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 11 from 10:30am-4:30pm ET** to discuss how to ensure a full range of stakeholders participate in AI system design and development, how labor processes may change with the development and use of AI, and how to build equitable infrastructure to address risk and build trust.

Tuesday, JUNE 11, 2024

10:30-10:45 am ET

Welcome by Workshop Co-Chairs and Committee
Mona Sloane (co-chair), Assistant Professor, University of Virginia
Ben Shneiderman (co-chair), NAE, Professor Emeritus, University of Maryland
Nathanael Fast, Jorge Paulo and Susanna Lemann Chair in Entrepreneurship, Executive Director of the Neely Center for Ethical Leadership, Associate Professor of Management and Organization, University of Southern California

10:45-11:00 am ET

Opening Remarks
Gabriella Waters, Research Associate, National Institute of Standards and Technology; Director of Operations, Director of the Cognitive & Neurodiversity AI (CoNA) Lab, Morgan State University; Professor, Propel Center

11:00-12:00 pm ET

I. Ensuring Stakeholder Participation in AI Design and Development

Moderator: **Nathanael Fast,** Jorge Paulo and Susanna Lemann Chair in Entrepreneurship, Executive Director of the Neely Center for Ethical Leadership, Associate Professor of Management and Organization, University of Southern California

Speakers:

- **Tawanna Dillahunt,** Associate Professor of Information, University of Michigan
- **Deep Ganguli,** Research Scientist, Anthropic
- **Brent Hecht**, Partner Director of Applied Science, Microsoft; Associate Professor of Engineering, Northwestern University

12:00-12:45 pm ET

Lunch

12:45 am-1:45 pm ET — **II. Labor Processes in AI**

Moderator: **Tamara Kneese**, Senior Researcher and Project Director of the Algorithmic Impact Methods Lab, Data & Society

Speakers:

- **Krystal Kauffman**, Research Fellow, the Distributed AI Research Institute; Lead Organizer, Turkopticon
- **Tom Kochan,** George Maverick Bunker Professor Emeritus of Management, Massachusetts Institute of Technology
- **Christina Colclough**, Founder, The Why Not Lab
- **Veena Dubal**, Professor of Law, University of California Irvine

1:45-2:15 pm ET — **Break**

2:15-3:30 pm ET — **III. Building Equitable Infrastructures to Address Risk**

Moderator: **Sheena Erete**, Associate Professor of Information Studies, University of Maryland College Park

Speakers:

- **Rayya El Zein**, Director of Partnerships, Code for Science & Society
- **Tina Park,** Head of Inclusive Research & Design, Partnership on AI
- **Aviv Ovadya**, Research Fellow, newDemocracy; Affiliate, Berkman Klein Center, Harvard; Affiliate, Centre for the Governance of AI; Founder, AI and Democracy Foundation
- **Ovetta Sampson,** Director of User Experience Machine Learning, Google
- **Gloria Washington**, Associate Professor of Computer Science, Howard University

3:30-3:45 pm ET — **Broadening Stakeholder Participation Closing Remarks**

Sheena Erete, Associate Professor of Information Studies, University of Maryland College Park
Tamara Kneese, Senior Researcher and Project Director of the Algorithmic Impact Methods Lab, Data & Society

3:45 pm ET — **Workshop Event 1 Adjourn**

Speaker Biographical Sketches

Dr. Christina J. Colclough founded The Why Not Lab with the aim to reshape the current digitalisation trajectory, so human rights, freedoms and autonomy are respected and protected. Christina's background is in labour market research and in the global labour movement, where she led their future of work policies, advocacy and strategies for a number of years. She was the author of the union movement's first principles on Workers' Data Rights and the Ethics of AI. Christina is a Fellow of the Royal Society of Arts in the UK and Advisory Board member of Carnegie Council's new program: AI and Equality Initiative. She is also a member of the UNESCO #Women4EthicalAI Platform, the OECD One AI Expert Group and is affiliated to FAOS, the Employment Relations Research Center at Copenhagen University. In 2021, Christina was a member of the Steering Committee of the Global Partnership on AI (GPAI).

Dr. Tawanna Dillahunt (she/hers) is an Associate Professor at the University of Michigan's School of Information (UMSI). Her research explores the nexus of technology, employment, and community well-being, focusing on developing digital tools for economically challenged communities in Southeastern Michigan and beyond. She uses community-centered approaches like Community Based-Participatory Research and Participatory Design to ensure stakeholder involvement. Her projects focus on creating accessible employment solutions for individuals facing digital literacy challenges, and enhancing services like ridesharing and online grocery delivery for underserved communities. Her recent explorations into speculative technology design reveal insights into anticipatory governance pertinent to discussions on AI development and stakeholder engagement. Professor Dillahunt was honored as a 2022-2023 William-Bentick Harvard Radcliffe Fellow and is currently a 2023-2024 MIT MLK Visiting Scholar. She received her Ph.D. in Human-Computer Interaction from Carnegie Mellon's Human-Computer Interaction Institute.

Dr. Veena Dubal's research focuses broadly on law, technology, and precarious workers, combining legal and empirical analysis to explore issues of labor and inequality. Her work encompasses a range of topics, including the impact of digital technologies and emerging legal frameworks on workers' lives, the interplay between law, work, and identity, and the role of law and lawyers in solidarity movements. Her research has been cited internationally in legal decisions, including by the California Supreme Court, and her research and commentary are regularly featured in media outlets, including The New York Times, The Washington Post, The Wall Street Journal, The Los Angeles Times, NPR, CNN, etc. Prof. Dubal received a B.A. from Stanford University and holds J.D. and Ph.D. degrees from the University of California, Berkeley, where she conducted an ethnography of the San Francisco taxi industry.

Rayya El Zein, PhD is Director of Partnerships at Code for Science and Society. She brings a global, justice-centered perspective to philanthropy, public policy, and community-led institution building. She works with technologists, artists, and researchers to imagine and build flexible infrastructures to support their work.

Dr. Deep Ganguli is a research scientist at Anthropic focusing on the interpretability, fairness, transparency, and societal impacts of AI. Prior to joining Anthropic, he was director of research programs at the Stanford Institute for Human-Centered Artificial Intelligence (HAI), as well as a science program officer at the Chan Zuckerberg Initiative where he developed grant making programs and convenings to foster interdisciplinary research at the intersection of artificial intelligence, data engineering, and cellular biology. He has led his own academic and industrial research in theoretical neuroscience, machine learning with humans in the loop, distributed computing, and biomedical data analysis. Deep has a PhD in computational neuroscience from New York University, and a BS in electrical engineering and computer science from the University of California at Berkeley.

Dr. Brent Hecht is Director of Applied Science at Microsoft, where he is helping lead the company-wide efforts to introduce generative AI into Microsoft's products and to adapt Microsoft's products to a hybrid work world. Dr. Hecht has an additional appointment as an Associate Professor at Northwestern

University. Dr. Hecht has been doing award-winning human-centered artificial intelligence (AI) research for over 10 years and his work has been particularly influential in language modeling, responsible AI and their intersection. His early work was central in identifying what we now call ‘algorithmic bias’. He is the recipient of a CAREER award from the U.S. National Science Foundation and his work has received Best Paper recognition at top-tier publication venues in human-centered AI (e.g. ACM SIGCHI, ACM CSCW, ACM Mobile HCI, AAAI ICWSM). Dr. Hecht’s research has been featured by The New York Times, the Washington Post, Wired, MIT Tech Review, the Atlantic, Le Monde, Der Spiegel, and other outlets. He was on the founding executive committee of ACM FAccT, the premiere publication venue for responsible AI research, and he played a key role in catalyzing the rapidly growing movement for AI researchers (e.g. NeurIPS authors) to more deeply engage with the societal impacts of their work.

Krystal Kauffman, a former political and issue campaigns organiser, co-founded Turkopticon, a non-profit organisation advocating for the rights of gig workers on Amazon's Mechanical Turk platform. Now a research fellow with the Distributed AI Research Institute, Kauffman works to build a community of workers aiming to rectify the wrongs of big-tech marketplace platforms. In her work, she has identified ethical challenges in data work and the lack of recognition for the global workforce. She also discusses the diverse impacts of periods, advocating for swift access to facilities for women with heavy flows, and the provision of hot water bottles and medication for cramp relief.

Dr. Thomas A. Kochan is the George Maverick Bunker Professor Emeritus at the MIT Sloan School of Management and a faculty member at the MIT Institute for Work and Employment Research. Kochan focuses on the need to update work and employment policies, institutions, and practices to catch up with a changing workforce and economy. His recent work calls attention to the need for a new social contract at work, one that anticipates and engages current and future technological changes in ways that can drive innovation and build a more inclusive economy and broadly shared prosperity. His most recent book is Shaping the Future of Work: A Handbook for Action and a New Social Contract (Routledge, 2021). He is a member of the National Academy of Human Resources, the National Academy of Arbitrators, and past president of the International Industrial Relations Association and the Industrial Relations Research Association. From 2018-2020 he served as a member of the MIT Task Force on Work of the Future. Kochan holds a BBA in personnel management as well as an MS and a PhD in industrial relations from the University of Wisconsin. Kochan is also a co-founder of the Worker Empowerment Research Network.

Aviv Ovadya works at the intersection of AI, platforms, democracy, and deliberation. He is a research fellow at newDemocracy; an affiliate at Harvard's Berkman Klein Center, and the Centre for the Governance of AI; a founder of the AI & Democracy Foundation (soon to be launched); and consults for civil society organizations, technology companies, and funders. Aviv's primary focus is on ensuring that the governance of AI can keep up with the rate of AI advances, building on lessons from applied deliberative democracy to enable effective transnational governance and alignment. This involves framing (e.g., "Platform Democracy"), theory (e.g., "Generative CI"), and applied work: accelerating efforts to build out and pilot the organizational and technical infrastructure for deliberative governance (formally or informally advising efforts at Meta, Twitter, and OpenAI). Aviv received his BS and MEng degrees in computer science from MIT and worked in industry in San Francisco and internationally before transitioning to focus exclusively on the societal implications of technology.

Dr. Tina M. Park is the Head of Inclusive Research & Design at the Partnership on AI. She focuses on equity-driven research frameworks and methodologies to support the responsible development of AI and machine learning technologies. She is currently working with PAI's Global Task Force for Inclusive AI to draft guidance on community-centering stakeholder engagement practices for AI and other digital technology development and deployment. Her scholarship has focused on the examination of research designs and theoretical frameworks used by sociologists and other socio-technical researchers to study social inequality and racial capitalism. Tina received a Ph.D. in Sociology from Brown University, a

Master's in Urban Planning from New York University, and a B.A. in Political Science (with minors in Chicana/o Studies and Public Policy) from the University of California, Los Angeles.

Ovetta Sampson is a tech industry professional who strives to amplify the beauty of humanity with design while avoiding practices that exploit its fragility. She has worked in multiple industries designing with various technologies including machine learning and artificial intelligence, IoT and AR/VR. Named one of the Top 15 People in Enterprise Artificial Intelligence by Business Insider in 2023, Ovetta has been designing and creating enterprise software solutions for companies all over the globe for more than a decade. She is currently the Director of UX, Machine Learning at Google LLC. A design executive who has led several multidisciplinary teams and organizations at global companies such as IDEO, Microsoft and Google, Ovetta's sweet spot is taking the deep expertise of engineering, product management and design and creating transdisciplinary practices, policies and teams that produce better, more human-centered products, platforms and services. She has a background in journalism, psychology, history and anthropology and, of course, equity and social justice and uses this foundation in her design practice. She currently is the Director of UX, Core ML at Google LLC.

Dr. Gloria Washington is an Associate Professor at Howard University in Computer Science. At Howard, she runs the Institute for Human-Centered Artificial Intelligence (HCAI@Howard) and the Affective Biometrics Lab. The HCAI@Howard Insititute is focused on performing research on all things in human-in-the-loop. Technology should not be biased, however unfortunately it is. HCAI@Howard seeks to create projects, collaborations, and policies on protecting Black and Brown people against the unintended consequences of AI technology. The Institute has several investigators and scientists performing research on topics like ethical AI, human-robot interaction, health artificial intelligence, and accessibility. HCAI@Howard currently houses the largest data collection of African American Vernacular English. This multimodal dataset is protected and housed at Howard University, but it is part of a larger initiative to create a consortium of HBCUs that protect historical African American history, culture, and language. This research is currently funded by Google, Amazon, the National Science Foundation, and the Office of Naval Research. Dr. Washington has more than fifteen years in Government service and has presented on her research throughout industry. Ms. Washington holds M.S. and Ph.D. in Computer Science from The George Washington University, and a B.S. in Computer Information Systems from Lincoln University of Missouri.

Dr. Gabriella Waters is an artificial intelligence and machine learning researcher on the ARIA team at NIST where she coordinates AI testing and evaluation across three teams. She is also the Director of Operations and the director of the Cognitive & Neurodiversity AI (CoNA) Lab at the Center for Equitable AI & Machine Learning Systems at Morgan State University in Baltimore, MD. She serves as the AI advisor and principal AI scientist at the Propel Center, where she is also a professor of Culturally Relevant AI/ML Systems. She is passionate about increasing the diversity of thought around technology and focuses on interdisciplinary collaborations to drive innovation, equity, explainability, transparency, and ethics in the development and application of AI tools. In her research, Gabriella is interested in studying the intersections between human neurobiology & learning, quantifying ethics & equity in AI/ML systems, neurosymbolic architectures, and intelligent systems that make use of those foundations for improved human-computer synergy. She develops technology innovations, with an emphasis on support for neurodiverse populations.