Leadership and Management, Harvard Business School
- **Dorothy Carter**, Associate Professor of Management, Michigan State University

3:15-3:30 pm ET

Closing Remarks

Madeleine Clare Elish, Head of Responsible AI, Google Cloud

3:30 pm ET

Workshop Event 3 Adjourn

Speaker Biographical Sketches

Dr. Cristina Banks is Associate Director of the California Labor Laboratory, a NIOSH Total Worker Health Center of Excellence at the University of California, San Francisco. Dr. Banks also founded and directs the Interdisciplinary Center for Healthy Workplaces at the University of California, Berkeley, where she leads an interdisciplinary team of scholars and practitioners in innovative research projects and in the distribution of state-of-the-art information for creating healthy workplaces. Currently, Dr. Banks is developing models of interdisciplinary teaming to increase the reach and impact of evidence-based approaches to improving worker health, safety, and well-being. In her role as a Senior Lecturer at the Haas School of Business, UC Berkeley, Dr. Banks has taught courses in management, organizational behavior and leadership to graduate and undergraduate students. She is a nationally recognized expert in management practices, human resource management, employment and labor law compliance, and worker health and well-being. She is also an elected Fellow of the American Psychological Association and the Society for Industrial-Organizational Psychology. Dr. Banks earned her doctorate in industrial-organizational psychology at the University of Minnesota in 1979 and a Bachelor's degree in Psychology with highest distinction at the University of California, Berkeley in 1974.

Dr. David Bates is the Medical Director of Clinical and Quality Analysis at Mass General Brigham (MGB), co-Director of the Center for Artificial Intelligence and Bioinformatics Learning Systems (CAIBILS) at MGB, and a Senior Physician at Brigham and Women's Hospital (BWH). He stepped down as he stepped down as Chief of General Internal Medicine and Primary Care at BWH after 25 years in June 2023. Dr. Bates is also a Professor at both Harvard Medical School and at the Harvard T.H. Chan School of Public Health; in addition, he directs the Center for Patient Safety Research and Practice at BWH. Previously, he served as Chief Quality Officer and later as Chief Innovation Officer at BWH. Dr. Bates is an internationally renowned expert in patient safety, using information technology to improve care, quality-of-care, cost-effectiveness, and outcomes assessment. He has served as president of the International Society for Quality in Healthcare (ISQua), the Board Chair of the International Society of Quality and Safety, and the editor of the Journal of Patient Safety. Dr. Bates is a member of the National Academy of Medicine, the Institute of Medicine, the American Society for Clinical Investigation, the Association of American Physicians, and the American College of Medical Informatics, and was Chairman of the Board of the American Medical Informatics Association. He has published more than 1,200 peer-reviewed papers which have been cited over 162,000 times; he is among the 400 most cited of all biomedical researchers and is listed as being among the top 250 medical scientists in the world.

Dr. Dorothy Carter (PhD, Georgia Institute of Technology) is an Associate Professor of Management in the Eli Broad College of Business at Michigan State University. Her research identifies how leaders, teams, and multiteam systems can effectively tackle complex challenges in a variety of organizational contexts including medicine, scientific research, the military, and space exploration. Her work on leadership and teamwork has appeared in numerous top-tier outlets in the organizational sciences and other disciplines. Her research program has been supported by funding from the National Science Foundation (NSF), the Department of Defense (DoD), and the National Aeronautics and Space Administration (NASA).

Dr. Mary (Missy) Cummings received her B.S. in Mathematics from the US Naval Academy in 1988, her M.S. in Space Systems Engineering from the Naval Postgraduate School in 1994, and her Ph.D. in Systems Engineering from the University of Virginia in 2004. A naval officer and military pilot from 1988-1999, she was one of the U.S. Navy's first female fighter pilots. She is a Professor in the George Mason University College of Engineering and Computing, and directs the Mason Responsible AI program as well as the Mason Autonomy and Robotics Center (MARC). She is an American Institute of Aeronautics and Astronautics and a Royal Aeronautical Society Fellow, and is a Commissioner for the Global Commission on Responsible Artificial Intelligence in the Military Domain. Her research interests include the application of artificial intelligence in safety-critical systems, assured autonomy, human-systems engineering, and the ethical and social impact of technology.

Dr. Amy C. Edmondson is the Novartis Professor of Leadership and Management at the Harvard Business School, a chair established to support the study of human interactions that lead to the creation of successful enterprises that contribute to the betterment of society. Edmondson has been recognized by the biannual Thinkers50 global ranking of management thinkers since 2011, and most recently was ranked #1 in 2021 and 2023; she also received that organization's Breakthrough Idea Award in 2019, and Talent Award in 2017. She studies teaming, psychological safety, and organizational learning, and her articles have been published in numerous academic and management outlets, including *Administrative Science Quarterly*, *Academy of Management Journal*, *Harvard Business Review* and *California Management Review*. Her 2019 book, *The Fearless Organization: Creating Psychological Safety in the Workplace for Learning, Innovation and Growth* (Wiley), has been translated into 15 languages. Edmondson's latest book, *Right Kind of Wrong* (Atria), builds on her prior work on psychological safety and teaming to provide a framework for thinking about, discussing, and practicing the science of failing well. First published in the US and in the UK (Penguin) in September, 2023, the book is due to be translated into 24 additional languages, and was selected for the Financial Times and Schrodgers Best Business Book of the Year award.

Aiha Nguyen leads the Labor Futures Initiative at Data & Society. The initiative strives to better understand emergent disruptions in the labor force as a result of data-centric technological development and to create new frames for understanding these disruptions through evidence-based research and collaboration. Aiha's recent publications include *The Constant Boss: Work Under Digital Surveillance* and *The Digital Doorstep: How Customers Use Doorbell Cameras to Manage Delivery Workers*. Prior to joining Data & Society, she worked to raise job standards for retail, airport, and other service workers, and addressed issues of food access, safety and security, and local governance at the Los Angeles Alliance for a New Economy (LAANE). Aiha has a Masters in Urban Planning from UCLA. Her interests lie at the intersection of labor, technology, and urban studies. Through these lines of inquiry, she hopes her work contributes to shifting the debate towards worker-centered analysis and solutions.

Dr. Mark Peters is the Executive Vice President for National Laboratory Management and Operations at Battelle Memorial Institute with responsibilities for governance and oversight of U.S. Department of Energy (DOE) and U.S. Department of Homeland Security national laboratories for which Battelle has a significant lab management role. Previously, he was the director of Idaho National Laboratory and president of Battelle Energy Alliance, LLC. He served two years as chairman of the National Laboratory Directors' Council, an independent body that coordinates initiatives and advises the DOE and other national laboratory stakeholders. Prior to joining Battelle, he served as the associate laboratory director for Energy and Global Security at Argonne National Laboratory. In recognition of his distinguished contributions to engineering, Dr. Peters was elected as a member of The National Academy of Engineering (NAE) in 2021, for leadership and contributions in advancing U.S. nuclear energy capabilities and infrastructure. He was honored as a Fellow of the American Nuclear Society (ANS) in 2015, for outstanding accomplishments in the area of nuclear science and technology. He received his doctorate in geophysical sciences from the University of Chicago and a bachelor's degree in geology from Auburn University.

Dr. Hatim A. Rahman is an Assistant Professor of Management and Organizations at the Kellogg School of Management, Northwestern University. His research investigates how artificial intelligence, undergirded by algorithms, is impacting the nature of work and employment relationships in organizations and labor markets. His research and teaching have received numerous awards, including the National Science Foundation CAREER award. His book *Inside the Invisible Cage: How Algorithms Control Workers* (University of California Press) investigates how digital platform organizations are using shifting, opaque algorithms to control workers in an "invisible cage."

Dr. Melissa Valentine studies how technology is changing work and organizations. Her work delves into the evolving dynamics between organizational design and AI/algorithms. Through in-depth field

studies, she explores how technology reshapes organizations, offering insights into the future of work in an increasingly digital and specialized world. Prof. Valentine is an Associate Professor in the Management Science Department at Stanford and a faculty affiliate at Stanford Institute for Human-Centered Artificial Intelligence. She and collaborators have received best paper awards for research in both management and computer science conferences, and she was awarded the prestigious NSF CAREER award. Her work has been covered in The New York Times, The Wall Street Journal, Harvard Business Review, Wired, Fast Company, and Financial Times. Prof. Valentine holds a bachelor's degree from Stanford University, a master's degree from NYU, and a Ph.D. from Harvard University.