Conference Programme

Many Worlds of AI
Intercultural Approaches to the Ethics of Artificial Intelligence

Date: 26–28 April 2023
Venue: Jesus College, University of Cambridge
Many Worlds of AI is the inaugural conference in a series of biennial events organised as part of the ‘Desirable Digitalisation: Rethinking AI for Just and Sustainable Futures’ research programme. The ‘Desirable Digitalisation’ programme is a collaboration between the Universities of Cambridge and Bonn funded by Stiftung Mercator. The primary aim of the programme is to explore how to design AI and other digital technologies in a responsible way, prioritising the questions of social justice and environmental sustainability.

The Keynote Lecture by Bing Song, Vice President of the Berggruen Institute and Director of the Institute’s China Center, and the Relational Philosophies and Ethical Diversity in the Intercultural Evolution of AI Ethics: A ‘Disruptive’ Conversation panel are co-organized with the Berggruen Center China.

The Ethics of Digitization in India panel is co-organized with Ashoka University.
The aim of this conference is to interrogate how an intercultural approach to ethics can inform the processes of conceiving, designing, and regulating artificial intelligence (AI).

Many guidelines and policy frameworks on responsible AI foreground values such as transparency, fairness, and justice, giving an appearance of consensus. However, this apparent consensus hides wide disagreements about the meanings of these concepts and may be omitting values that are central to cultures that have been less involved in developing these frameworks. For this reason, scholars and policymakers have increasingly started to voice the need to acknowledge these disagreements, foreground the plurality of visions for technological futures, and centre previously overlooked visions – as the necessary first steps in establishing shared ethical and regulatory frameworks for responsible AI.

While planetary-scale challenges demand international cooperation in search of new solutions – including those that rely on AI – to address the crises ahead of us, feminist, Indigenous, and decolonial scholars, among others, have pointed to potential problems arising from the techno-solutionism and techno-optimism implied by the universalising ‘AI for Good’ paradigm. They recognise that some groups of humans have been multiply burdened under the current, dominant system of technology production, and that this system – if unchanged – is unlikely to bring about positive transformation. To ensure that new technologies are developed and deployed responsibly, we must, therefore, acknowledge and draw on ontological, epistemological and axiological differences, in ways that do not privilege a particular worldview. Yet in doing so, we must also work to avoid essentialising other nations or peoples, erasing extractive colonial histories, diversity washing, and cultural appropriation.

By foregrounding the many worlds of AI, we aim to create a space for dialogue between different worldviews without reifying the notion of discrete and unchanging cultural approaches to AI. The question central to Many Worlds of AI is therefore: How can we acknowledge these complexities to facilitate intercultural dialogue in the field of AI ethics, and better respond to the opportunities and challenges posed by AI?
Detailed Agenda  (In British Summer Time, GMT+1)

DAY 1: Wednesday, 26 April

9:00 am – Welcome: Dr Stephen Cave, Director of LCFI, University of Cambridge  
(Venue: Frankopan Hall)

Session 1: Fundamental Questions

Venue: Frankopan Hall

9:30 am – Panel 1: Common Vocabularies (Chair: Professor Maurizio Ferraris)

Presentations:
1. To Build “Fairer AI”, First Thoroughly Understand “Fairness”: A Multidisciplinary Review Through an Intercultural Lens by Was Rahman (Coventry University)
2. Towards a Praxis for Intercultural Ethics in Explainable AI by Chinasa T. Okolo (Cornell University)
3. Automating Desire: Laws of sex robotics in the US and South Korea by Michael Thate (Princeton University and Northwestern Pritzker School of Law)

10.30 am – Tea/ Coffee Break

11.00 am – Panel 2: Shared Policies? (Chair: Professor Jocelyn Maclure)

Presentations:
1. Emerging policy landscape around AI in South Asia and Sub-Saharan Africa by Amber Sinha (Policy Data Institute, Kampala and Mozilla Foundation)
2. “Made in Europe”: exporting European values through the regulation of Artificial Intelligence: an exploratory analysis of the Morocco–EU relationship by Oumaima Hajri (Rotterdam University of Applied Sciences)
3. AI Regulation in Brazil: National Knowledge or Foreign Appropriation? by Marina Garrote (Brasil Research Association and University of São Paulo), Paula Guedes (Catholic University of Portugal), and Bruno Bioni (Data Privacy Brasil Research Association and Brazilian Data Protection Authority)
12.00 pm – Panel 3: Accounting for AI Harms (Chair: Dr Kanta Dihal)

Presentations:

1. AI Colonialism – Karen Hao (Pulitzer Centre)

2. A taxonomy of AI-mediated epistemic injustices – Suvradip Maitra (University of Melbourne)

3. Can the Ghost Worker Speak? De-colonializing Digital Labor – Sergio Genovesi (University of Bonn)

1.00 pm – 2.00 pm: Lunch Break

Session 2: Many worlds...

Venue: Frankopan Hall | Parallel Venue: Bawden Room

2.00 pm – Panel 4: African AI (Chair: Tonii Leach)

Presentations:

1. Artificial Intelligence, Data Capitalism, and Bioethics in Sub-Saharan Africa – Golden Lwando Mwinsa, Frances Griffiths, and Benjamin Ferguson (University of Warwick)


3. African world of AI: a people-centered approach to responsible AI – Makouchi Sam Nkwo (University of Namibia), Annastasia Shipepe (University of Namibia) Shaimaa Lazem (City of Scientific and Technological Applications) Anicia Peters (University of Namibia)

2.00 pm (Parallel Session) – Panel 7: Alternative histories of AI in Europe and the Anglophone West (Chair: Christiane Schäfer)

Presentations:

1. Praxis, or the Yugoslav Search for Man: Thinking and Human Self-Realization in the Age of Generative AI – Ana Ilievska (Stanford University)

2. Conceptions of Ethics in World-Making Machines: Iconographies of AI in Ireland and the UK – Peter Rees (We and I)

3. Contentious Others: Logo and Dilemmas of Difference in the US, Britain, and France – Apolline Taillandier (University of Cambridge and University of Bonn)
3.00 pm – Panel 5: AI Cultures in MENA (Chair: Dr Audrey Borowski)

Presentations:

1. AI Projects in the Gulf Region and their Ethical Questions: An Analytical Overview – Ala Al-Fuqaha (Hamad Bin Khalifa University)
2. Islamic Ethical Discourse on AI: Three Challenges in Focus – Mohammed Ghaly (Hamad Bin Khalifa University)
3. Developing a Legal Framework for AI in Qatar and Beyond – Barry Solaiman (Hamad Bin Khalifa University and Weill Cornell Medicine–Qatar)
4. Lost History and overlooked Present: Mechanical and Artificial Intelligence in Arabic culture – Reham Hosny (The University of Cambridge/ Minia University)

3.00 pm (Parallel Session) – Panel 8: AI Histories in India (Chair: Dr Maya Indira Ganesh)

Presentations:

1. Taking off with ease or Face-off with Justice? Mapping Digital Citizenship and ‘Ways of Seeing’ the Indian Biometric State – Madhavi Shukla (Jawaharlal Nehru University)
2. Data power, AI and the “doubtful citizens”: The case of India’s National Population Register – Anirban Mukhopadhyay (University of Illinois Urbana–Champaign)
3. Palmistry, Predictive Analytics and Imprints of Colonized Bodies – Charu Maithani (University of New South Wales)
4. The Digital Afterlives of Brahminical Colonialism: Biometric Surveillance, Facial Recognition Technology, & AI Ethical Complicities in India, 1858–2022 – Nikhil Dharmaraj (Harvard College)

4.00 pm – Tea/ Coffee Break

4.30 pm – Panel 6: Confronting AI at the Margin: Conflicts around Faith, Hope, and Identity in Bangladesh (Chair: Abdullah Safir)

Panel Discussants:

1. Sharifa Sultana (Cornell University and Facebook Fellow)
2. Mohammad Rashidujjaman Rifat (University of Toronto and Schwartz Reisman Institute for Technology and Society)
3. Dr Syed Ishtiaque Ahmed (University of Toronto)
4.30 pm (Parallel Session) – Panel 9: Contemporary China and AI (Chair: Dr Kerry McInerney)

Presentations:

1. A community-of-practice approach to understanding Chinese policymaking on AI ethics – Guangyu Qiao-Franco (University of Radboud and University of Southern Denmark)

2. From Accuracy to Alignment: The Practical Logic of ‘Trustworthy AI’ among Chinese Radiologists – Wanheng Hu (Cornell University/Harvard University)

3. AI Ethics and Governance in China: from Principles to Practice – Rebecca Arcesati (Mercator Institute for China Studies)

5:30 pm – Art Performances and Presentations (Chair: Dr Stephen Cave)

Venue: Frankopan Hall

1. Post-modern dance performance and a group conversation about responsible design and social impact of AI – Betsy Campbell (Penn State University)

2. ΑΠΟαποκοποΠΟΙΗΣΗ: Decolonising Cypriot AI through poetry – Alexia Achilleos (CYENS Centre of Excellence & Cyprus University of Technology), Spyros Armostis (University of Cyprus), Eleftheria Sokratous (Ypogia Skini)

3. Cultural Memory: Artistic Experiments in AI – Yasmine Boudiaf (Royal Society of Arts and the Ada Lovelace Institute)

6.45 pm – Finish

7.00 pm – Conference Dinner (by invitation only)

DAY 2: Thursday, 27 April

Session 3: Many worlds... (contd.)

Venue: Frankopan Hall

9:30 am – Panel 10: In search of new fundamentals (Chair: Professor Alan Blackwell)

Presentations:

1. Korean value of ‘jeong’ – Robert M Geraci and Yong Sup Song (Manhattan College; Youngnam Theological Seminary)


3. What would an anti-casteist AI system look like? – Shyam Krishna (Alan Turing Institute)
10.30 am – Tea/ Coffee Break

11.00 am – Panel 11: Relational Philosophies and Ethical Diversity in the Intercultural Evolution of AI Ethics: A ‘Disruptive’ Conversation (co-organized with the Berggruen Center China)

Panelists:
1. Dr Robin R Wang (Loyola Marymount University)
2. Dr Peter D. Hershock (East-West Center)
3. Dr Osamu Sakura (The University of Tokyo)

12.00 pm – Panel 12: The Ethics of Digitization in India (co-organized with Ashoka University)

Moderator: Dr Subhashis Banerjee (Ashoka University) and Jaspreet Bindra (TechWhisperer)

Panelists:
1. Malavika Raghavan (LSE)
2. Shirin Madon (LSE)
3. Amber Sinha (Mozilla Foundation)

1.00 pm – 2.00 pm Lunch Break

Session 4: Practical Approaches

Venue: Frankopan Hall | Parallel Venue: Bawden Room

2.00 pm – Panel 13: Intercultural and Decolonial Approaches in Practice (Chair: Dr Chelsea Haramia)

Presentations:
1. Operationalizing decolonial AI through Ethics-as-a-Service – Saif Malhem (AI Future Lab), Daricia Wilkinson (Microsoft Research), Kathy Kim (Booz Allen Hamilton), Paul Sedille (Harvard Kennedy School), Nupur Kohli (European Health Parliament)
2. Multicultural design and ubuntu ethics – Bev Townsend (University of York and University of KwaZulu-Natal), Bongi Sholi (University of California, San Diego), Donrich Thaldar (University of KwaZulu-Natal).
2.00 pm (Parallel Session) – Panel 15: NLP and Intercultural Ethics (Chair: Dr Eleanor Drage)

Presentations:

1. Building NLP models to teach local languages in Africa - Aderonke Busayo Sakpere (University of Ibadan) and Makuochi Samuel Nkwo (University of Namibia, Windhoek)


3. Towards Building a Gender-Inclusive Khaya AI for English— or →Twi Text Translator – Abigail Oppong (NLP Ghana)

3.00 pm – Panel 14: Alternative Practices: New Datasets and Archives (Chair: Dr Miri Zilka)

Presentations:

1. Artificial intelligence as a decolonisation tool: Lessons from libraries, archives and museums – Maribel Hidalgo-Urbaneja (University of the Arts London); Lise Jaillant (Loughborough University)

2. Sharp Image, Vague Face: Disrupting the Facial Transparency in A.I. through a Diasporic Approach – Yifeng Wei (National College of Art and Design, Ireland)

3. AI’s Colonial Archives – Rida Qadri (Google), Huma Gupta (MIT); Katrina Sluis (ANU); Fuchsia Hart (Victoria and Albert Museum), Emily Denton (Google)

3.00 pm (Parallel Session) – Panel 16: Alternative Practices: Design and Participation (Chair: Dr Dorian Peters)

Presentations:

1. Participatory speculative design framework for fostering an accessible AI discourse: Japan and Germany – Michel Hohendanner (Munich University of Applied Sciences) and Chiara Ullstein (Technical University of Munich)

2. Imagining AI through participatory design in Nigeria, Brazil, and South Korea – Cornelius Onimisi Adejoro and Tom Yeh (University of Colorado Boulder)


4.00 pm – Tea/ Coffee Break
Workshops

**Venue:** Frankopan Hall | **Parallel Venue:** Bawden Room

**4.30 pm – Workshop 1:** Envisioning equitable representation in ML evaluation - Stevie Bergman and Boxi Wu (Deepmind)

**4.30 pm (Parallel Session) – Workshop 2:** Provotypes for embodiment of value tensions across cultures Dasha Simons (IBM)

**Keynote Lecture | Venue:** Frankopan Hall

**5:30 pm – Keynote:** Approaches to AI Ethics: “Sparks of Ideas” (Inspirations) from East Asian Philosophies  
**Speaker:** Bing Song (Senior Vice President of the Berggruen Institute and Director of the Institute’s China Center)

**7.00 pm – Finish**

**DAY 3: Friday, 28 April**

**Session 5: Just AI Futures**

**Venue:** Elena Hall | **Parallel Venue:** Bawden Room

**9:30 am – Panel 17:** AI and the Planetary (Chair: Dr Tomasz Hollanek)

Presentations:

1. Occupying Urgency: How AI Solutionism Shapes the Narrating of Urgency around the Climate Crisis – Eugenia Stamboliev, Mark Coeckelbergh (University of Vienna)

2. An Approach Based in Eastern Philosophy to Identify Ethical Issues in Early Stages of AI for Earth Observation Research – Mrinalini Kochupillai (Technical University of Munich)


**9:30 am (Parallel Session) – Panel 19:** Many Stories of AI (Chair: Dr Jan Voosholz)

Presentations:

1. Media coverage of AI in Sweden and Chile – Claudia Wladdimiro Quevedo (Uppsala University)

2. Narratives of weaponised AI: France, Japan, the US – Ingvild Bode, Hendrik Huelss, Anna Nadibaidze (University of Southern Denmark) and Tom Watts (Royal Holloway, UoFL)

3. Responsible AI reporting requires cross-border collaboration – Boyoung Lim (Pulitzer Centre)
10:30 am – Tea/ Coffee Break

11.00 am – Panel 18: Sustainability of AI (Chair: Charlotte Bander)

Panel Discussants:
1. Şebnem Yardımcı Geyikçi (University of Bonn)
2. Tijs Vandemeulebroucke (University of Bonn)
3. Larissa Bolte (University of Bonn)
4. Sophia Falk (University of Bonn)

11.00 am (Parallel Session) – Panel 20: AI with/for the Youth and the Elderly (Chair: Dr Rune Nyrup)

Presentations:
2. Human First Innovation for AI ethics? : a Cross-cultural Perspective on Youth and AI – Toshie Takahashi (Waseda University)
3. (Old) age in the age of artificial intelligence – crossing generational borders in AI research and development – Justyna Stypinska (Weizenbaum Institute)

12.00 pm – Imagining AI: How the World See Intelligent Machines: Book launch and Discussion (Venue: Elena Hall) (Chair: Dr Stephen Cave)

Discussants:
1. Anzhelika Solovyeva, PhD (Charles University)
2. Hirofumi Katsuno, PhD (Doshisha University)
3. Dr Kanta Dihal (LCFI, University of Cambridge)

1.00 pm – 2.00 pm: Lunch Break

Conclusion

2.00 pm – Concluding Remarks: Professor Markus Gabriel, Director of CST, University of Bonn (Venue: Elena Hall)

3.00 pm – Finish
Presentation 1: To Build “Fairer AI”, First Thoroughly Understand “Fairness”: A Multidisciplinary Review Through an Intercultural Lens

Abstract: News reports of “unfair” AI have become all too familiar, stretching back beyond Amazon’s inadvertently sexist recruitment system and Google’s facial recognition system confusing black humans with gorillas. Few would disagree these are wrong, but different academic disciplines offer very different views of not just how to address such issues, but what the actual issue is to address. This paper explores what we could and should learn about building fairer AI, if we treat these disciplinary differences as a form of cultural difference. As with other forms of culture, each discipline brings its own perspective, context and ontology to the subject. For example, computer science typically treats AI unfairness in explicit mathematical terms, whereas social scientists may point to implicit attitudes and behaviours rooted in colonial history. Meanwhile, legal scholars are likely to focus on demonstrable, material harm linked only to specific protected characteristics. Psychologists, neuroscientists, politicians and philosophers add further variety, alongside other, equally valid, academic perspectives. This paper examines implications of this diversity on building fairer AI systems. It is based on an ongoing investigation into discrimination towards under-represented groups in business decision-making. The findings are based on and illustrated by unpublished interview data and publicly available examples. The paper lays out a foundation of diverse literature, including critical review of common AI fairness assurance approaches and mechanisms. It focuses on underlying concepts, definitions, indicators/metrics, and language, and considers different cultural models. It compares and contrasts how different disciplines perceive (AI) fairness, using a common structure. This multidisciplinary review highlights some valuable insights into fairness that are rarely considered in AI work rooted in a single field. It reminds us of a deceptively simple finding of the overall work – to create fairer AI, we must ensure we understand fairness (at least) as much as we understand AI.

Author bio: Was Rahman is a researcher and consultant in ethical AI. He is a doctoral researcher at Coventry University, and CEO of AI Prescience, a consultancy helping organisations use AI ethically and responsibly. He has over 30 years global experience using data and technology to improve business performance. His research interests include the use of AI in organisational decision-making, ethical AI governance, and the impact of AI on social division. His current work is an investigation into the ontological diversity of different disciplinary approaches to AI fairness, focusing on AI-enabled business decision-making. In business, Was has worked with large corporates, start-ups and SMEs around the world, working with CxOs, Boards and Investors. He has held leadership roles at Accenture, Infosys and Wipro, managing business in the US, EU and Asia Pacific. He has also run start-ups and raised funding. For governments, Was has advised UK and Indian political leaders on technology industry policy. Was graduated in Physics at Oxford, and Computing at Coventry University. His AI and data science education is courtesy of Stanford, Johns Hopkins, Amazon and Google. He has been a guest lecturer at Oxford’s Said Business School, Cambridge’s Judge Business School, London Business School and IIT Madras.
**Presentation 2: Towards a Praxis for Intercultural Ethics in Explainable AI**

**Abstract:** Explainable AI is often promoted with the idea of helping end users understand how machine learning models arrive at respective predictions. Still, the majority of these benefits are reserved for those with specialized domain knowledge, such as machine learning developers. Recent research has argued that making AI explainable can be a viable way of making AI more useful in real-world contexts, especially within low-resource domains and emerging markets. While AI has transcended borders, a limited amount of work focuses on democratizing the concept of explainable AI to the “majority world”, leaving much room to explore and develop new approaches within this space that cater to the distinct needs of users within this region. This work introduces the concept of an intercultural ethics approach to AI explainability. It aims to examine how cultural nuances impact the idea of “explaining”, how existing cultural norms and values influence the adoption of modern digital technologies such as AI, and how situating local knowledge in the development of AI technologies can improve user understanding and acceptance of these systems.

**Author bio:** Chinasa T. Okolo is a fifth-year Ph.D. Candidate in the Department of Computer Science at Cornell University. Before coming to Cornell, she graduated from Pomona College with a B.A. in Computer Science. Her research interests include explainable AI, human–AI interaction, global health, and information & communication technologies for development (ICTD). Within these fields, she works on projects to understand how frontline healthcare workers in rural India perceive and value artificial intelligence (AI) and examines how explainability can be best leveraged in AI-enabled technologies deployed throughout the Global South, with a focus on healthcare.

**Presentation 3: Automating Desire: Laws of sex robotics in the US and South Korea**

**Abstract:** Oliver Wendall Holmes, the great American Supreme Court Jurist, famously asserted that basic principles do not get one very far. Legal disputes or thorny ethical challenges live betwixt and between competing and equally valid claims of rights and interests. What is needed in such scenarios is an ability to draw a line between them. “Automating Desire” adopts this legal realism by considering the legal and policy challenges introduced by the new industry of A.I. Sex Robotics. As opposed to an “ethics of” paper, or a basic principles argument, “Automating Desire” is an exercise in legal reasoning through case law as it relates to the legal framing of A.I. and desired policy outcomes in the sex industry. The emerging technology is creating new legal challenges—and not a few uncomfortable policy scenarios. The paper will proceed in three broad moves. Part One, “Obscene Desire,” will work through three American cases as it relates to import bans on the technology of child sex robots. In U.S. law, case law distinguishes between pornography and obscenity. Pornography, insofar as does not encroach the so-called “Miller Standard” of obscenity, is protected speech. Obscenity, which includes child pornography, is unprotected speech (Roth v. U.S.; and New York v. Ferber). With respect to child sex robots, there are three landmark cases to consider: Ashcroft v. Free Speech Coalition; Williams v. U.S.; and Miller v. California.
What emerges from these cases are fundamental challenges of legal classification and the designation of real v. artificial. In the Supreme Court’s holding of U.S. v. Williams, for example, the court stated, leaning on Ashcroft v. Free Speech Coalition, that child sex robots are protected speech insofar as the person who solicits the material reasonably believes the material involved does not involve real children. This section considers not only the legal framing of the matter, but takes seriously our intuitive sense of wrongdoing irrespective of the standard of real v. artificial in policy questions of therapy through legal philosopher, Joshua Kleinfeld’s rubric of “victimization” (2013). Part Two, “Governing Desire,” considers the South Korean court’s ban of imported sex robots on the grounds of what it deemed: threats to disrupt the constitutional order (Article 234 of their Customs Law). This ruling, I suggest, alerts us to the governing function of “Desire” on the one hand, and, on the other, the policy question of control: who gets to control what I desire. Part Three, “Automating Desire,” concludes by thinking prospectively from the edge of where law and policy currently stand on the question of sex robots. This section reflects on both the discomforts and pro-social applications of the technology of A.I. Sex Robotics.

Author bio: Michael J. Thate, Ph.D. is an Associate Research Scholar and Lecturer at Princeton University’s Faith & Work Initiative, SEAS, and the Keller Center for Innovation in Engineering Education; and a Law Student at Northwestern Pritzker’s School of Law. He has held visiting fellowships and lectureships at Yale, Harvard, Durham University (U.K.), and l’École normale supérieure, Paris. He was a recipient of the Alexander von Humboldt Award, spending three years at Universität Tübingen in Germany. Michael’s academic interests are informed and complemented by his corporate consulting experience on matters relating to brand equity, communication strategy, and corporate trust. Michael is the author of two monographs: Remembrance of Things Past? (Mohr Siebeck 2013) and The Godman and the Sea (UPenn Press, 2019). He has edited four other volumes and written several articles on subjects ranging from suicide, philosophy of religion, participation, labor, time, money, the second space age, the attention economy, design thinking, and business ethics. His work attempts to track genealogies of thought and set into comparison the assemblages of ethical questions. He is currently working on two books: Scented Life and Natural Prayers of the Soul. The former considers the ethical challenges of difference among and between lifeforms. The latter is an ethical reflection on the so-called attention economy.

Chair Bio:

Professor Maurizio Ferraris wrote more than sixty books that have been translated into several languages. The last ones are Post-Coronial Studies (Einaudi) and Documanità. Filosofia del mondo nuovo (Laterza). Full Professor of Philosophy, he is the President of the Labont – Center for Ontology. He is columnist for ‘Corriere della sera’ and for ‘Neue Zürcher Zeitung’. He is also the director of “Scienza Nuova”, an institute of advanced studies – dedicated to Umberto Eco and uniting the University and the Polytechnic of Turin – aimed at planning a sustainable future, both from a cultural and from a political point of view.
Presentation 1: Big Tech and its adversaries: Situating platform power within the geopolitical battle for data

Abstract: Perhaps there is no clearer indication of the primacy of data in this age than the overworked metaphors that are often used to describe it. In the last few years, data has been likened, aside from the hackneyed comparison to ‘oil’, to any number of (tangible entities) such as mineral deposits, (Hooper, 2017) dividend deposits, (Sumagaysay, 2019), and even the Alaskan Permanent Fund (Hughes, 2018). On the other end of the spectrum, commentators have also compared data to radioactive materials such as uranium and pollutants such as carbon dioxide (Tisne, 2019). As tired or inventive these metaphors may be, they signify a desperate need for a clear conceptual model through which we can think through the legal, social and economic ramifications of data. The comparisons of data to an asset are intrinsically linked to the question who controls this asset. In their 2006 book, Who Controls the Internet?: Illusions of a Borderless World, Jack Goldsmith and Tim Wu narrated how the idea of a borderless Internet first ran into territorial governments. A decade and a half later, we see new battlelines being drawn, only now with large American BigTech companies and governments, more notably from the Majority World at its centre. The struggle over who controls the Internet has only continued its journey to becoming the important geopolitical tug-of-war of our times, with data—how it is collected, stored, protected, used, and transferred over national borders—being its most recent site of battle. This new geopolitical war has a diverse cast of characters and interests. In the last decade, the biggest technology firms Facebook, Apple, Google, Amazon, and Microsoft, and their foreign counterparts such as Alibaba, Huawei, Baidu and Tencent have created a ‘new dimension in geopolitics’ I will situate platform power as a consequence of data governance practices within this emerging geopolitical battle for data. Significant attention has been paid to the struggles between BigTech companies and governments in the US and EU. However, there has been little or no scholarly analysis of states in the Majority World flexing their power to not only rein in power exerted by BigTech players in their jurisdictions but also the use of regulatory muscle.

Author bio: Amber Sinha works at the intersection of law, technology and society, and studies the impact of digital technologies on socio-political processes and structures. Until June 2022, he was the Executive Director of the Centre for Internet and Society (CIS), India. He has led programmes on privacy, identity, AI, and free speech. He is a Senior Fellow–Trustworthy AI at Mozilla Foundation studying models for algorithmic transparency, and Director of Research at Pollicy Data Institute, Kampala. Amber is a member of the Steering Committee of ABOUT ML, an initiative to bring diverse perspectives to develop, test, and implement machine learning system documentation practices. He also serves on the GPA Reference Panel of Global Privacy Assembly. His first book, The Networked Public, was released in 2019. Amber is the Director of Research at Policy.
Presentation 2: “Made in Europe”: exporting European values through the regulation of Artificial Intelligence: an exploratory analysis of the Morocco–EU relationship

Abstract: Much has been written about the transformative power of AI systems and how countries are racing for global AI dominance to reap the economic and geopolitical power expected to result. Nevertheless, this ‘race to AI’ is bringing forth a ‘race to AI regulation’ where a new playground for global regulatory competition seems to emerge. With the AI Act, the European Commission is introducing an extraterritorial regulatory framework for AI systems to ensure that systems placed on- and used within the European Union (EU) market comply with EU values. Notably, despite this being a good step in the right direction to protect fundamental rights - it remains a rather pompous and self-proclaimed aim to produce and foster universal AI systems that are ‘made in Europe’. Specifically, aware of the power of norms, the EU seems to strategically capitalise on the opportunity to spread its normative influence, export its values and promote its vested interests through regulating AI systems. Consequently, the question arises whether the AI Act can be perceived as a tool for Western normative dominance, as it denies the diversity of humankind’s ethical stances. Stepping outside of this self-referential Western point of view requires a view from elsewhere, which this thesis aims to accommodate by analysing the case of Morocco. To date, Morocco is unconsciously still tied to its colonial masters in Europe and is actively trying to independently boost its economic and social development through digital transformation – of which the national AI strategy is a prime example. However, being a postcolonial and a ‘periphery’ country, it will soon find itself at the receiving end of the EU’s dictum of what ‘ethical AI’ is – leaving no other choice than compliance considering its significant dependencies on the EU market. Therefore, the central thesis investigates whether the EU’s attempt to regulate AI through the AI Act can be charged with ‘normative imperialism’ when investigating Morocco’s case. The aim is to conduct a literature review to locate the existing knowledge on this topic and to fill the cavity with qualitative semi-structured interviews covering different institutional perspectives in Morocco. The aim is to fill the void left by previous research on how the EU’s normative imperialism, tied to its suppressed and colonial history of ‘peace’ and ‘prosperity’, continues to impact Global South countries to date through contemporary forms.

Author bio: Oumaima Hajri is a researcher and lecturer at the Rotterdam University of Applied Sciences. Her work focuses on the social impact of AI. For the Designing Responsible AI Media Applications project, she is investigating, in collaboration with media organisations in the Netherlands, how AI can be applied in a responsible manner. She also deals with regulations (such as the AI Act), in particular how ethical guidelines can make an important contribution in translating strategic interventions into practice. She is currently part of the first cohort of the MSt AI Ethics & Society at the University of Cambridge, mainly focusing on decolonization and demystification. Additionally, Oumaima co-founded the platform “AI Better World” to raise awareness, educate about and deconstruct AI.
Abstract: The implementation of Artificial Intelligence (AI) systems in Brazil has grown in recent years. Along with it came the first concerted effort to regulate the technology through different instruments: first, the Brazilian Strategy for Artificial Intelligence (EBIA), then a first Draft Bill (PL 21/2020), and, finally, the most recent step, a Commission of Jurists in the Senate for the writing of an amendment to the draft bill. Regulating AI in the South presents a more significant challenge than in Global North countries, as northern citizens have older and more structured right-protecting mechanisms regarding data protection as well as more robust democracies and institutional protections and hence more extensive protection of human rights. Thus, any regulation in the South needs to consider their population’s local context and vulnerability to the risks of AI and the utmost relevance of popular participation in the regulations (Arun, 2020). This article aims to analyze the amendment’s text to the Brazilian Draft Bill written by the Commission of Jurists. Is there a consideration of the local context and the population vulnerabilities reflecting the participatory processes in which local knowledge was emphasized? Is there a prevalence of foreign knowledge and appropriation of models of regulation? Are these models a rights-based or risk-based approach? This is a new analysis, and answering the question will allow the evaluation of the state of AI regulation in the country. An analysis of the efforts so far to regulate AI demonstrate that, although there was the big goal of producing a framework to regulate the technology in the country, there was no proper consideration of the contributions made by stakeholders and the public during participatory processes as well as the resulting instruments were not fit to their purpose (Belli et al., 2023).

Author bios: Marina Garrote is a Researcher at Data Privacy Brasil Research Association. Master’s student at the University of São Paulo. Lawyer and academic in Data Protection, Digital Rights, Access to Justice, and Gender and Sexuality. Specialist in Gender and Sexuality at the University of the State of Rio de Janeiro.

Paula Guedes is a Law and Technology Specialist. Phd candidate at Catholic University of Portugal – Faculty of Law in Porto. Master in International and European Law Catholic University of Portugal – Faculty of Law in Porto.

Bruno Bioni is the Executive Director at Data Privacy Brasil Research Association. Lawyer and professor in the field of regulation and new technologies. Ph.D. in Commercial Law at the University of São Paulo, Masters degree in Private Law at University of São Paulo. Founder of Data Privacy Brazil and founding partner of Bioni Consulting. Member of the Board of Directors of the Brazilian Data Protection Authority, as a representative of civil society. Member of the Jurists’ Commission responsible for subsidizing the elaboration of the substitute draft for the Bill (no.5051, of 2019), which aims to originate Brazil’s Legal Framework for Artificial Intelligence.
Chair bio:

After a B.A. at Laval University and an M.A. at the University of Victoria, Professor Jocelyn Maclure earned a DPhil from the University of Southampton. Before joining the Department of Philosophy at McGill University in 2021, he taught for 17 years in the Faculty of Philosophy at Laval University. He held visiting appointments in several universities. Jocelyn’s research is primarily in ethics and political philosophy. His book Secularism and Freedom of Conscience (Harvard University Press, 2011), co-authored with Charles Taylor, appeared in 9 languages. His recent work on artificial intelligence and on end-of-life issues led him to explore different metaphysical questions ranging from the mind-body problem to the enigma of personal identity. As a public philosopher, he served as an expert-analyst for the Bouchard-Taylor Commission on cultural and religious accommodations, co-chaired the group of experts on medical assistance in dying, advanced medical directives and neurodegenerative diseases, and participated in numerous public debates. He is the current president of the Quebec Ethics in Science and Technology Commission.

----------
Presentation 1: AI Colonialism

Abstract: Over the past few years, a growing number of scholars have argued that AI development is repeating colonial history. European colonialism was characterized by the violent capture of land, extraction of resources, and exploitation of people for the economic enrichment of the conquering country. While it would diminish the depth of past traumas to say the AI industry is repeating this violence, it is now using more insidious means to enrich the wealthy and powerful at the great expense of the poor.

My AI Colonialism series for MIT Technology Review, supported by the MIT Knight Science Journalism program and Pulitzer Center, dug into these parallels between AI development and our colonial past by examining communities around the world that have been profoundly changed by the technology. I will share the stories I gathered from South Africa, Venezuela, Indonesia, and New Zealand, which together reveal how AI is impoverishing the communities and countries that don’t have a say in its development—the same communities and countries already impoverished by former colonial empires. They also suggest how AI could be so much more—a way for the historically dispossessed to reassert their culture, their voice, and their right to determine their own future.

Author bio: Karen Hao is a Hong Kong-based reporter at the Wall Street Journal, covering China’s technology industry and its impacts on society. She was previously a senior editor at MIT Technology Review, covering artificial intelligence. Her work is regularly taught in universities and cited in government reports and by Congress. She has received numerous accolades for her coverage, including an ASME Next Award for Journalists Under 30, two Front Page Awards, and several Webby Award nominations.

Presentation 2: A taxonomy of AI-mediated epistemic injustices

Abstract: Epistemic injustices are defined by Miranda Fricker (2007) as harm to our capacity as knowers by reason of social prejudice. A burgeoning literature considers how the use of Artificial Intelligence (‘AI’) systemically perpetuate such epistemic injustices. I propose a taxonomy of AI-mediated causes of harms that are recognised within the literature as being distinctly epistemic in nature. Based on the causative factor of the underlying social prejudice at play as understood with insights from social psychology, I propose a typology of epistemic harms: ‘AI-exacerbated’, ‘AI-generated’ and ‘AI-consolidated’. The taxonomy facilitates identification of remedies for distinct causations within a comprehensive account of AI and epistemic injustice. ‘AI-exacerbated’ epistemic harms are caused by pre-existing societal prejudices, the effect of which is exacerbated by AI. An illustration is Amazon’s infamous resume.
sorting algorithm that automated pre-existing social prejudice (ie. gender bias) in the company’s hiring practices which led to ‘hermeneutical’ injustice. (Rafanelli, 2022) ‘AI-generated’ epistemic harms are caused where the deployment of AI systems generate new forms of social prejudice. An illustration is where a cognitive bias such as ‘automation bias’ is engaged whereby AI systems are improperly considered more reliable than humans. This may lead to prejudice against humans in their capacity as epistemic agents, resulting in ‘testimonial smothering’. (Nihei, 2022) ‘AI-consolidated’ epistemic harms are caused where a pre-existing social prejudice is consolidated with a new form of social prejudice generated by AI usage. An illustration is the 2016 decision of State v Loomis where Mr Loomis was unable to challenge the COMPAS risk-assessment system leading to ‘hermeneutical injustice’. Pre-existing prejudice (ie. racial bias) preventing Black defendants from having a fair trial was consolidated with a distinct form of prejudice caused by the proprietary nature of COMPAS which prevented the sharing of the risk-assessment calculations with Mr Loomis. (Symons & Alvarado, 2022)

**Author bio:** Suvradip is currently completing a Master of Laws at the University of Melbourne. As part of his Masters, Suvradip will be a Visiting Student at the Leverhulme Center for Future of Intelligence. At LCFI Suvradip will be working under the mentorship of Dr Kanta Dihal to explore the ethics of AI systems from an intercultural lens. In his research, Suvradip is interested in combining understandings of colonial histories of knowledge with theories of epistemic justice and intercultural ethics to understand the impact of AI on marginalised populations. Suvradip graduated with First Class Honours from the University of Queensland with a Bachelor of Science/LLB (Hons) majoring in physics. Since graduating, Suvradip has been involved in various projects researching the impact of technology on society, including at Harvard University’s Berkman Klein Centre for Internet and Society, and Global Catastrophic Risk Institute. His research has been published in the Australian Law Journal and Proceedings of the AI and Ethics Society Conference. He has previously practised as a commercial lawyer and was associate to a Judge in the Queensland Court of Appeal.

**Presentation 3:** Can the Ghost Worker Speak? De-colonializing Digital Labor

**Abstract:** The current training and development processes of AI systems are based on the exploitation of “ghost work” (Gray, Sury 2019). Typical tasks of ghost workers include labeling images and sentences, verifying data, and moderating content. Based on remote work, ghost work represents a case of work outsourcing and offshoring and can be seen as part of the larger phenomenon of “algorithmic coloniality” or “data colonialism” (Mohamed et al. 2020). For example, web-based microwork platforms such as Amazon’s Mechanical Turk, Samasource, and CrowdFlower enabled new forms of labor offshoring from corporations mostly based in the U.S., U.K., India, and Australia to workers in Africa, Latin America, and Southeast Asia (Rani, Furrer 2020; Anwar, Graham 2020; Royer 2021). In the first part of my talk, I outline the main ethical concerns related to digital labor (Fuchs, Fischer 2015) focusing especially on the practices leading to its outsourcing, offshoring, and exploitation, and considering the perspectives of scholars from the global south (Albrieu 2021).
In the second part of my talk, I explore solutions targeting AI systems’ design process and regulation. On the one hand, I suggest that ethical questions concerning the fulfillment of tasks usually performed by ghost workers should be already addressed by product design and not be left to chance. On the other hand, I stress that ghost workers’ rights should be protected by regulations. In order for this to happen, digital labor performed by ghost workers should be acknowledged as work and regarded as part of the software development process by lawmakers, corporations, and consumers (Snower, Twomey 2022). Moreover, in a globalized economic context, international supply chain laws addressing new forms of digital labor are necessary to prohibit the sale of products developed through work offshoring and exploitation.

**Author bio:** Sergio Genovesi is a postdoctoral researcher at the Center for Science and Thought of the University of Bonn and a team member of the KI.NRW flagship project “Zertifizierte KI” (Certified AI). His research focuses on the ethics and ontology of technology, an specifically of AI systems. He holds a Ph.D. in theoretical philosophy from the University of Bonn.

**Chair bio:**

**Dr Kanta Dihal** is a Senior Research Fellow at the Leverhulme Centre for the Future of Intelligence, University of Cambridge. She leads two research projects, Global AI Narratives and Decolonizing AI, in which she explores intercultural public understanding of artificial intelligence as constructed by fictional and nonfictional narratives. Kanta’s work intersects the fields of science communication, literature and science, and science fiction. She has a PhD in science communication from the University of Oxford: in her thesis, ‘The Stories of Quantum Physics,’ she investigated the communication of conflicting interpretations of quantum physics to adults and children. She is co-editor of the books AI Narratives: A History of Imaginative Thinking About Intelligent Machines (Oxford University Press, 2020) and Imagining AI: How the World Sees Intelligent Machines (Oxford University Press, 2022) and has co-authored a series of papers on AI narratives with Dr Stephen Cave, including ‘The Whiteness of AI’ (Philosophy and Technology, 2020). She is currently writing the book Stories in Superposition. For more of her publications, see the Publications page. Kanta has advised the World Economic Forum, the UK House of Lords, and the United Nations on portrayals and perceptions of AI. She has been an invited speaker for national and international events, radio, and TV.
**Presentation 1: Artificial Intelligence, Data Capitalism, and Bioethics in Sub-Saharan Africa**

**Abstract:** There has been an increase in the digitalisation of healthcare data in Sub-Saharan Africa because of its potential to improve the health outcomes of the people. This includes individual’s clinical records (symptoms, diagnoses, test results), image data such as photographs and X-rays and self-monitoring data (exercise, heart rate). Where there is large volume digital data, there is potential for the use of Artificial Intelligence (AI). In this article we focus on non-knowledge-based AI such as data mining. Often little attention is given to how the collection, collation, usage, ownership, and control of digital health data fits with local ethical systems. This paper seeks to explore the application of Ubuntu ethics to digital data in Sub-Saharan Africa’s healthcare system. Ubuntu is an ancient ethical system practiced across Sub-Saharan Africa by the Bantu speaking people, and it is best explained by the Nguni proverb, “Umuntu ngumuntu ngabantu”, translated as “a person is a person through other persons”. Ubuntu ethics put the community and the social good before an individual and does not regard an individual as an autonomous private being. Consequently, the individual has little latitude for self-determination outside the context of the community. Fundamentally, Ubuntu is not opposed to personal data sharing as it does not believe that an individual is a private being. However, it is on condition that collected data brings about social good for the community. In western cultures, the collection of electronic data raises serious ethical challenges about patient privacy, autonomy, consent, and confidentiality. In this paper, using the example of a specific health issue, we will consider the challenges in relation to electronic health data within the Ubuntu ethic. Who is the community? What is a community benefit? Who decides?

**Author bios:** Golden Lwando Mwinsa is Chevening and Commonwealth Scholar studying for a PhD in bioethics at the University Warwick, in the Social Science and Systems in Health Unit (SSSU), Division of Health Sciences, Warwick Medical School. He is a Public Health and Social Development Professional with 13 years practical experience in Zambia and Sierra Leone.

Frances Griffiths is Professor of Medicine in Society at University of Warwick, South Africa Chairs Initiative Professor of South Africa Research Chairs Initiative Award Holder, Centre for Health Policy, University of the Witwatersrand, Fellow of the Alan Turing Institute and co-chair of the Institute’s Data Ethics Group. As a social scientist she specialises in research on social, health, health service and health policy implications of innovation in health care including the use of digital communication. She is an experienced general medical practitioner (GP) currently working as a GP in Coventry, UK.

Benjamin Ferguson is Professor of Philosophy and Director of PPE at the University of Warwick, United Kingdom, and member of the Warwick International Higher Education Academy, among other memberships. He has published widely in peer-reviewed journals and has co-authored several books. His research interests include exploitation, consumer choices, labour, trade justice, migration, and human rights. He has also presented at several international conferences.
Presentation 2: AI, Journalism, and the Ubuntu Robot in Sub-Saharan Africa: A quest for a normative framework

Abstract: The notion of how Artificial Intelligence (AI) can be used for the greater good has continued to elude scholars. While some focus on how AI would change the world for the better, others hanker on the idea that AI is built on a framework that impedes equality and perpetuates biases. The debates have led to a chasm between the content creators and the users. As a result, a number of calls for a normative framework have emerged to understand the symbiotic relationship. This study, therefore, investigates the Ubuntu framework as one of the proposed typologies for understanding AI. Drawing from interviews conducted among 43 journalists in five sub-Saharan African countries, the study examined the Ubuntu framework in the context of the core critiques of AI: 1. exclusion of marginalized communities in the design of automated decision-making systems, 2. biases in data selection, 3. failure to recognize societal interconnectedness, 4. commodification of digital selves, and 5. data centralization. Essentially, we interrogate how each of the core critiques shapes the collection, production, and dissemination of news content in Sub-Saharan Africa. We employ both mathematical and theoretical insights to underscore the process of news production and thus argue for Ubuntu as a universal framework.

Author bio: Greg Gondwe, PhD, is an Assistant Professor of Journalism at California State University – San Bernardino, and A Visiting Scholar with the Institute for Social Media Rebooting at Harvard.

Presentation 3: African World of AI: A people-centered approach to responsible AI

Abstract: Globally, AI stakeholders have continued to engage with contemporary discussions on values such as ethics, transparency, explainability, data integrity, fairness, and governance, which are germane to responsible AI discourse. Many guidelines and policy frameworks have been recommended and are being adopted in this direction. In Africa however, awareness of these policies is at their early stages despite increasing activities in AI R&D on the continent. Conflicts in cultural values, digital divide, and age-long inequalities perpetuated through non-involvement of local AI stakeholders in developing these frameworks have impacted its acceptance on the continent (Gwagwa et al., 2021). But one-size-does-not-fit-all. The wholesale adoption of AI solutions developed without the consideration of afrocentric perspectives could aggravate its potential negative impacts on users, communities, and environment. Ethical and responsible products need to be seen as inherently related to the local context. There is a need for a new approach that allows local designers to develop socially responsible frameworks instead of depending on existing models that don’t align with communities’ values. This approach will promote responsible AI by allowing startups to collect sentiments of their technology ideas from peoples’ perspectives—how it might affect users in the communities. Through public engagement starting with startups in Africa and Namibia’s AI ecosystem, we attempt to uncover how startups and designers interpret users’ opinions, map them to
relevant principles and guidelines for responsible AI, and how they tailor them in practice. The results will support AI-startups to harness local and situated design ideas. The results will further translate and integrate them into existing AI frameworks. It will be invaluable in building AI capacity that addresses contextual issues of ethics, equity, and inclusion especially as it affects the cherished values of marginalized communities. It is expected to contribute to creating AI-solutions that are unique and appropriate for the African market.

Author bios: Dr. Makuochi Samuel Nkwo is a Postdoctoral Research Fellow, in the Research Innovation and Development Unit, University of Namibia, Windhoek. He has done high-quality and impactful single and collaborative research works in health & wellness, teaching & learning, eCommerce & workplaces, as well as in sustainable environments. Also, he has authored and co-authored, published, and presented over 26 peer-reviewed papers in reputable journals and conferences around the world. He is currently leading the Responsible Human-Centered Artificial Intelligence Design in Africa project, hosted at the University of Namibia, Windhoek.

Annastasia Shipepe is a Lecturer at the Department of Computing, Mathematical & Statistical Sciences at University of Namibia.

Prof Anicia Peters is the Pro-Vice-Chancellor for Research, Innovation and Development at University of Namibia, also the Chairperson of the Namibia Presidential Task Force on the Fourth Industrial Revolution. She serves as ACM CHI 2023 Technical Program Co-chair (Hamburg, April 2023).

Dr Shaima Lazem is an Associate Research professor at City of Scientific Research and Technological Applications SRTA-City, Egypt, and Co-founder of ArabHCI. She demonstrated leadership skills in 2017 UK-Egypt Newton-Mosharafa funded project “The Hilali Network: Exploring Lived Cultural Heritage Through Design in Higher Education” in partnership with Kingston University, UK. She was awarded the Leaders in Innovation Fellowship with the Royal Academy of Engineering in London in 2018 and prestigious Grace Hopper Celebration of Women in Computing scholarship in 2011 & 2016.

Chair bio:

Tonii Leach is a Research Assistant at the Leverhulme Centre for the Future of Intelligence. She is interested in the societal and ethical implications of AI narratives, and the impact of science fiction narratives on global AI policies. Tonii is a PhD candidate in the Centre for Computing and Social Responsibility at De Montfort University, Leicester. She holds an MA in Modern Literature from the University of Leicester and a BA from The Open University.
Panel/Group Presentations: AI Ethics in the Muslim World: Perspectives from the Gulf Region

Convener: Mohammed Ghaly

Abstracts: The oil-rich Gulf Cooperation Council (GCC) countries are leading the Muslim world in the field of Artificial Intelligence (AI). In their efforts to move toward knowledge-based economies, these countries hope to achieve returns of up to USD billions from AI. In 2016, an autonomous transportation strategy was launched in Dubai and the world’s first AI Minister was appointed in the UAE. In 2019, a dedicated AI university was established in Abu Dhabi. Moreover, different GCC countries developed their respective AI strategies principles. The Muslim world is no exception to the general rule that technological developments do not work in a vacuum and thus cannot be isolated from people’s moral worlds and ethical frameworks. Throughout the Gulf region, values rooted in the Islamic tradition are an integral part of people’s value system and moral compass. This is true to the extent that many AI applications have found their way into the very domain of religious affairs. Just as representative examples, AI technologies are increasingly employed to make the pilgrimage (hajj) a smart ritual so that it becomes more efficient and convenient. Also, the centuries–old tradition of issuing religious advice (ifta’) has recently been made available through an AI–powered service. On the other hand, available literature shows awareness of the need to address AI ethical questions and associated social risks through the lens of the socio-religious values and cultural norms rooted in the Islamic tradition. Against this background, the multidisciplinary presentations in this panel will explore the various dimensions of AI ethics in the Gulf region, drawing from disciplines such as Computer Engineering, Islamic Studies and Ethics, and Legal Studies.

Presentation 1: AI Projects in the Gulf Region and their Ethical Questions: An Analytical Overview (Ala Al–Fuqaha)

Artificial intelligence (AI) has seen significant growth and development in the Gulf region in recent years. AI driven systems are increasingly utilized in personal and mission–critical applications, including medical diagnosis, autonomous vehicles, smart city infrastructure, smart manufacturing, financing, news, etc. This presentation will provide an overview of (a) the types of AI projects that have taken place in the Gulf region and (b) the main ethical questions that computer engineers, who are active in this region, usually grapple with. Recent years have witnessed several AI projects in the region that cut across a wide range of fields and sectors. With the aim of improving people’s quality of life, different AI projects engage with key sectors in society such as healthcare, transportation and finance. Examples of such projects include the “smart” robots that were introduced by the Dubai Health Authority (DHA). Also research institutes and hospitals jointly work on different projects tailored to the national and regional needs of the healthcare sector, e.g., those focusing on prevailing diseases such as diabetes and genetic conditions. To show awareness of the religio–cultural context of the region, different AI projects employ AI technologies to address issues that relate to the Islamic Scriptures (Quran and Sunna) and the overall heritage of Islamic scholarship in addition to the Arabic language Arabic assistant technology, and Arabic sentiments on social media outlets.
To pave the way for the second presentation in the panel, this presentation will review the key questions that computer engineers usually face during their work in the Gulf region. Some of these questions relate to the nature of AI technologies, whereas other questions are generated by the religio-cultural context of the region.

**Presentation 2: Islamic Ethical Discourse on AI: Three Challenges in Focus**
(Mohammed Ghaly)

Cutting-edge AI technologies pose tough ethical questions and challenges that cut across people’s geographies, moral convictions and religious beliefs. As for the Islamic ethical discourse, this presentation will analyze three main challenges. First, the applications of AI technologies are creeping into the religious domain with questions about the (im)permissibility of “robotizing” significant Islamic rituals, e.g., replacing the leader of ritual prayer (imam) or the scholar who issues religious advice (mufti) with a robot or AI system. To analyze this challenge, reference will be made to a number of relevant fatwas in addition to the 2019 Dubai initiative to launch the first-of-its-kind AI-powered ‘Virtual Ifta’ service in both English and Arabic. Secondly, the deep and complex AI questions require not only direct answers to subject-specific questions, but also a reconsideration of one’s overall worldview. Many Muslims believe that Islam has a comprehensive religio-moral system (Sharia), whose scope is flexible enough to respond to such grand challenges. The concept of ensoulment (nafkh al-ruh), with its theological and juristic dimensions, will be used to illustrate this Sharia-based dimension. This element will also be analyzed in relation to religious scholars’ responses to computer engineers’ questions about making humanoid robots. Thirdly and of particular relevance to the Gulf region context, are the international discussions on creating “transcultural AI principles”. To highlight this challenge, the relevant (semi-) official documents produced by GCC countries will be critically reviewed. These documents extensively engage with widely circulated AI principles like fairness, justice and explainability. On the other hand, they also stress the need to consider the religio-cultural context of the Gulf region. To pave the way for the next presentation in the panel, it will examine how these documents tried to strike a balance between these two aspects of AI ethics so that they can produce an appealing legal and regulatory framework.

**Presentation 3: Developing a Legal Framework for AI in Qatar and Beyond**
(Barry Solaiman)

AI is the target of much research in the GCC region and is being deployed in various sectors for efficiency gains. As with all other countries, there is no specific law regulating this space. The pattern that has arisen in the region is much the same as elsewhere. Namely, experts have identified specific concerns and challenges of AI. They seek to adapt existing guidelines and laws to regulate the use of AI but find that those laws are lacking. Thus, as a stopgap, they develop soft-law guidelines until legislation is created. To elucidate this process in the region, this presentation examines the developments in Qatar, which is a regional leader on the development
of AI. The health sector is used to highlight the dynamics that have arisen owing to the significant attention given to AI in the medical realm. AI raises unique legal challenges owing to its autonomous nature, limitations surrounding explainability, and a plethora of complex data. In medicine, these raise specific issues for patient privacy, informed consent and medical liability, in addition to cultural norms specific to the region. There are existing laws that touch upon those issues, setting the standard of care for medical liability, outlining requirements for informed consent and data processing rules. However, the legal framework as it applies to AI is piecemeal and untargeted. As such, researchers at HBKU in conjunction with the Ministry of Public Health in Qatar have been working to develop soft-law guidelines in the interim to resolve some of the legal gaps. Those guidelines are being tailored for AI healthcare researchers in Qatar. Moving forward, consideration must be given as to how a comprehensive legal framework can be developed from those existing initiatives that applies broadly across sectors.

Author bios: Dr. Ala Al-Fuqaha is a professor and associate provost at Hamad Bin Khalifa University (HBKU). His research interests include the use of machine learning in general and deep learning in particular in support of the data-driven and self-driven management of large-scale deployments of IoT and smart city infrastructure and services, ethical aspects of AI deployments, Wireless Vehicular Networks (VANETs), cooperation and spectrum access etiquette in cognitive radio networks, management and planning of software-defined networks (SDN), and engineering education. He is a senior member of the IEEE, a senior member of the ACM, and an ABET Program Evaluator (PEV) and commissioner. He serves on editorial boards of multiple journals, including IEEE Communications Letters, IEEE Network Magazine, and Springer AJSE. He also served as chair, co-chair, and technical program committee member of multiple international conferences, including IEEE VTC, IEEE Globecom, IEEE ICC, and IWCMC.

Dr. Barry Solaiman is an Assistant Professor of Law (HBKU Law, Qatar) and an Adjunct Assistant Professor of Medical Ethics in Clinical Medicine (Weill Cornell Medicine-Qatar). He completed his PhD in law at the University of Cambridge. He is the Lead Project Investigator for a grant research project at HBKU and Project Investigator on another HBKU grant, both examining AI in healthcare. He is co-editing a major book on Health, AI and the Law with Glenn Cohen to be released in 2023. Recent Publications on AI •Barry Solaiman and I. Glenn Cohen (eds), Research Handbook on Health, AI and the Law (Edward Elgar, Forthcoming 2023) • Barry Solaiman and Mark Bloom, ‘AI, Explainability, and Safeguarding Patient Safety in Europe: Towards a Science-Focused Regulatory Model’ in The Future of Medical Device Regulation: Innovation and Protection (Cambridge University Press, 2022).

Dr. Mohammed Ghaly is professor of Islam and Biomedical Ethics in the Research Center for Islamic Legislation & Ethics (CILE) and adjunct professor in the College of Health and Life Sciences at Hamad Bin Khalifa University, Qatar. He is the founding editor-in-chief of the Scopus-indexed Journal of Islamic Ethics. He has a B.A. in Islamic Studies from Al-Azhar University (Egypt) and M.A. and PhD in the same specialization from Leiden University (the Netherlands). Islamic Ethics and its intersection with biomedical sciences and artificial intelligence belong to Ghaly’s main research interests. He is the Lead Principal Investigator (LPI) and research consultant of a number of funded research projects related to these research areas. He also published widely on these topics in both
Presentation 4: Lost History and overlooked Present: Mechanical and Artificial Intelligence in the Arabic culture

Abstract: The main aim of this presentation is to investigate how social needs and political circumstance inform the use of mechanical and artificial intelligence in Arabic culture. Precursors of AI are engendered in many significant inventions that appeared in the Islamic golden age as a response to human needs at that time. The religious activities and social events such as finding the Qibla, determining prayer times and the initial days of Ramadan and Eid were the impetus behind these inventions. For example, helping with ablution and reminding with the prayer time was the reason behind the invention of the first programmable humanoid robot by the Muslim inventor and engineer al-Jazari, who was known as the “father of robotics” in the 13th century. This mechanical kind of intelligence developed by social and religious needs enabled artificial intelligence in the modern age. Artificial intelligence is currently used in the Arabic postcolonial context to perform an activist role for the social good. AI Arabic literary writings use AI techniques and capabilities to enable literary activism and at the same time save authors’ lives living under authoritarian regimes. This study reflects on these two types of mechanical and artificial intelligence and the ideological and socio-political factors of their development.

Author bio: Dr Reham Hosny is an award-winning digital creative writer and a British Academy Visiting Fellow at the University of Cambridge. She is an Assistant Professor of digital literary studies and critical theory at Minia University and previously, she was a Lecturer at the University of Leeds, UK. Her interdisciplinary research focuses on investigating the cultural, social, and political contexts of Arabic and Anglo-American electronic literature and digital culture. Her work appears in peer-reviewed journals such as the Journal of Postcolonial Writing and her forthcoming book @ArabicELit: Electronic Literature in the Arab World (Bloomsbury) highlights new aesthetics and perspectives of electronic literature outside the Western electronic literature community. She is an editorial board member of academic journals such as the Journal of Digital Islamicate Research by Brill, and a member of various international research networks such as the Intersections, Feminism, Technology & Digital Humanities network (IFTe) and the Global AI Narrative (GAIN) network in the MENA region. Her co-authored novel, Al-Barrah [The Announcer] (2019, 2021), the first Arabic artificial intelligence novel, won the 2022 Robert Coover Award’s Honorable Mention, and her short story collection Amma Ba’ḍ [and thereafter] (2012) won the Ihsan Abdel Quddous Literary Prize for short story writing. Dr. Hosny is the first Arab and African to be elected as a director at the international Electronic Literature Organization (ELO). She is directing arabicelit, the first initiative focusing on globalizing Arabic electronic literature in English. Dr. Hosny is an invited speaker at many international conferences, workshops, and symposiums in different places around the world.
Chair Bio:

Dr Audrey Borowski is a Research Fellow at the Centre for Science and Thought in Bonn and an Associate Member of the Leverhulme Centre for the Future of Intelligence in Cambridge. She was a postdoctoral Research Fellow at the Munich Centre for Mathematical Philosophy as well as a research associate at the University of Oxford. She completed her doctorate (D.Phil) in the History of Ideas at the University of Oxford. Audrey’s research background lies at the intersection of philosophy, history and science and in the last few years she started working more closely on the philosophical history and philosophy of computing and AI. Audrey works on various projects pertaining to the philosophy of algorithms, the philosopher Hans Blumenberg, Myth and AI, as well as the analog and the digital. Her main project is entitled Philosophers of the Digital Age: A Philosophical History of Computing and AI from Leibniz to the Present.
Abstract: Bangladesh has been largely excluded in the conversations around the socio-technical and cultural implications of Artificial Intelligence (AI) despite observing a huge inflow of ‘intelligent’ systems and applications like the rest of the world in recent times. However, our decade-long ethnographic and design studies with various groups including faith-based communities, indigenous and rural populations in the country demonstrate (a) how the ‘secular’ proposition of AI is often refused and resisted (Interactions 2022, CHI 2022), (b) how AI prediction tools fall short in addressing the hopes and aspirations of the locals (CHI 2019, 2021) and (c) how the empiricist model of data denies the local histories, traditions, and identities (CSCW 2020). Based on these original and empirically grounded studies from one of the most marginalized territories in the map of global AI practices, we aim to organize a one-hour virtual interactive ‘Panel Discussion’ in this conference. We intend to discuss how ‘intelligence’ and associated ethics are conceptualized by the local communities and how those could be accommodated in designing scalable AI technologies and policies and contribute to develop alternative and competing values such as diversity and plurality in AI Ethics scholarship. Discussions will commence with three short case studies (7-minutes each) reflecting on the local and situated ‘knowledges’ around AI based on the diverse oral culture, myths, legends, wisdom, and visual and performative art practices in Bangladesh. Later, we will invite participants to engage with these communications by sharing their questions, concerns, and thoughts. We intend to publicize the session via relevant professional mailing lists, appropriate social media outlets, as well as personal connections with various research communities from diverse geographical regions. This panel discussion is our ongoing scholarly effort towards strengthening ‘AI Narratives’ from Bangladesh after successful similar initiatives in other venues, including, CHI, FAccT and Mozilla Festival.

Panelists:

Sharifa Sultana is a PhD Candidate at Cornell University, USA and Facebook Fellow. She conducts research in the intersection of HCI, ICTD, wellbeing, and feminist-HCI. She designs to address the challenges for the rural low-education population while accessing information. She actively engages with rural communities, local NGOs, and traditional healthcare support providers in rural Bangladesh to design computational tools and systems.

Mohammad Rashidujjaman Rifat is a Ph.D. Candidate in the Department of Computer Science at the University of Toronto and a Schwartz Reisman Institute for Technology and Society Fellow. His research is at the intersection of HCI, ICTD, and faith. Rifat conducts qualitative studies and computational analysis to explore faith-based values, rationality, and politics; and designs technologies to mitigate faith-based intolerance and make technologies more inclusive and just.

Dr Syed Ishtiaque Ahmed is an Assistant Professor of Computer Science at University of Toronto. He conducts research in the intersection between Human-Computer Interaction (HCI) and Information and Communication Technology and Development (ICTD). He received his PhD in Information Science from Cornell University in 2017. He established the first HCI research lab in Bangladesh in 2009, and still maintains it. His
research work is built around the concept of ‘voice’ that connects various branches of political philosophy to technology intervention. His current research focuses on the politics of faith and justification in computing.

Chair bio:

Abdullah Safir is a Research Assistant at LCFI, University of Cambridge whose current research interest lies in reimagining AI particularly from the Global South perspectives. He has recently finished his Master’s in Digital Media and Culture from the University of Warwick with prestigious Commonwealth Scholarship, achieved distinctions and was awarded for his academic excellence. His previous publications engage with design and development issues at the intersection of digital technologies and their implications on Rohingya refugees and Internally Displaced Populations in Bangladesh.
**Presentation 1: Praxis, or the Yugoslav Search for Man: Thinking and Human Self-Realization in the Age of Generative AI**

**Abstract:** In this talk I will propose praxis as an essential concept for the study of human-centered Artificial Intelligence. The importance of this concept, outlined by the Yugoslav Praxis school of philosophy in the 1960s and ’70s, is particularly evident in the analysis of the impact of generative Artificial Intelligences (GenAI) such as ChatGPT on human self-realization, thinking, and creativity. For the Yugoslav philosophers, emancipated from Stalinist thought and ideological censorship already in the 1940s, the central concern of Marx’s thought was man as a being of praxis, i.e., “as a being capable of free creative activity by which he transforms the world, realizes his specific potential faculties, and satisfies the needs of other human individuals” (Mihailo Marković). Unlike “practice,” usually understood in opposition to theory, praxis stands for human potentiality which, in certain adverse situations, may be impeded. Does GenAI present such an impediment? Debates about ChatGPT’s potential in the tech industry, business, and Anglophone universities have centered on its creative or destructive capabilities. However, what this technology makes clear above all is that the mastery over writing and form is not exclusive to humans and should not be the primary focus in the cultivation of the mind. The real issue at stake is, again, our ability to think, as Hannah Arendt wrote around the same time that the Yugoslav philosophers were developing their critical Marxist Humanism. With these propositions in mind, I ask: how does ChatGPT impede or, conversely, facilitate human self-realization? What happens when thinking, or the “soundless dialogue (eme emautô) between me and myself, the two-in-one” (Arendt) is externalized and relegated to an AI? I conclude by making specific proposals about what a New Praxis school of philosophy can contribute to the understanding of thinking as the essential feature of human self-realization in the age of GenAI.

**Author bio:** Ana Ilievska holds a Ph.D. degree in Comparative Literature from the University of Chicago (2020), and a BA and MA in Romanistik and Comparative Literature from the Eberhard Karls Universität Tübingen (2011, 2013). Prior to joining Stanford, she was Humanities Teaching Fellow in the College and the Department of Comparative Literature at the University of Chicago (2020–2021) and Adjunct Lecturer at the Università degli Studi di Catania in Sicily (2020) where she was also a Fulbright doctoral scholar. Currently, she is Andrew W. Mellon Postdoctoral Fellow at the Stanford Humanities Center as well as board member and membership secretary of the Pirandello Society of America.

**Presentation 2: Conceptions of Ethics in World-Making Machines: Colonial Iconographies of AI in Britain**

**Abstract:** This paper examines an intercultural dispute in which rival conceptions of ethics were mobilised to deploy cultural models of artificial intelligence in society. The public dispute in question took place in mid-c20th Britain between C S Lewis, a British colonist, and J D Bernal, an Irish colonial subject. Lewis forged a neomedievalist model...
whereas Bernal campaigned for scientific communism on the Soviet model. I argue that their conflict between fundamentally opposed conceptions of artificial intelligence and its regulation can only be understood through consideration of their cosmological worldviews: their rival conceptions of knowledge and ontological assumptions. On their own terms, they lived in different worlds. I draw on John Tresch’s studies of worlding with cosmograms and Arturo Escobar’s notion of the ‘Pluriverse’ to provide a methodologically symmetrical approach to achieve meaningful and productive comparison of their rival approaches to AI ethics. The paper responds especially to the questions pursued at ‘Many Worlds of AI’ under the ‘intercultural AI’ and ‘AI across borders’ themes. Lewis’s and Bernal’s rival programmes grew out of their education and experiences of growing up in Ireland, the British Empire’s most perturbing colony. Their biographies provoke consideration of their diasporic thinking about AI and global justice: Lewis descended from Ulster Scots settlers and Bernal descended from Sephardic Jews. I also show how their iconographies of intelligent machinery trained their audience to discern what counts as machine ‘intelligence’ and produce assent to their cultural model. Lewis forged what he called the ‘Medieval Model’ by subjecting readers to emotional training through the intelligent text-based ‘machine’ of medieval literary tradition, whereas Bernal co-opted the scientific methods of X-ray crystallography to explicitly visualise the living machinery which he argued must organise the whole of social life.

Author bio: Peter Rees was trained in natural sciences and history & philosophy of science at the University of Cambridge. His research on the history of science and ethics addresses cultural models of artificial intelligence. He is especially interested in how such models of intelligent machines and their place in society are deployed to co-produce emotion and political order. Currently he is working on a monograph provisionally titled ‘The Wars of the Human Machines’ which investigates how iconographies of intelligent machines were used in Cold War public polemics to produce humans and even forge worlds. He is also investigating the media strategies employed to marketise neoliberal accounts of the free-market as a ‘machine’ or ‘price mechanism’. Rees also has a background in neuroscience and biochemistry and completed laboratory research at the University of Edinburgh and Harvard Medical School. He is enthusiastic about interdisciplinary collaboration and is also a member of software start-up Cambridge BioNexus.

Presentation 3: Contentious Others: Logo and Dilemmas of Difference in the US, Britain, and France

Abstract: Recent historiography suggests that social and ethical concerns have long been central to AI research. This paper traces ambitions to develop socially responsible AI in the history of the Logo computer programme between the late 1960s and early 1990s in the US, the UK, and France. The main inspirations for Logo were Jean Piaget’s developmental psychology and Marvin Minsky’s decentralised theory of intelligence. For designers and educators, the hope was that Logo would enable the
flourishing of a diversity of learning and programming styles, thereby undermining the dominant culture within computer science. I suggest that distinct national imaginaries shaped understandings of the programme’s political and epistemological possibilities across the three cases. Logo was developed at MIT’s AI lab as part of a libertarian, anti-authoritarian education project, later recast as a tool to undermine patriarchy. In France, it was accommodated with republican ideals and welcomed as a first step in developing ‘informatology’, the study of people and computers across cultural differences. At Edinburgh, researchers emphasised the structural making of difference (esp. along gender lines) that limited the programme’s alleged revolutionary potential. This paper bridges AI historiography and the history of feminist thought. Studying how imaginaries of nationhood, debates about the aims and nature of AI, and conceptions of justice and otherness contributed to shaping distinct and partly contradictory ideals of computer emancipation in various national contexts, it helps to complexify AI historiography but also to recover alternative conceptualisations of moral development and subject configurations for current AI ethics.

Author bio: Apolline Taillandier is a postdoctoral research associate at the Leverhulme Centre for the Future of Intelligence and POLIS at the University of Cambridge, and the Center for Science and Thought at the University of Bonn. Apolline studied political theory at Sciences Po in Paris before joining the Max Planck Sciences Po Center on Coping with Instability in Market Societies, where she wrote her dissertation under the supervision of Prof. Jenny Andersson. During her PhD, she studied the history of contemporary transhumanism as articulating a set of projects about liberalism’s future. She was a Fulbright student researcher at the University of California, Berkeley Sociology Department in 2018 and a Cambridge Sciences Po visiting student in POLIS in 2019. In her postdoctoral research, she investigates the historical role of feminist thought and activism in the critique of computer technology and the remaking of artificial intelligence as a scientific project from the 1980s onwards. In the context of rising concerns about the discriminatory and stratifying effects of AI, she studies the transnational circulation of ethics and gender justice norms and their reinterpretation and appropriation by scientists and industry actors, focusing on European and U.S. American sites of technical AI research.

Chair bio:

Christiane Schäfer is the academic and administrative project coordinator for the team in Bonn and works at the Center for Science and Thought (University of Bonn). With a background in Comparative Literature and Eastern European Studies she previously worked as a research fellow at the Peter–Szondi–Institute for Comparative Literature of the Free University of Berlin and at the Institute of Slavic Studies of Potsdam University. She is the co-editor of Literatur und Arbeit (2018, with Maria Fixemer, Elisa Purschke and Vid Stevanović) and the online journal novinki.de as well as the author of several articles. Christiane’s research deals with adventurous narratives in the depiction of work and labor and with frontiers geographical as well as digital.
Presentation 1: Taking off with ease or Face-off with Justice? Mapping Digital Citizenship and ‘Ways of Seeing’ the Indian Biometric State

Abstract: In November 2022, the Planning Commission of India or the Niti Aayog released a discussion paper titled ‘Responsible AI for All’ for public comments introducing Facial Recognition Technology (FRT) at domestic airports. While the document awaited feedback, earlier this month the Indian civil aviation ministry brought out the biometric system named as DigiYatra (an acronym for Digital travel) app for domestic passengers travelling from New Delhi airport. The official website promotes FRT as the ‘future of air travel’ encouraging users for voluntary participation in the policy. With face as the boarding pass, people, data and technologies thereby, intersect with one another promising ‘a simple and easy’ travel. However, the Indian case pertaining to use of FRT by law enforcement has drawn criticism from human rights activists for targeting individuals belonging to marginalized backgrounds. This notwithstanding the fact that the Indian capital city with 1862.6 CCTVs per mile, is the site of greatest government surveillance project internationally. There are additional concerns regarding accuracy of algorithms and, given that there are no Indian laws regarding data protection, the official recommendations cited making a case for ‘Responsible AI’ (RAI) have further raised concerns of legal experts regarding privacy violations as well. Situated within this context, my paper reviews the policy formulations listed in of DigiYatra that describes the futuristic images of FRT as desirable AI. It ‘looks’ into how a gaze of camera categorically renders the body of citizen–subject as visual portrait in a digital database, while the visual persuasion furbishes a process of governance that the state deems as just, efficient and good. The rise of the AI-enabled ethnographic state thus, in this manner reveals how the ocular-centric capital of FRT pivots on the promise of an efficient digital future, imaging ‘ways to see’, positing itself as currency of modern state rule, and ‘law by other means’.

Author bio: Madhavi Shukla is a doctoral candidate studying at Centre for the Study of Law and Governance, Jawaharlal Nehru University at New Delhi. She works on law and visuality, law and society, and cultures of technologies for her research. Additionally, she is part of a study group on ‘AI and Facial Recognition Technology’ anchored at the Centre for the Study of Law and Governance that has shared its recommendation on the Niti Aayog discussion paper titled ‘Responsible AI for All: Adopting the Framework – A use case approach on Facial Recognition Technology.’ As a legal feminist she is passionate about feminist jurisprudence and imagines possibilities for equal futures for everyone. When she is not dreaming of equality, her other hobbies involve tending to her plants and playing with cats.

Presentation 2: Data power, AI and the “doubtful citizens”: The case of India’s National Population Register

Abstract: This paper analyzes how the National Population Register (NPR) project in India uses techniques of data power to shape the boundaries of citizenship rights. With religion becoming a shadowy marker of identity in framing “real Indians,” the
The present governmental regime in India actively seeks to identify the “doubtful citizens” within the national space by linking AI-powered biometric data and other forms of legal identification. The notion of the “real citizen” is reframed through biometric identification, documentation, and the bureaucratic web of this governmental reframing affects who will access essential services like opening a bank account or becoming part of the food distribution system. The entanglements of data infrastructures and systemic discrimination in this context are shaped through surveillance machinations, neoliberal governance, and the erasure of citizenship rights. In Documents and Bureaucracy, Matthew Hull emphasizes “the way documents link to people, places, things, times, norms, and forms of sociality” (Hull, 2012; p. 255). He notes that documents are often thought of as offering access to things and processes they document. However, they have broader capacities, particularly as they relate to administrative control and the active construction of subjects and socialites. This paper questions how the NPR in India seeks to re-construct socialities and rights using AI and data power. Through an examination of the discourses in the mainstream media, the rhetoric of the politicians of the ruling party, and the official government rhetoric, this paper explores how the logic of NPR engenders specific forms of statehood, governance, and citizenship. Using Foucault’s ideas of biopower and Agamben’s conceptualization of Homo Sacer, the notion of citizenship building in relation to legal documents and biometric identification is analyzed in this paper.

**Author bio:** Anirban is a Ph.D. candidate in Communications and Media at the Institute of Communications Research at the University of Illinois Urbana-Champaign. He is interested in the history of communication/media technologies, media and space, race and media, infrastructures of media, media policies, and critical information studies. He is intrigued by how the media affect cultural changes in the public sphere, build cultural citizenship, and (re)produce frameworks of everyday life.

**Presentation 3: Palmistry, Predictive Analytics and Imprints of Colonized Bodies**

**Abstract:** This paper aims to draw links between the ancient practice of palmistry or chiromancy and the colonial history of fingerprinting that became the basis of biometrics that are widely used to identify, control and surveil bodies through a critical cultural analysis of machine learning techniques. In 1858, William James Herschel, a British officer in the administrative services in Bengal, India got Rajyadhar Konai’s handprint as a testament for Konai to honour his road building material supply contract. Frances Galton’s study on fingerprints made use of finger and palm prints documented by Herschel in Bengal leading to the swift institution of finger and palm prints for identification purposes in administrative and legal areas. Although finger and palm printing have legacies of anthropometry, they also have a key place in biometrics. Moreover, current machine learning and data practices also inherit some of the methods established by Galton in the study of fingerprints. This way contemporary AI and data practices bear imprints of colonized bodies. On the other hand, palmistry continues to be widely practiced in parts of South and East Asia where the lines of a person’s palms are interpreted to predict the future. This aim common to the application of machine learning in predictive analytics has been explored by developing machine learning programs (including various mobile apps) using the
principles of palmistry to predict the future. My paper is interested in exploring the connections between machine learning programs, palmistry, predictive analysis and colonized bodies to two somewhat opposing ways. Firstly, to think of current practices in AI in the framework of a cultural practice such as palmistry to challenge the universalizing vision of AI. Secondly, cloaking AI in ancient palmistry practices as seen in the popular apps continue the extractive practices of historic colonialism.

Author bio: Charu Maithani is a researcher who organises her inquiries in the form of writing and curated projects. She is currently a sessional academic at UNSW, Sydney.

Presentation 4: The Digital Afterlives of Brahminical Colonialism: Biometric Surveillance, Facial Recognition Technology, & AI Ethical Complicities in India, 1858-2022

Abstract: This project is a historically-informed investigation of artificial intelligence (AI)-powered surveillance systems in modern day India from a decolonial perspective. In particular, I aim to trace the evolution of biometric surveillance in India from British-Brahminical analog schema relying on data such as the fingerprint to digital Hindu nationalist reincarnations powered by facial recognition technology. In 2021, Ameya Bokil et al published “Settled Habits, New Tricks: Casteist Policing Meets Big Tech in India,” examining the important rearticulations of Brahminical surveillance under emerging techno–capitalist regimes. Building off of such scholarship, this paper will posit a novel historical framework for linking the colonial origins of state surveillance in British India to contemporary facial-recognition technology concerns under the Hindu Right — focusing in particular on three strategies (1) the colonial history of surveillance technology informed by age-old Brahminical ideas of criminality (2) facialization as a site of biometric power under the British Raj (3) legal resonances such as the Criminal Procedure (Identification) Bill of 2022 that permit targeted surveillance of ghettoized communities in India today by the Delhi and Hyderabad police forces. Applying these perspectives to contemporary discourses, I wish to audit contemporary technical AI Ethical papers and frameworks from the Anglophone West in terms of their (in)ability to capture the violence of such Brahminical, colonial power and multidimensional social variables such as religion, caste, etc. How can we contextualize the emerging popularity of facial recognition technology in India, both as a continuation of colonial biometrics as well as an instrument of current Hindu nationalist/Brahminical regimes? How well do contemporary (Western) AI ethical frameworks capture these complex sociopolitical realities? What politics of solidarity does the current moment necessitate from us instead? I wish to intentionally attend to my own complicities and positionality within these matrices of power, as an upper-caste Hindu-raised class-privileged Indian-American with family ties to the Silicon Valley technology industry.

Author bio: Nikhil is a student at Harvard College Class of 2023 and A.B. Candidate in History & Literature (Modern World Track) and Computer Science, Secondary Degree in South Asian Studies.
Chair bio:

**Dr. (des) Maya Indira Ganesh** is Course Co-leader (MSt) of the Master of AI Ethics and Society programme at CFI. She is a media and digital cultures theorist, researcher, and writer who has worked with arts and cultural organisations, academia, and NGOs. She earned a Drphil in Cultural Sciences from Leuphana University, Lüneburg, Germany in 2022. Her doctoral work examined the re-shaping of what we mean by the ‘ethical’ and the shifting role of the human in the emergence of the driverless car. Maya’s dissertation investigates what initiatives for governance of such a complex technology implies for human social relations, spaces, and bodies. Prior to academic work Maya spent 15 years working at the intersection of gender justice, digital security and data privacy, and digital freedom of expression in a variety of global regions. Hence her work has consistently brought questions of power, justice, and global inequality to those of the body, the digital, and knowledge. She continues to be associated with feminist movements in Asia and works with arts and cultural organisations in Germany.
Presentation 1: A community-of-practice approach to understanding Chinese policymaking on AI ethics

Abstract: Extant literature has not fully accounted for the changes underway in China’s perspectives on the ethical risks of artificial intelligence (AI). Some of the ethical principles promulgated in Chinese policies on AI, such as privacy, fairness, justice, and inclusiveness, bear some similarity to those developed in Western countries, but they embody different connotations and philosophical assumptions in Chinese culture. This article develops a community-of-practice (CoP) approach to the study of Chinese policymaking in the field of AI. It shows that the Chinese approach to ethical AI emerges from the communication of practices of a relatively stable group of actors from three domains – the government, academia, and the private sector. This Chinese CoP is actively cultivated and led by government actors. This paper draws attention to CoP configurations during collective situated-learning and problem-solving among its members that inform the evolution of Chinese ethical concerns of AI. In so doing, it demonstrates how a practice-oriented approach can contribute to interpreting Chinese politics on AI governance.

Author bio: Dr Guangyu Qiao-Franco is Assistant Professor of International Relations at Radbound University and Senior Researcher of the ERC funded AutoNorms Project, University of Southern Denmark. Her research leverages practice theory, norm contestation, norm diffusion, and actor-network theory to interpret legal and foreign policy instruments developed by the Chinese government, as well as other developing countries. Her work has been published in International Affairs, The Pacific Review, International Relations of the Asia Pacific, and Policy Studies, among others.

Presentation 2: From Accuracy to Alignment: The Practical Logic of ‘Trustworthy AI’ among Chinese Radiologists

Abstract: The increasing use of machine learning algorithms to support human decision-making has brought about the popular notion of “trustworthy AI”. Accuracy and explainability, among other things, are considered two key elements in the trustworthiness of machine learning systems and are formulating ethical AI guidelines as well as major research efforts in computer science. The underlying assumption is that, if the output of AI systems is more “accurate” and “explainable,” then they become more trustworthy and trusted by users. Drawing on extensive participant observations and interviews with radiologists in China, this paper problematizes such universal assumptions and proposes an alternative, locally-rooted framework centered on “human-machine alignment” to understand AI trustworthiness. I argue that radiologists in China develop their trust based on the degree of alignment between their own judgment and the algorithmic output, including “direct alignment” and “adjusted alignment.” Regardless of the claimed performance indicated by statistical parameters, Chinese radiologists are still prompted to judge if the algorithmic decisions directly align with their own because of two factors. First, the probabilistic nature of evaluation metrics cannot guarantee algorithms’ correctness in individual cases in the clinical setting, for which typically no ready “ground truths” are available. Second, under current Chinese legal and regulatory regimes,
radiologists are held accountable for medical reports and are therefore motivated to doublecheck AI’s recommendations. Yet, even if the direct alignment is low, radiologists may still trust and use the algorithmic output if they can observe certain patterns of, and thus explain away, the misaligned algorithmic output. This leads to an “adjusted alignment” based on the radiologist’s own interpretations. In conclusion, the paper suggests that universal notions of accuracy and explainability are misplaced in conceptualizing and regulating trustworthy AI in the real world; instead, trust in AI is a result of human–machine alignment that is subject to social and institutional shaping and could not be reduced to some intrinsic technical features of the algorithms.

Author bio: Wanheng Hu is a Ph.D. candidate in Science and Technology Studies at Cornell University and a research fellow in the Program of Science, Technology and Society of the Harvard Kennedy School. At Cornell, he is also a member of the Artificial Intelligence, Policy, and Practice (AIPP) initiative and a graduate affiliate of the East Asia Program. His dissertation research examines the use of machine learning algorithms to cope with expert tasks, with an empirical focus on the development, application and regulation of AI systems for image-based medical diagnosis in China. The project has been supported by the National Science Foundation, China Times Cultural Foundation, and a Hu Shih Fellowship in Chinese Studies, among others. His research is broadly situated at the intersection of the sociology of expertise, medical sociology, critical data/algorithm studies, and development studies. Wanheng holds an M.Phil. in Philosophy of Science and Technology, a B.L. in Sociology, and a B.Sc. in Biomedical English, all from Peking University.

Presentation 3: AI Ethics and Governance in China: from Principles to Practice

Abstract: In the recent past, China’s government has recently taken remarkable steps at regulating Artificial Intelligence (AI). In March 2022, China’s regulation on algorithmic recommendation came into effect, breaking new ground internationally as regulators in several jurisdictions are beginning to approach the technical challenge of promoting algorithmic transparency and explainability. The Cyberspace Administration of China (CAC) is primarily focused on the role algorithms play in disseminating information, which is unsurprising given its censorship authorities. However, the regulator is also concerned with how recommendation systems impact consumers and shape labor conditions for platform workers. A separate regulation dated January 2023 addresses AI-generated content, such as deepfakes. With input from industry and research institutes, other branches of China’s government are pursuing parallel efforts, such as developing testing and certification methods for ‘trustworthy’ AI systems. Despite these developments, China’s AI governance efforts remain poorly understood abroad, particularly outside scientific circles. Justified outrage at the Chinese Communist Party’s use of AI for mass surveillance and ethnic profiling, as well as at the associated human rights abuses, has led to skepticism towards the ability of the country’s ethical and political tradition to produce responsible AI. In fact, misconceptions about the relationship between AI and social credit experiments in China have even made their way into the European Union (EU)’s draft AI Act. Against this backdrop, this working paper seeks to investigate emerging
approaches to ethical AI in China by studying the country’s first local regulations addressing AI, namely those issued by Shanghai and Shenzhen in the fall of 2022. Through an examination of original government documents, Chinese media coverage and local expert commentaries, which will be complemented by interviews with practitioners in a subsequent phase, this ongoing research aims to shed light on the actors as well as the ethical, political, and sociocultural forces that shape AI governance in China. Following an informed overview and analysis of the main Chinese AI governance developments at the national level, the paper will contextualize and examine the cases of Shanghai and Shenzhen to identify similarities and differences between emerging approaches to AI ethics and governance in China and in the EU, as well as any instances of—or opportunities for—mutual learning and intercultural dialogue.

Author bio: Rebecca Arcesati is an Analyst at the Mercator Institute for China Studies (MERICS) in Berlin, Germany. Her research focuses on China’s technology and digital policy and regulation. She covers the global footprint of Chinese tech firms, digital infrastructure and surveillance tools, governance of data and artificial intelligence, and Europe–China relations in the technology and innovation spaces, including tech transfer. Prior to joining MERICS, Rebecca gained experience helping Italian tech startups scale in China and as a research assistant in the UN Women China office. She holds an LL.M. in China Studies with a focus on politics and international relations from Peking University, where she was a Yenching Scholar. Rebecca received an MA degree in International Studies from the University of Turin and a BA in Language Mediation and Cross-Cultural Communication from the University of Milan. She has studied and worked in Beijing, Shanghai and Dalian, Liaoning.

Chair bio:

Dr Kerry McInerney (née Mackereth) is a Research Fellow at the Leverhulme Centre for the Future of Intelligence, where she researches AI from the perspective of gender studies, critical race theory, and Asian diaspora studies. She is also a Research Fellow at the AI Now Institute and in February 2023 will be a Visiting Fellow at the Sarah Parker Remond Centre for the Study of Race and Racialisation at the Institute of Advanced Studies, UCL. Previously, she was a Christina Gaw postdoctoral researcher in Gender and Technology at the University of Cambridge Centre for Gender Studies. Her scholarship has appeared in journals such as Feminist Review, Public Understanding of Science, and Philosophy and Technology. Her work on AI–powered hiring tools has also been covered by media outlets like the BBC, Forbes, the Register, and the Daily Mail. She is the co-editor of the upcoming collection Feminist AI (Oxford University Press), the co-editor of the upcoming volume The Good Robot: Feminist Voices on the Future of Technology (Bloomsbury) and the co-host of The Good Robot podcast on feminism and technology, which has received over 20,000 downloads to date. She has appeared on shows like The Guilty Feminist and the Radical AI Podcast, has been recognised as one of the 100 Brilliant Women in AI Ethics 2022, and has been shortlisted for the Champion of Women – Champion of Innovation (2022), Women of the Future – Technology and Digital (2022), and Women in Tech Excellence – Rising Star (2022) awards.
Presentation 1: Korean value of ‘Jeong’

Abstract: Developing ethical artificial intelligence has become a crucial problem, especially as advancements in machine learning lead to its increasing deployment across a broad spectrum of social and political processes. The frequent assertions of the independence of science from culture and religion have been widely debunked, and the social impact of digital technology makes this reality increasingly obvious. Drawing on religious and cultural values helps expose the lacunae in current approaches to robotics and AI, and creates opportunity to design AI for human flourishing. The Korean value of jeong offers a specific example of such cultural theology, and can be applied in the ongoing development of AI and related technologies. Jeong is a complex social phenomenon including empathy, solidarity, and mutual obligation. Making jeong a priority in the generation of new AI technologies will be relevant to the use of AI in human-human and, theoretically, human-robot interactions. The conjunction of theological, religious studies, and social AI approaches shows that ethical AI depends on more than the current focus on western philosophical ethics. If AI design incorporates human-human and human-AI jeong, the challenges of surveillance, algorithmic bias, and even hypothetical AI superintelligence become more manageable.

Author bios: Robert M Geraci is Professor of Religious Studies at Manhattan College. He is the author of several books, including Futures of Artificial Intelligence: Perspectives from India and the US (Oxford 2022) and Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality (Oxford 2010). His research has been supported by the US National Science Foundation, the American Academy of Religion, and two Fulbright-Nehru research awards. He is a Fellow of the International Society for Science and Religion.

Yong Sup Song is an Assistant Professor of Christian Ethics and Theology at Youngnam Theological University and Seminary, South Korea. His recent interests are focused on the ethical issues of Artificial Intelligence. As a Korean Christian ethicist, he emphasizes the priority consideration of the poor and the marginalized, and the inclusion of regional values in the development of AI. He is currently working on discovering and introducing cultural values in Korean society as regional values for moral AI. He has conducted three research projects on theology and artificial intelligence for the National Research Foundation of Korea.

Presentation 2: The Five Tests: Designing and Evaluating AI According to Indigenous Māori Principles

Abstract: As AI technologies are increasingly deployed in work, welfare, healthcare, and other domains, there is a growing realization not only of their power but of their problems. AI has the capacity to reinforce historical injustice, to amplify labor precarity, and to cement forms of racial and gendered inequality. An alternate set of values, paradigms, and priorities are urgently needed. How might we design and evaluate AI from an indigenous perspective? This article draws upon the 5 Tests
developed by Māori scholar Sir Hirini Moko Mead. This framework, informed by Māori knowledge and concepts, provides a method for assessing contentious issues and developing a Māori position. This paper takes up these tests, considers how each test might be applied to data-driven systems, and provides a number of concrete examples. This intervention challenges the priorities that currently underpin contemporary AI technologies but also offers a rubric for designing and evaluating AI according to an indigenous knowledge-system.

Author bio: Luke Munn is a Research Fellow in Digital Cultures & Societies at the University of Queensland. His wide-ranging work investigates the sociocultural impacts of digital cultures, from data infrastructures to platform labor and far-right radicalisation, and has been featured in highly regarded journals such as Cultural Politics, Big Data & Society, and New Media & Society as well as popular forums like the Guardian, the Los Angeles Times, and the Washington Post. He has written five books: Unmaking the Algorithm (2018), Logic of Feeling (2020), Automation is a Myth (2022), Countering the Cloud (2022), and Technical Territories (2023 forthcoming). His work combines diverse digital methods with critical analysis that draws on media, race, and cultural studies.

Presentation 3: What would an anti-casteist AI system look like?

Abstract: This paper fills an acknowledged gap in AI and wider digital research which positions ‘caste’ as a complex phenomenon and explores anti-casteist principles interpreted for AI contexts. This paper then seeks to distil an ontological interpretation of caste and to forward an anti-casteist ethical approach to the design, policy and governance of AI and algorithmic technologies. A conceptual indeterminacy and a subjective experiential nature add to the challenge of how caste can be interpreted within digital contexts. Caste cannot be explained either solely through the conceptualisation of race and racial biases, or, more markedly, through prevalent class-based analysis. Further caste is constituted of and experienced through largely communicative practices and contexts. We propose as first principles, going back to Dr. B.R. Ambedkar’s treatise on caste and its ‘mechanism, genesis and development’. From this and other literature, some main explanatory aspects can be identified and interpreted for AI contexts. We describe caste-based enclosures – by which we mean factors that allow for ideological, or functional boundedness formed in the (digitally mediated) real world or its virtual representations, that pre-suppose or permit caste-centric in/exclusion and caste homogeneity. Next is a preponderance of caste-centric purity in social relations, either implicitly or explicitly found and which can be queried in human contexts or digital content of AI systems. There is also an intersubjectivity of caste among groups and individuals which establishes differing positionalities, in/visibility of caste markers, and caste-originated power differentials, all of which can be questioned among the various stakeholders of AI systems. The paper maps the above conceptualisation to AI’s sociotechnical lifecycle of design, development and deployment. Through this, the paper argues for a practical, ex-post and ex-ante embedding of Ambedkarite ideals to present a framework for an ideologically anti-casteist mechanism for the ethical assurance of AI.
**Author bio:** As an engineer-turned-researcher in critical data studies, my research interests lie in developing an ethical and social justice-oriented view of emergent digital innovations and the technopolitical ecosystems they inhabit. My research interest includes algorithmic and data practices of gig work, digital identity, and fintech platforms. I am interested in policy and governance issues surrounding AI, particularly in the global South, and how this overlaps with contexts of labour and credit. Currently, at The Alan Turing Institute, I research and advise organisations on national and global ethical policy agendas for AI and how this can be directed by principles of data justice.

**Chair bio:**

Dr Alan Blackwell is a Professor in the Computer Laboratory, Director of the Crucible network for research in interdisciplinary design and co-Director of Research for the University of Cambridge Global Challenges Initiative. Crucible is a research network within and around the University of Cambridge. Its purpose is to encourage interdisciplinary collaboration of technologists with researchers in the Arts, Humanities and Social Sciences (AH&SS). The main focus of this collaboration is on design as a meeting point for widely differing research disciplines. Crucible activities include the establishment of new research programmes, training of researchers, input to policy bodies, and identification of suitable funding sources for research in interdisciplinary design. Crucible provides both a scientific and organisational framework for this research. Alan is also co-Director of Research for the University of Cambridge Global Challenges Initiative, the Cambridge Strategic Research Initiative for the Sustainable Development Goals. Alan has a long-standing mission to make meaningful contributions to the lives of the world’s poorest 3 billion people through the design of new technologies, including software that is directly programmable and customizable by end-users, and making IT accessible to those with a wide range of social backgrounds and abilities. Alan’s concerns for global equality were acquired with the egalitarian traditions of his native New Zealand, as a result of which he has often collaborated with technology-focused Cambridge NGOs such as Aptivate and Africa’s Voices Foundation. His motivation to contributing to the Sustainable Development Goals through the participation in the Global Challenges Forum is also shown by his previous role as Director of the EPSRC Ideas Factory “Bridging the Global Digital Divides”, and by the active publishing of research investigations of modern technology infrastructure in the light of universal human rights.
Panel 1: Relational Philosophies and Ethical Diversity in the Intercultural Evolution of AI Ethics: A 'Disruptive' Conversation

Co-organized with the Berggruen Center China at Peking University

Panelists:

1. Dr. Robin R Wang, Professor of Philosophy, Loyola Marymount University
2. Dr. Peter D. Hershock, Director, Asian Studies Development Program, East-West Center; Coordinator, Humane AI Initiative, East-West Center
3. Dr. Osamu Sakura, Professor, Interfaculty Initiative in Information Studies, The University of Tokyo; Team Leader, RIKEN Center for Advanced Intelligence Project (AIP)

Panel 12: The Ethics of Digitization in India

Co-organized with Ashoka University

Moderator: Prof. Subhashis Banerjee, Professor and Head, Department of Computer Science, Ashoka University and Jaspreet Bindra, Founder, TechWhisperer

Panelists:

1. Malavika Raghavan, PhD candidate in Information Systems and Innovation, Department of Management, LSE
2. Shirin Madon, Associate Professor of Information Systems, Department of Management, LSE
3. Amber Sinha, Senior Fellow Trustworthy AI at Mozilla Foundation studying models for algorithmic transparency

---------
**Presentation 1: Operationalizing decolonial AI through Ethics-as-a-Service**

**Abstract:** With more than 80% of papers published at AI conferences since 2018 attributed to authors in East Asia, North America, or Europe, efforts in AI Ethics risk being futile if they continue to fail to account for the cultural and regional contexts in which AI operates. Meanwhile, two concepts have garnered increasing prominence in the same time period. The first is decolonial AI, and the second is Ethics as a Service. Each has its own merits that offer needed contributions and improvements in the design and deployment of Artificial intelligence. Decolonial AI acknowledges the evolution of value and power, and leverages historical hindsight to explain patterns of said power that shape our intellectual, political, economic, and social world. Employing foresight, it provides tactics to better align research and technology development with established ethical principles, centering vulnerable peoples who continue to bear the brunt of negative impacts of innovation and scientific progress. Meanwhile, Ethics as a Service offers an on-demand customizable approach to examining AI development and deployment on a case by case basis, and in a manner that can satisfy both the agreed upon principles and the technical translational tools tasked to fulfill them. It does so by calibrating said tools in a balanced fashion so they are not too flexible (and thus vulnerable to ethics washing) or too strict (unresponsive to context). Our research connects the two concepts and offers a practical framework for operationalizing the foresight and tactics provided by decolonial AI when deploying Ethics as a Service. In doing so, our research first provides a global list of some of the most prominent regional and cultural values and a replicable methodology for sourcing and identifying more. Second, our research takes a given AI deployment scenario at hand and provides answers to the following questions: how to select the values (Western or not) that befit said scenario, how to interpret each value selected, and what’s the roadmap for operationalizing said value, via software tools, akin to existing approaches with more literature-popular values such as explainability and fairness.

**Author bios:** Daricia earned her Ph.D. in Human-Centered Computing at Clemson University. Her dissertation investigated alternative pathways for the design of justice-oriented safety countermeasures particularly for people in non-Western contexts. During her tenure as a graduate student, she has been fortunate to have been selected as Meta Fellow (formerly Facebook), Google Scholar, and a Trailblazer in research by the United Nations for her work on online safety in the Caribbean. Saif Malhem is the founding co-chair of the AI Future Lab: the largest global lab for millennials and Gen Z’s in artificial intelligence, built by members of the World Economic Forum’s Global Shapers Community. In 2022, the AI Future Lab launched the world’s youth AI manifesto at the International Telecommunication Union’s Generation Connect conference in Kigali, Rwanda. For his leadership in AI and climate technology, Saif was named one of Canada’s Top 30 Under 30 in sustainability in 2020. He is an engineering professional with experience in Fortune 500, nonprofit and start-up environments. Within the Global Shapers Community, Saif sits on the impact council and was one of the #Davos50 Global Shapers invited to attend the World Economic
Panel 13: Intercultural and Decolonial Approaches in Practice

Forum’s Global Shapers Community. In 2022, the AI Future Lab launched the world’s youth AI manifesto at the International Telecommunication Union’s Generation Connect conference in Kigali, Rwanda. For his leadership in AI and climate technology, Saif was named one of Canada’s Top 30 Under 30 in sustainability in 2020. He is an engineering professional with experience in Fortune 500, nonprofit and start-up environments. Within the Global Shapers Community, Saif sits on the impact council and was one of the #Davos50 Global Shapers invited to attend the World Economic Forum’s Annual Meeting in 2022. Saif has been a public speaker and public speaking coach for over 10 years and has spoken at a number of international conferences in India, Germany, and Canada.

Paul Sédille is a Belfer Center Student Fellow pursuing a joint degree between the Harvard Kennedy School and Stanford Graduate School of Business. Prior to this degree, he lived 10 years in China, working as a writer and videographer in Hong Kong and Beijing. His research interests cover China, media, and tech, from US-China relations to new media business models. He is a member of Global Shapers, where he has worked on digital literacy, public involvement in AI, refugee rights, and ocean conservation. Paul is a graduate of the Beijing Film Academy, Sciences Po Paris, and Sorbonne University.

Kathy Kim is a lead data scientist and data strategist with the Booz Allen Hamilton CTO Artificial Intelligence (AI) Integrated Management Team (IMT). She has extensive experience in engaging federal agencies on topics including data governance, policy, architecture, other specialized technologies, and national security. She previously supported the Millennium Challenge Corporation’s MCC-PEPFAR Data Collaboratives for Local Impact (DCU) program and also the Aspen Institute Philanthropy & Social Innovation Nonprofit Data Project as a William Randolph Hearst Fellow. Kathy received her bachelor’s degree in international studies from American University’s School of International Service and also the SIS Resonator’s Award for Outstanding Service upon graduation. She recently received a certificate the Center for Asian Pacific American Women (CAPAW) "Unleash the S(He)ro in You", a one–year long Women’s Leadership Program funded by the Walmart Center for Racial Equity. In her spare time, Kathy runs a 501c7 organization, WEF Global Shapers Community DC, and leads pro bono career coaching services for first-generation college students and BIPOC professionals.

Nupur Kohli is an award–winning healthcare leader, advisor and medical doctor. She previously worked as a medical advisor to the largest health insurance company in the Netherlands. Nupur is an appointed member of the European Health Parliament which aims to create a resilient European Health Union. She is also a supervisory board member for UNICEF Netherlands and member of International Advisory Board of Amsterdam Economic Board. Nupur aspires to make healthcare better, more efficient and accessible. Her expertise extends to building resilient health systems, social determinants of health, health equity, stress and productivity beyond. Nupur actively contributes to the World Economic Forum projects: Chatbots RESET, Generation AI: Developing Artificial Intelligence Standards for Children and a shared learnings platform on the transformative role of women and girls in health.
Presentation 2: Multicultural AI design and Ubuntu philosophy

Abstract: Much has been recently published on the core ethical values guiding policy frameworks on responsible AI. While many of these ethical principles form a common core or corpus of values shared between and across applications and locations, their realisation must be articulable through lenses that are relevant and appropriate to a particular context. Lenses, we argue, should be multicultural in formulation. One such multicultural lens is the traditional philosophy of sub-Saharan Africa of Ubuntu which prizes communitarianism and values ideas of humanness, co-operation, and reciprocity. The development and deployment of AI cannot be divorced from important socio-political, philosophical, and normative debates involving inclusion and diversity. While foreground values such as transparency, fairness, and justice give an appearance of consensus, their use is highly contextually-sensitive and application-specific. Informing an algorithmic outcome requires consideration of a multitude of normatively-relevant reasons – both quantitative and qualitative – and includes consideration of underlying prevailing values, interests, and duties. Reasons that provide a scaffold for building an ethical case and providing a legitimate pathway for explicitly selecting (either by justifying or refuting) the course of action an algorithmic system should follow. We argue that any account of meaningful embedded intelligence should include as part of the conversation previously marginalised, silenced, and under-represented voices in both establishing this common core of values and in articulating how these values find application. Not only is the system itself a network of functions, but it is one embodied and embedded within a broader holistic and connected functioning system of real life and is not independent from it. A real-world that is by nature and design multicultural. Thus, the system must not only account for, and follow, the rules of – and be integrated within – a cultural and societal domain but actively participate and contribute to it.

Author bios: Dr. Bev Townsend is a Postdoctoral Researcher at the York Law School at the University of York, UK and an Honorary Research Fellow at the University of KwaZulu-Natal, South Africa. Her expertise is in integrating law and ethics into safe and resilient autonomous systems (robots). Her research has focused on law, ethics, human rights, artificial intelligence, and governance.

Dr. Bongi Shozi is a Postdoctoral Scholar at the Institute of Practical Ethics at the University of California, San Diego and an Honorary Research Fellow at the University of KwaZulu-Natal, South Africa.

Professor Donrich Thaldar is an academic at the Law School of the University of KwaZulu-Natal, Durban, where he chairs the Health Law & Ethics Research Interest Group. He is the Principal Investigator of a research project on the legal aspects of the use of data science in health innovation in Africa, funded by the NIH. He also has a private legal practice, and has served as legal counsel in 13 reported cases.
Abstract: In computer vision AI ethics, a key challenge is to determine how digital systems should classify human faces in images. Across different fields, there has been considerable scholarly debate about normative guidelines that inform policy-making for facial classification. In our previous work, we have applied an experimental philosophy approach to investigate how non-experts and experts in AI deliberate about the validity of AI-based facial classifications [1, 2]. Our analysis of 30,000 written justifications using the transformer-based language model RoBERTa quantified the normative complexity behind classifying human faces. Experts and non-experts found some AI facial classifications morally permissible and others objectionable. We also found justificatory pitfalls that legitimized invalid facial AI classifications. These justifications reflected an over-confidence in AI capabilities, while others appealed to narratives of bias-free technological decision-making or cited the pragmatic benefits of facial analysis for specific decision-making contexts such as advertisement or hiring. Thus, contrary to popular justifications for facial classification technologies, these results suggest that there is no such thing as a “common sense” facial classification that accords simply with a general, homogeneous “human intuition.” However, cross-cultural perspectives have been missing entirely in this debate. In ongoing work, we add such missing cross-cultural perspectives working with collaborators in Japan, Argentina, and Kenya to extend this research project to an analysis of non-experts’ justifications of facial AI classification in these countries. We are curious to understand whether there are cultural commonalities and differences in the ethical evaluation of facial AI classifications. At the Desirable AI conference, we would present the quantitative and qualitative results of our cross-cultural study in Japan, Argentina, Kenya, and the US. This research supports critical policy-making by documenting cross-cultural perceptions and judgments of computer vision AI classification projects with the goal of developing ethical digital systems that work in the public’s interest.

Author bios: Severin Engelmann With a background in philosophy of technology and computer science, Severin is an ethicist focusing on the ethics of digital platforms and systems. Currently, he studies how non-experts in AI ethically evaluate AI inference-making across computer vision decision-making scenarios. In this research project, he also investigates whether and to what extent participatory approaches to AI ethics help advance the ethical governance of algorithmic systems.

Chiara Ullstein Chiara Ullstein is a Ph.D. student at the Chair of Cyber Trust. With a background in Politics and Technology, Chiara’s research explores public participation in the development and regulation of AI applications. Chiara applies both qualitative and quantitative research methods.
Chair bio:

Dr. Chelsea Haramia is Associate Professor of Philosophy and Director of Gender Studies at Spring Hill College, USA. She holds a PhD in Philosophy and a graduate certificate in Gender and Women’s Studies from the University of Colorado, Boulder. She is also co-editor of the journal 1000-Word Philosophy. Her recent research focuses on the intersections of science, technology, and values, and she is the author of several articles and book chapters on space exploration, astrobiology ethics, and the search for extraterrestrial technology — as well as public philosophy news articles and outreach. She has a chapter forthcoming in the Routledge Handbook of Social Studies of Outer Space and another forthcoming in the volume Reclaiming Space: Progressive and Multicultural Visions of Space Exploration, published by Oxford University Press.
**Presentation 1: Artificial intelligence as a decolonisation tool: Lessons from libraries, archives and museums**

**Abstract:** The decolonial turn in cultural institutions has shed light on existing power dynamics behind galleries, libraries, archives and museums (GLAMs) surfacing the existing biases and exclusions generated by Western forms of knowledge production. Cataloguing frameworks used to structure and document collections are modeled upon Western epistemologies. GLAMs’ widespread digitisation efforts and adoption of collection management systems and digital methods have supported preservation, accessibility, and research Artificial intelligence is increasingly used by these institutions to facilitate decision making tasks around documenting and cataloging activities as well as to improve user access to the information they hold about their objects. While AI can be perceived as an assisting tool it functions as a threatening one too. This perception holds relevance for GLAMs as postcolonial digital humanities and decolonial computing emphasize how technologies rehearse colonial dynamics (Risam 2021; Adams 2021). Many of these technologies are built upon classification systems and methods, such as statistics, used to control populations in colonial territories and racialized neighborhoods, and employ datasets that misrepresent non-dominant cultures with the use of derogatory terms. However, AI, like other digital tools, can be used as “technologies of recovery” (Gallon 2016) that unmask, repair, and remodel existing inequalities, biases, and other forms of colonial violence. For example, these projects use AI to identify discriminatory and problematic terms in documentation; tackle omissions of historical marginalised people in documentation; and seek to resurface hidden and forgotten objects. Through a series of examples of projects that use AI to decolonise museums and archives, this presentation will highlight strategies proposed by critical, postcolonial, and decolonial digital humanities that can be relevant to a wider AI community that tries to make AI fairer and more equitable – In particular, AI practitioners that are interested in developing systems that address issues around biased and problematic datasets.

**Author bio:** Dr Maribel Hidalgo Urbaneja a postdoctoral research fellow at the University of the Arts London working on the Worlding Public Cultures project and was research associate for LUSTRE, Unlocking our Digital Past with Artificial Intelligence, at Loughborough University. She obtained a PhD in Information Studies from the University of Glasgow. Her research interests span digital humanities, digital museum and heritage studies, digital narratology, and critical and decolonial approaches in digital humanities. Additionally, she has held positions in digital departments at The Getty in Los Angeles and the National Gallery of Art in Washington DC.

Dr Lise Jaillant is a senior lecturer in Digital Humanities at Loughborough University. She has a background in publishing history and digital humanities. Her expertise is on issues of Open Access and privacy with a focus on archives of digital information. She was the first researcher to access the emails of the writer Ian McEwan at the Harry Ransom Center in Texas. Her work has been recognised by a British Academy Rising Star award. Since 2020, she has led several international networks on Archives and Artificial Intelligence: LUSTRE Network, AURA Network, AEOLIAN Network and EYCon project.
**Presentation 2: Sharp Image, Vague Face: Disrupting the Facial Transparency in A.I. through a Diasporic Approach**

**Abstract:** Algorithm bias occurs when there is a lack of data diversity. A commonly adopted solution is to better an A.I. with data sets coming from minorities, implying a more general and severe process of data collection. Unfortunately, recent discussion of A.I. ethics fails to consider this procedure as a constant exposure of the marginalised, including the diaspora, and latent risks brought by inevitable watching and listening. Meanwhile, the potential of the diaspora’s elusive identity has gained scant attention when reflecting on a possible way to resist persistent contemplation. Based on this knowledge gap, my research revolves around criticising the compulsory transparency required by A.I. and imagining an alternative, nebulous A.I. ethics learning from the diaspora. First, my study elucidates how the data collection process colludes with colonialism photography as they both repel obscurity and uncertainty. Second, my research unpacks a poetic opacity that originates from the nomadic identity of the diaspora. Such ambiguity has the potential to contribute to the A.I. ethics focusing on marginalised communities and to resist top-down viewing. My study further uses the documentary Welcome to Chechnya as a case study. It argues that the obfuscation created by Deep Fake technology in this work not only protects the Chechen LGBTQ diaspora’s privacy and dignity but also gives rise to a chance to challenge the dictatorship and surveillance system. Last, my research articulates the potentiality of A.I. for the weak lies not in how accurate and transparent an algorithm can be but to which extent those people can retain their opacity and invisibility with A.I. in front of the viewing entangled with power.

**Author bio:** Yifeng Wei is an artist, curator, and PhD candidate in Visual Culture at the National College of Art and Design in Ireland. While reflecting on the legacy of cybernetics and system theory, Wei’s research interests lie in digital colonisation and emancipation, as well as resistance against algorithm bias and surveillance capitalism. His study relates current technological surveillance to the desire for certainty in cybernetics and system theory and manages to find a possible way of resistance by resorting to the aesthetics of opacity. Wei’s investigation of such aesthetics involves writing an alternative art history focusing on anonymous and incognito artists. Also, it touches upon the analysis of artistic practices that protect and liberate the oppressed by adopting nontransparent technologies, including the black box mechanism in artificial intelligence. Revolving around artists who apply such a nebulous approach to resist the power structure looming behind technology society, his recent curatorial practice, “The Cloud of Unknowing”, was shortlisted as a finalist for Hyundai Blue Prize Art+Tech 2023.

**Presentation 3: AI’s Colonial Archives**

**Abstract:** Generative AI technologies, such as text-to-image models, have received a lot of attention recently. As with all AI technologies, critics note these models enact various forms of simplifications, erasure and bias in their outputs. Yet, to understand the visual representations AI models produce (and to disrupt them), we must understand the politics and history of the canonical archives they build upon,
attending to the power differences and gazes that have historically been amplified in these archives. In this paper, we evaluate AI-generated images by situating them within broader global histories of cultural preservation and representation of visual archives. We specifically focus on the representations AI models generate of the global south. By juxtaposing AI-generated images about communities and practices from the global south with historic examples from visual archives we can trace the lineage of the racial, ethnic, gender and class narratives these models reproduce. We use critical scholarship on colonial archives, museum curatorial practices, and history of photography, to show how the visual archives underpinning AI models are sites of miscalcategorization, produced through an elite imperialist understanding of the “other” to perpetuate an orientalist gaze. Importantly, this gaze persisted in the cultural archives produced not just “by the west” but from within the south also, complexifying the question of where we can find an ‘inclusive’ archive.

Author bios: Rida Qadri is a Research Scientist at Google. Her research interrogates the overlaps between culture and AI. She is interested in the organizational, epistemological and geographical cultural assumptions underpinning the design and deployment of AI systems. She also studies the tensions and frictions that emerge when mono-cultural AI design choices are universalized through a global deployment at scale. Prior to joining Google she completed her PhD at the Massachusetts Institute of Technology in Urban Studies.

Huma Gupta is Assistant Professor in the Aga Khan Program for Islamic Architecture at MIT. Gupta holds a PhD in the History and Theory of Architecture and a Master’s in City Planning from MIT. Currently, she is writing her first book The Architecture of Dispossession, which is based on her research examining state–building through the architectural production of the dispossessed. Her broader research interests include the economic, cultural, and political relationships between discourses of architecture, development, and urban planning. Developing methodologies using sonic, visual, and other sensory archives to construct histories of subaltern spaces and subjects is of particular interest to her.

Fuchsia Hart is Iran Heritage Foundation Curator for the Iranian Collections at the Victoria and Albert Museum in London. She holds a BA in Persian, an MPhil in Islamic Art with Arabic, and is working towards the completion of a PhD at the University of Oxford, also in Islamic Art, with a focus on shrines in 19th-century Iran and Iraq.

Katrina Sluis is Associate Professor and Head of Photography & Media Arts in the School of Art & Design at Australian National University where she convenes the Computational Culture Lab. Katrina’s research is broadly concerned with the politics and aesthetics of art and photography in computational culture, its social circulation, automation and cultural value. As a curator and educator, for the past decade she has worked with museums and galleries to support digital strategy, digital programming and pedagogy. Her present work addresses the emerging paradigms of human–machine curation, as a contemporary response to the massive intensification of global image production and circulation.
Emily Denton (they/them) is a Senior Research Scientist at Google, studying the societal impacts of AI technology and the conditions of AI development. Prior to joining Google, Emily received their PhD in machine learning from the Courant Institute of Mathematical Sciences at New York University. Though trained formally as a computer scientist, Emily draws ideas and methods from multiple disciplines, in order to examine AI systems from a sociotechnical perspective. Emily’s recent research centers on a critical examination of the histories of datasets — and the norms, values, and work practices that structure their development and use — that make up the underlying infrastructure of AI research and development.

Chair bio:

Dr Miri Zilka is a Research Associate in the Machine Learning Group at the University of Cambridge where she works on Trustworthy Machine Learning. Her research centres around the deployment of algorithmic tools in criminal justice. Before coming to Cambridge, she was a Research Fellow in Machine Learning at the University of Sussex, focusing on fairness, equality, and access. Miri obtained a PhD from the University of Warwick in 2018. She holds an M.Sc. in Physics and a dual B.Sc. in Physics and Biology from Tel Aviv University. Miri was awarded a Leverhulme Early Career Fellowship to develop a human-centric framework for evaluating and mitigating risk in causal models, set to start in May 2022. She is a College Research Associate at King’s College Cambridge and is currently on a part-time secondment to the Alan Turing Institute.
Abstract: Language is key for communicating the values, norms and beliefs of a group of people. Research has shown that a lot of indigenous languages in Africa face the risk of extinction. It has also been noted that kids are best taught their mother tongue in their formative years. It’s evident from the various research and a preliminary exploratory study carried out that there is indeed a wide gap that exists when it comes to designing technology that can support kids to learn indigenous language. As a result, in this study we attempted to bridge the existing gaps by designing a platform that can support kids in Nigeria to learn a native language. As a first step, the proposal demonstrated how a prototype developed on the concept of gamification and constructivism pedagogy gave positive results such as fun in learning language, ability to improve critical thinking. However, on further analysis we realized that the NLP tool that supports the application doesn’t produce the native language using local accent. As a result, it negatively affects the learning of our kids to correctly pronounce Yoruba language. We therefore propose a native NLP tool that can help kids correctly learn indigenous or local accents. We propose using the Naive Bayes, which is an NLP model that gives the best accuracy, when compared with others such as Decision Tree and Random Forest Algorithm. We will incorporate the Yoruba language dataset with which we use to train the NLP model. Building the NLP pipeline step-by-step, we will integrate the app with tools that will help kids to correctly learn indigenous or local accents of the Yoruba language.

Author bios: Aderonke Busayo Sakpere holds a PhD in Computer Science from the University of Cape Town, South Africa. She is a faculty member at the University of Ibadan with about 12 years of experience in lecturing, research and mentoring. As a lecturer, she teaches both undergraduate and postgraduate students, with research interests spanning Human Computer Interaction (HCI), Data Privacy, ICT for Development. Dr Sakpere’s continued academic and research excellence has earned her various honors and awards. They include MIT Empowering The Teachers Fellowship, Selected Young Scientist by Lindau Nobel Laureate Meetings, Top 200 young researcher by Heidelberg Laureate Foundation (HLF), Hasso Plattner Institute (HPI) Fellowship, among others. She is the founder of Tech Girls Club which was set up to foster networking and interaction among females in technology. She recently served as a Consultant/Research Assistant with a digital innovation company named Qhala, on a Wikimedia Foundation Project.

Makuochi Samuel Nkwo is a Postdoctoral Research Fellow, Research Innovation and Development Unit, University of Namibia, Windhoek. He is also a lecturer in Computer Science with over 8 years of teaching and research experience. He has done high-quality and impactful single and collaborative research works in health & wellness, teaching & learning, eCommerce & workplaces, as well as in sustainable environments. Also, he has authored and co-authored, published, and presented over 26 peer-reviewed papers in reputable journals and conferences around the world. He is currently working in the field of Responsible and Human-Centered Artificial Intelligence and Ethics.

Abstract: This paper critically examines the use of AI-powered digital dialogue tools by United Nations’ peacemakers in Libya and Yemen. AI-powered digital dialogues claim to transform the peacemaking field as peace mediators deploy AI-powered digital dialogues to facilitate real-time conversations between thousands of individuals simultaneously (Alavi et al. 2022). Peace mediation teams inform and receive instant feedback from large numbers of civil society actors. As such, proponents of these innovations claim that AI-powered digital dialogue tools have the potential to speed up the signing of peace agreements without sacrificing principles of inclusivity (Bilich et al. 2019). Since June 2020, mediators of the United Nations have been using AI-powered digital dialogue tools in the Libyan and Yemeni peace processes. However, the evidence supporting such positive contributions of technological innovations to peacemaking efforts is thin (Richmond and Visoka 2021; Richmond and Tellidis 2020). Through qualitative interviews with individuals who participated in digital dialogue sessions, this study provides insights into the benefits, shortcomings, and risks of using digital dialogue tools in peacemaking processes. Drawing on the cultural turn in Peace Studies, the study critiques the techno-solutionist promise of AI-powered digital dialogue tools. Preliminary findings suggest that using AI-powered digital dialogues exacerbates the impact of digital divides, produces new privacy concerns, and violates conflict sensitivity principles and do-no-harm approaches. The application of AI-based digital dialogue tools in Libya and Yemen also signals that the penetration of techno-solutionism in the peacemaking field reproduces widely criticized colonial power dynamics between the global North and South, and between the global and local. Finally, the study shares critical reflections on whether and how AI-powered digital dialogue should be deployed to make peace processes more inclusive.

Author bio: Ali Altiok is a doctoral student in the joint Ph.D. program on Peace Studies and Political Science at the University of Notre Dame. His research focuses on political inclusion and the securitization of young people in the context of peacebuilding processes. Previously, he has worked as a researcher, policy officer, and coordinator for several international peacebuilding organizations in New York City. As an in-house research consultant at the United Nations Peacebuilding Support Office, he supported data analysis and narrative development of the Missing Peace: Independent Progress Study on Youth, Peace, and Security, mandated by the United Nations Security Council Resolution 2250. He co-authored the policy paper, We Are Here: An Integrated Approach to Youth Inclusive Peace Processes for the UN Secretary General’s Envoy on Youth. His work highlights what is transformative to youth agency in peacebuilding and social change. He is a former member of the United Networks of Young Peacebuilders. He holds an M.A. in Political, Legal, and Economic Philosophy from the University of Bern (Switzerland) and an M.A. in Peace and Security Studies from the University of Hamburg (Germany). He is originally from Turkey.
Abstract: This paper presents the preliminary findings of a study critically examining the use of AI-powered digital dialogue tools by United Nations’ peacemakers in Libya and Yemen. AI-powered digital dialogues claim to transform the peacemaking field as peace mediators deploy AI-powered digital dialogues to facilitate real-time conversations between thousands of individuals simultaneously. Peace mediation teams inform and receive instant feedback from large numbers of civil society actors. As such, AI-powered digital dialogue tools have the potential to speed up the signing of peace agreements without sacrificing from principles of inclusivity. Since June 2020, mediators of the United Nations have been using AI-powered digital dialogue tools in the Libyan and Yemeni peace processes. Yet, no study has evaluated the use of AI in the peacemaking context through listening to the experience of participants—a key measure of success. Through qualitative interviews with individuals who participated in digital dialogue sessions, this study provides insights into the benefits, shortcomings, and risks of using digital dialogue tools in peacemaking processes. Drawing on the cultural turn in Peace Studies, the study critiques the techno-solutionist promise of AI-powered digital dialogue tools. Preliminary findings suggest that using AI-powered digital dialogues exacerbate the impact of digital divides, produce new privacy concerns, and violate conflict sensitivity principles and do-no-harm approaches. The application of AI-based digital dialogue tools in Libya and Yemen also signals that the penetration of techno-solutionism in the peacemaking field reproduces widely criticized colonial power dynamics between the global North and South, and between the global and local. Finally, the study shares recommendations on how AI-powered digital dialogue tools should be improved to make the inclusion of civil society more meaningful in peace processes.

Author bio: Abigail Oppong is the lead/Co-Investigator for AI ethics at NLP Ghana, part of the 2023 list for 100 Women in AI Ethics from the Women AI Ethics due to her work in examining gender in AI Applications for African AI systems. She is an early career researcher focused on AI ethics in the sector of health and communication and an enthusiast about fairness and ethics in Natural Language Processing. She is a young Ambassador for Ariel Foundation International under the United Nations, focusing on issues affecting youths and women in Africa and making the voices of African Youths heard. Changemaker, Young Activist, and a Social Entrepreneur. As a member of the Ghana Natural Language Processing group. Her research project is based on mitigating gender bias in NLP applications. Recipients of the “Young Leaders Creating a Better World for All award” at the Women Economic Forum, 2019. She has made several AFI publications that seek to speak, influence, and impact life. She is part of the Board Members of Women Media and Changes working hard to ensure young women’s voices are represented well in digital media. She is also a Co-founder of Photo4her, an initiative that seeks to empower teenage mothers with photography skills. Currently, she is running a project called “The Level Her Up project” that aims to help female student entrepreneurs with the needed digital skills to thrive in managing their businesses while in their studies.
**Chair bio:**

**Dr Eleanor Drage** is a Research Fellow on the Desirable Digitalisation project, for which she is creating playbooks, games and tools to help AI ethics designers and project managers be responsive to AI Ethics. She is also helping companies across Europe respond to the EU AI act. Her other research investigates how AI relates to structural inequality, including systems of race and gender. Her work on AI-powered hiring tools has been published in Philosophy & Technology and covered by media outlets like the BBC, Forbes, and international news outlets. As part of this project, Eleanor led a team of computer science students in developing a tool that demonstrates the problematic logics behind AI-powered tools that claim to debias hiring by detecting a candidate’s personality from a video interview. Dr Drage is the co-host of The Good Robot Podcast, where she interviews top scholars and technologists about AI ethics, has appeared on popular shows such as The Guilty Feminist, and is a TikToker for Carole Cadwalladr’s All The Citizens’ data rights channel. She is the co-editor of the forthcoming edited collections Feminist AI with Oxford University Press and The Good Robot: Why Technologies of the Future Need Feminism with Bloomsbury Academic. Her other projects can be found at www.eleanordrage.com. She has presented findings to a range of audiences including the United Nations, NatWest, The Open Data Institute (ODI), the AI World Congress and the Institute of Science & Technology.

----------
Presentation 1: Imagining AI and a prospective Metaverse: A Participatory Speculative Design Case Study from Japan and reflections from Germany

Abstract: AI development and AI ethics are continually shaped by imaginations of future technology and narratives about socio-technical societies. The emergence and distribution of these narratives is limited with respect to both, how lay citizens and experts differently form and influence these narratives and how narratives from global north are more dominant than from global south. Local perspectives and the expertise of individual everyday life realities represent highly important factors for technology adoption processes, yet they are often neglected in development processes. In our presentation, we introduce a participatory speculative design framework for fostering an accessible AI discourse based on situated knowledge with the aim of contributing to the plurality of technology visions. Building on previous cross-cultural research about speculative perceptions of AI future societies and digital common good imaginations we present how the role and ethics of AI are imagined by Japanese citizens in a workshop on prospective Metaverse societies. Our approach enables sensemaking processes of and through technology design, revealing participants’ anticipations of pressing societal problems, imagined problem solving capacities of AI technology, and underlying societal value systems. The resulting visions discuss AI enhanced prospective democratic practice in the light of pressing problems of Japan’s silver democracy. Furthermore, future social relationships influenced by Japanese VR-culture and AI technology impacting the construction of identity, the preservation of Japanese food culture, and possibilities of natural disaster preparedness are discussed. These results show that the applied methodological approach allows to highlight how situated knowledge informs the imagination of future technologies. We conclude the presentation by illustrating how intercultural discourse can emerge from presenting the speculated futures as discussion starters to people in other countries, here Germany, and how the reflection on local visions can contribute to the creation and appreciation of futures taking into account local contexts.

Author bios: Michel Hohendanner is a research associate at the Munich Center for Digital Sciences and AI at the Munich University of Applied Sciences (MUAS), Germany. He is also a doctoral candidate at the Faculty for Industrial Design at the University of Wuppertal. Previously he was a visiting researcher at KEIO University in Tokyo, Japan, with Prof. Osawa and at the Kyoto Institute of Technology, Japan, with Prof. Mizuno. His research interests include the social impact of increasingly digitalized living environments and the role of design in these contexts.

Chiara Ullstein is a research associate and doctoral candidate at the Chair of Cyber Trust (Department of Informatics) at the Technical University of Munich (TUM). Previously she was a visiting researcher at KEIO University in Tokyo, Japan, with Prof. Osawa, and at the Kyoto Institute of Technology, Japan, with Prof. Mizuno. In her research she investigates citizen perception of AI and citizen participation in AI development and AI regulation. Her current cross-cultural studies focus on the perception of facial processing technologies and future visions of AI and the Metaverse.
**Presentation 2: Intercultural AI: Exchange, Dialogue, and Conflict**

**Abstract:** Students from low-resource language backgrounds, like those from the Global South region, are secluded within the AI system, they need an opportunity to express themselves and put forward their opinion on how best to come up with a more inclusive AI system that allows diverse cultural backgrounds and collaboration. AI systems developed for students from the Global South may not conform to the exact need of the students. In addressing the problem, we, the Imagine AI team based at the University of Colorado Boulder in the USA, collaborated with partners in Nigeria, Brazil, and South Korea to organize a series of writing contests for local high school students in those countries to write about AI. The students wrote short articles and stories about AI to compete for prizes. They wrote in their local languages to express their opinions about AI and make recommendations based on their desires. Our US-based team consists of international students who also came from Nigeria, Brazil, and South Korea to ensure an accurate and unbiased understanding of students’ writings. We analyzed the written stories collected across the countries we consider to be underrepresented within the AI system to get insight into the sentiment they share. Additionally, we surveyed existing applications that allow for collaborative learning and recommended possible areas of improvement. The pilot work we did in Latin America (Brazil and Mexico) was a success and hence a pointer to the viability of the scope of this project. The implications of this study show that students across Africa, Asia, and Latin America appreciate the potential and trends in technological advancements. However, the majority of responses from the students suggest the need for deliberate efforts to be channeled toward incorporating fairness and diversity while upholding ethical values in the AI system. Finally, we recommend the concept of participatory design and usability testing with identified end users in developing these systems.

**Author bio:** Cornelius Adejoro is a Ph.D. student at the University of Colorado Boulder USA, advised by Prof. Tom Yeh. Mr. Adejoro received a Master degree in Computer Science from the University of Nigeria in Nsukka. Mr. Adejoro has authored and co-authored several research publications in the area of human-computer interaction, big data, and cybersecurity, such as “Users’ Perception of Social Media Opportunities and Challenges” and “Contemporary View of Human System Integration in System Development Process.” After receiving his Master degree in Computer Science, rather than working in a large technology company, Mr. Adejoro wanted to put his skills and knowledge into serving underprivileged children and decided to take up a teaching job by joining Nigeria-Korea Model School where he taught children in basic school for six years. As a member of Black in AI, Mr. Adejoro is passionate about AI’s potential to transform education in Nigeria and across Sub-Saharan Africa. Given his current research interest, extensive computer science training, and deep AI knowledge, Mr. Adejoro is uniquely qualified to lead the project.
Presentation 3: Responsible Future-Making? Testing Intercultural AI Ethics through the Use of Generative Tools

Abstract: The recent proliferation of off-the-shelf generative AI tools such as Stable Diffusion had both intrigued and unsettled the artistic and academic world in equal measure. As part of an international consortium project in exploring how traces of the past shape policy development and society’s future, the working group on “Traces and Technology” developed a project to examine if generative AI tool can be responsible used to envision a more sustainable and fairer future. As part of the project, members of the working group (WG) were asked to visualize a more accurate and fairer representation of the world in a more sustainable future. Members of the WG were made aware of the fact that historic/ mainstream representations of the world (cartography), such as the Robinson and the Universal Transverse Mercator System, often reflect geographical distortion, and had long been viewed as associated with colonial emphasis on the “centre” and the “periphery”. The group, comprised of members from diverse cultures and backgrounds, ranging from artists, geographers, to aerospace scholars, produced vastly different visualizations. The group will next attempt to synthesize these images with the help of generative AI tool, such as Stable Diffusion, to see if the application of AI and generative software could be used responsibly to synergize diverse visions. The project seeks to put AI ethics to the test, with close monitoring of AI application through participatory decision-making. In doing so, it will not only document diverse conceptions of the “future” by professionals from various sectors and disciplines, but also to make a concrete attempt at synergizing diverging, and at times contrasting opinion of a fairer and more responsible vision for the future world.

Author bio: Dr Nikita Chiu is Senior Lecturer in Innovation Policy at the University of Exeter. She is also Ad Astra Distinguished Fellow in Robotic and Outer Space Governance at the Space Engineering Research Center at USC. Dr Chiu is a former recipient of the Ernst Mach Grant. Named after the Austrian physicist, the grant enabled the investigation of multilevel governance policy at the city, national, and international (UN) level. Dr Chiu is also a mentor for Space4Women, a network organized by the United Nations Office for Outer Space Affairs (UNOOSA) to support the UN Sustainable Development Goals. She was selected as one of “20 under 35” future leaders to watch by the Space & Satellite Professionals International in 2019 for her work on space sustainability. She is currently working group lead of “Traces and Technology” - part of an international consortium project that examines how historic traces (e.g. data, archives) could inform and shape the governance of future technologies in a more sustainable and responsible manner.
Chair bio:

As Associate Director and Senior Research Fellow at CFI, Dr Dorian Peter applies multidisciplinary expertise in technology research and design to the development and enhancement of CFI’s research projects, educational initiatives, communications, and strategy. She is also Deputy Chair of the CFI Strategy Group and Postdoctoral Affiliate at Trinity College, Cambridge. Dorian is a designer, design researcher and author who specialises in design for wellbeing, design for learning, and digital ethics in practice. She guided the launch of the Centre’s Master of Studies in AI Ethics and Society and is on the editorial board of the journal AI & Ethics. She is also a Research Associate at the Dyson School of Design Engineering at Imperial College London where she works with stakeholders to design technologies for physical and psychological health. Her books include Positive Computing: Technology for Wellbeing and Human Potential (MIT Press), and Interface Design for Learning: Design Strategies for Learning Experiences (Pearson). With over 20 years’ experience in technology design, she works together with users, engineers, and social scientists to co-create human-centred, context-sensitive and research-driven technologies in ways that respect psychological needs. She has also done work for non-profit and corporate institutions including Movember Foundation, IESO Digital Health, Google, Sony and Phillips. She received her PhD in Design from the University of Sydney and Bachelor’s from Carnegie Mellon. Her research in human–computer interaction includes development of the METUX model for evaluating the wellbeing impact of technologies within six spheres of human experience. Sitting squarely at the intersection of technology and the humanities, she often fulfils the role of cross-disciplinary bridge, helping colleagues in disparate fields work together through shared language and design tools. She also acts as a bridge between research and practice by translating academic discoveries into actionable knowledge for professionals.
Presentation 1: Occupying Urgency: How AI Solutionism Shapes the Narrating of Urgency around the Climate Crisis

Abstract: Climate data narrates the climate crisis and its urgency, which is crucial to mitigate its effects and urge for preventive and globally aligned actions. However, climate data is bound to powerful technologies like artificial intelligence (AI), and these are not neutral storytellers but powerful narrators. Lately, AI has fostered a new culture of prediction, but such culture does not only provide us with large-scale information over climate events but also manoeuvres between nudging us to urgency while also manifesting institutional and corporate power. We explore AI’s role in climate narratives that aim for establishing climate urgency (to mitigate the climate crisis). To do this, we position AI into a conceptual paradox between a technological mediator and an occupying dominator. On the one hand, we see AI as a mediator of urgency, one that is otherwise not representable. For instance, algorithmic modelling helps us explore the causality and dynamics of climate data globally, and AI can heighten space–time resolutions and clarify stochastic aspects. On the other hand, we see the role of AI as a dominator over (situated) urgency that uses climate narratives to implement Western techno-solutionism by claiming total power over global predictions. In the latter, AI stands for more than the attempt to mediate a global urgency but for an attempt to implement corporate powers and colonialist narratives into climate narratives. By exploring how AI moves between shaping, representing, and deciding over a climate urgency that is global or total, we will outline that AI acts out a complex multitude between being a representational and/or a colonising tool for climate stories and one for global and/or situated climate urgency. While we see the value in AI’s large-scale technology to achieve binding climate actions, we debate if AI can resolve this conceptual (and practical) conflict between mediating and dominating global climate urgency while aiming for a middle ground that is ethically justified.

Author bios: Eugenia Stamboliev is a postdoctoral scholar of ethics of technology and media at the University of Vienna. As a fellow in the WWTF project ‘Interpretability and Explainability as Drivers to Democracy’, she explores the political power of complex algorithmic models. Her work equally looks at the explainability of AI and political authority, and how to think dis/trust in platform labour.

Mark Coeckelbergh is a Professor of Philosophy of Media and Technology at the Department of Philosophy, University of Vienna since 2015 and was Vice Dean of the Faculty of Philosophy and Education until 2020.

Presentation 2: An Approach Based in Eastern Philosophy to Identify Ethical Issues in Early Stages of AI for Earth Observation Research

Abstract: AI and Machine Learning models have been used in Earth Observation (EO) and Remote Sensing (RS) research (“AI4EO research”) for decades to study and analyze the petabytes of data that would otherwise be almost impossible to process and understand. Ethical issues are taking center stage in this field of research as the
resolution of EO/RS data increases rapidly, as newer sources of data are fused to achieve better results at lower costs and greater speeds, and newer use cases of AI4EO research emerge. Nevertheless, not all ethical issues can be identified in the present – partly because of rapid technological evolution and almost blind focus on innovation as an end in itself, and partly because of uncertainties inherent in AI4EO research methods, analysis, and results. Real-world applications of research findings also give rise to uncertainties vis-à-vis ethical impact. Recent academic research and surveys conducted by the author suggest that AI ethics guidelines are not practically useful for many AI(4EO) researchers. Yet, ethically mindful choices at the early stages of research can help in conceiving, designing, and regulating AI, and in developing applications that are more acceptable to the global community. This paper recommends a novel approach to identifying and avoiding ethical issues in the early stages of AI(4EO) research, based on a combination of Eastern and Western philosophical thought. More specifically, it explores the imagery associated with Skanda, a mythological figure that appears in both Indian (Hindu) and Tibetan (Buddhist) philosophies to develop and implement a novel approach to ethical decision-making. This approach can be especially useful for scientists engaged in research and innovation with emerging technologies. Using a concrete example from AI4EO, the paper also describes a step-wise process and questionnaire based on this approach that can help researchers identify major ethical issues and opportunities in the early stages of their research.

Author bio: Mrinalini Kochupillai (Nalini) is a guest professor and core scientist at the Artificial Intelligence for Earth Observation (AI4EO) Future Lab at the Technical University of Munich (TUM), where she leads the ethics working group. She is also a faculty member at the Munich Intellectual Property Law Center (MIPLC) and an affiliated senior researcher at the Institute for Ethics in Artificial Intelligence (IEAI), TUM. Nalini has been a senior research fellow with the Max Planck Institute for Innovation and Competition (2014–2018), and with the Chair for Business Ethics, TUM (2018–2019). She has also been a Program Director (2014–2017) at the MIPLC, and an adjunct faculty at the EU Business School, Munich. She has over 15 years of experience in teaching and research in the field of business law, business ethics, intellectual property (patent) law, plant variety protection, and Incentive systems for sustainable innovations. Nalini obtained her B.A. LL.B (Hons.) degree from the National Law Institute University, Bhopal and an LLM. in Intellectual Property, Commerce & Technology from the University of New Hampshire. She completed her Ph.D. at the Ludwig Maximilian University (LMU), Munich as a full scholar and fellow of the International Max Planck Research School for Competition and Innovation (2009–13).

Presentation 3: AI for Datong: A Normative Framework for Sustainable AI

Abstract: In an early article, I proposed datong, a normative ideal in Confucianism akin to the common good, offers an alternative way to formulate AI for Social Good’s agendas; and, I called this approach AI for Datong (Wong 2021). More specifically, I argued that the idea of datong requires AI–based projects, if they are to contribute to the social good, to be (i) public–centered that they ought to be motivated and
justified by the good of the general public but not the interests of specific groups of individuals, (ii) care-centric that they are based on altruistic care for all but not aim for any personal advantages, and finally (iii) transformative that they should not merely attempt to prevent, mitigate, or resolve problems adversely affecting human beings and the environment but more fundamentally transform the individuals and social conditions such that the problems do not arise. The approach of AI for Datong, I contend, can also contribute to the discussion of sustainable AI (van Wynsberghe 2021). In addition to the (i) public-centeredness, (ii) care-centricity, and (iii) transformativeness, the idea of datong also accounts for—or, indeed, transcends—temporality and intergenerality. Together, they provide an account of our moral obligations to sustainability in design, development, and deployment of AI. In this paper, I first review the normative challenges related to sustainable AI, highlighting in particular the questions about moral obligations arise from the temporal and intergenerational gap in design, development, and deployment of AI (see, e.g., Halsband 2022; Robbins & Wynsberghe 2022). Next, I rehearse my approach of AI for datong and then make explicit its temporal and intergenerational dimension. Finally, I elaborated on how the approach of AI for Datong, with its temporal and intergenerational dimension, offers a unique perspective to answer the normative challenges related to sustainable AI.

Author bio: Pak-Hang Wong is a philosopher and ethicist of technology working in the industry, where he explores and addresses the social, ethical, and political aspects of AI, data, and other emerging digital technologies. Wong received his doctorate in Philosophy from the University of Twente in 2012 and then held academic positions in Hamburg, Oxford, and Hong Kong prior to his current position in the industry. Most recently, he co-edited with Tom Wang Harmonious Technology: A Confucian Ethics of Technology, where they provide an alternative, non-Western approach to the ethics of technology.

Chair bio:

Dr Tomasz Hollanek is a design and technology ethics researcher with a background in cultural studies, philosophy, UX design, and communications. His PhD work at the University of Cambridge, funded by the AHRC and the Isaac Newton Trust, is at the intersection of design theory, ethics, and critical AI studies, and focuses on the role of critique in the design process in general – and the development of artificial intelligence in particular. His research was featured or is forthcoming in peer-reviewed journals, including AI & Society and Design Issues, popular magazines, such as Aeon/Psyche, as well as an edited collection on intercultural perceptions of artificial intelligence (Oxford University Press, 2022). He has contributed to various research projects and initiatives, including the Global AI Narratives Project at CFI and the Ethics of Digitalization Research Sprint at the Berkman Klein Center for Internet and Society (Harvard). In 2021, he was also a visiting research fellow at the Observatoire des humanités numériques at the École normale supérieure in Paris (PSL).
Abstract: In light of the ongoing climate crisis and loss of biodiversity and natural environment, increasing hopes are directed to artificial intelligence (AI) technologies to counter and mitigate negative effects of human activities. This use of AI technologies for sustainability inadvertently leads to the question of how sustainable these technologies themselves are. As AI begins to affect more and more areas of human life, it must be scrutinized what sustainability means in different settings, for different communities, at different scales and for the planet. To address this multidimensionality of the relationship between AI and sustainability, AI needs to be conceived as a world object, not only affecting a small corner of the world, but the world as a whole. Hence, to focus on the relationship between AI and sustainability means focussing on the interplay between the local ecologico-socio-technical and the global ecologico-socio-technical impact of the design, development, use, and disposal of AI in different settings. Therefore, using AI for sustainable purposes is not a straightforward endeavour. It relates to questions from a variety of fields of inquiry and thus needs to be approached from an interdisciplinary angle. The panel participants will present and discuss how they conceive interdisciplinarity in regard to the relation between sustainability and AI. Each of the four participants work from within a specific research setting, namely healthcare, philosophy, environmental sciences, and political sciences, each with its own methodologies, to understand how AI relates to the idea of sustainability. During the panel discussion their insights will be shared and critically assessed with explicit attention to what the theoretical and practical implications are for the interplay between the local ecological-socio-technical and the global ecological-socio-technical impact of the design, development, use, and disposal of AI in different settings.

Panelists: Dr. Sebnem Yardımcı-Geyikçi is a Research Associate in the Institute for Science and Ethics at University of Bonn. Before joining the IWE, she worked as associate professor of politics at Hacettepe University, Department of Political Science and Public Administration. She completed her Ph.D. studies in Government at the University of Essex in 2013. She also has a second Ph.D. degree in political science from Bilkent University, received in 2015. Her articles appeared in a number of leading political science journals including Party Politics, Democratization, Government and Opposition, Political Quarterly. Her current research concerns the political impact of AI, party politics, questions of representation, contentious politics and digitalization of political parties. She was a Fellow at the Netherlands Institute for Advanced Studies (NIAS) and a Visiting Fulbright Scholar in the Weiser Center for Emerging Democracies at the University of Michigan. In 2022, she received the Young Scientist Award in the field of Political Science given by the Academy of Science in Turkey. Currently, she serves as a board member of Turkish Political Science Association (Siyasi İlimler Türk Dünyası Derneği, SITD) and a steering committee member of the European Consortium for Political Research (ECPR) Standing Group on Political Parties.
Tijs Vandemeulebroucke (PhD) is a postdoctoral researcher at the Sustainable AI Lab of the University of Bonn, Germany. He holds a Ph.D. in Biomedical Sciences (KU Leuven, 2019), and MA degrees in theology and religious studies (KU Leuven, 2013) and philosophy (KU Leuven, 2015). He researches the ethical tension between the use of AI in healthcare settings and the environmental, health, and social impact of the development and recycling of AI on local communities across the world. His research relies on philosophical-ethical approaches as care ethics, global bioethics, critical theory of technology, and empirical approaches as grounded theory. Tijs won the 2020 Doctoral Dissertation Award on Artificial Intelligence & Ethics jointly given by the Microsoft Corporation and the Pontifical Academy for Life. His work is published in journals as Science & Engineering Ethics, American Journal of Bioethics, Journal of Medical Ethics.

Larissa Bolte (MA) is a PhD student at the Sustainable AI Lab of the Institute for Science and Ethics at the University of Bonn, Germany, under the supervision of Prof. Dr. Aimee van Wynsberghe. She holds a MA degree in philosophy (University of Bonn, 2021). Larissa investigates the conceptual foundations of sustainability, considering both the notion's central characteristics and its normative implications. Within the Sustainable AI Lab, she examines how sustainability, construed as a theoretical lens, can inform AI ethics. Larissa has previously worked as a research assistant at the Institute for Science and Ethics and at the German Reference Centre for Ethics in the Life Sciences. She has also been a tutor for both moral philosophy and logic and basic research at the philosophy department of the University of Bonn. Larissa was a co-organiser of the Sustainable AI Conference 2021 and has published in the journal Sustainability.

Sophia Falk (MSc) is a PhD student at the Sustainable AI Lab of the University of Bonn, Germany. She holds a M.Sc. degree in Environmental and Resource Economics (Kiel University, 2022) and a BA in International Culture and Business Studies (Passau University, 2019). Her research focus is the environmental impact of AI hardware components throughout their entire life cycle (cradle-to-grave). Beyond that, the local environmental, health and social impacts of mining activities in resource-rich countries and electronic waste sites in Africa are of particular interest to her research. Her work is based on analytical methodological skills such as statistical-econometric methods, intertemporal resource management, sustainability economics, and climate policy.

Chair bio:

Charlotte Bander came to Bonn to study in 2013 and now works as a science manager at the University of Bonn after stops in Ireland and Maastricht. At the GRÜNEN in Bonn, she co-founded and headed the Feminism working group.
**Presentation 1:** Artificial Intelligence in National Media: How the North–South Divide Matters?

**Abstract:** This study addresses the issue of how discourses around Artificial Intelligence have been presented in national media. To explore this topic, I analyze news articles to identify narratives and imaginaries that contribute to building the concept of AI from a North–South perspective. To attempt to answer these questions, I have selected two different countries to gather the data from, one in the Global North (Sweden) and one in the Global South (Chile). However, both are located in the same “Large/Medium” cluster when combining land area and population. And both countries are known as innovative leaders in the utilization of new technologies. Drawing on data collected from 103 news articles, I found that in both cases, AI is presented as a positive tool for the development of local and global economies. Furthermore, AI is seen as driving the creation of exciting and disruptive businesses. However, my analysis shows that there is uncertainty about the future of the current status quo, both regarding the labor market and the current geopolitical power balance if China were to win the so-called ‘AI race.’ The data was coded and analyzed using a combination of critical discourse analysis and a data extractivism and the approach introduced by Sheila Jasanoff, the sociotechnical imaginaries. These perspectives can help to understand the relations between scientific and technological projects, and political institutions and power. Throughout the sample, the hegemonic (dominant) voice prevailed through discussions of the economy having a particular North–centric representation. This is important to explore as it can shed light on whether the AI is to provide real opportunities or if it is replicating the power relations of the globalized world. In this sense, the study also criticizes the sociotechnological imaginaries since, despite the fact that they propose a local view of power relations, they confirm that technological developments are often subject to global, political and corporate planning, regardless of the particular reality of each country.

**Author bio:** Claudia Wladimiro Quevedo. MSc Digital Media and Society from Uppsala University. Research Assistant Department Informatics and Media, Human–Computer Interaction (HCI) unit, Uppsala University, Sweden. She focuses on communications and how through new technologies and human centered focused work, a more empathetic society can be channeled.

**Presentation 2:** Cross–Cultural Narratives and Imaginations of Weaponised Artificial Intelligence: Comparing France, Japan, and the United States

**Abstract:** “Thinking machines” have long featured in popular culture, as Cave, Dihal, and Dillon demonstrated in their research on artificial intelligence (AI) narratives. Cave and Dihal identify several narratives related to AI and argue that these influence the perceptions, actions, and decisions of developers, political actors, as well as the public. However, existing research of AI imaginaries do not specifically cover
perceptions of weaponised AI. Moreover, few studies go beyond the perceptions of the English-speaking public. In this article we investigate narratives and imaginations surrounding weaponised AI technologies across different cultures. Our analysis is based on data from a public opinion survey conducted in France, Japan, and the United States in 2022–23. In a first step, we assess the extent to which publics in these three states are familiar with the narratives of weaponised AI determined by previous research such as the literature in media and literature studies, as well science and technology studies. Based on our survey data, we also identify alternative narratives of weaponised AI that go beyond existing categories. In this way, we contribute to understanding how cultural contexts and embeddings foster different imaginations of weaponised AI. In a second step, we address the question of whether, and if so in what form, such narratives shape public attitudes towards regulating weaponised AI. In the context of global discussions surrounding the prohibition of some forms of these technologies, examining public perceptions is essential. Imaginaries linked to AI and autonomous weapons, for example in science-fiction, and have been linked to political discourses and decisions in the sphere of security and defence. This article therefore also contributes to the literature seeking to understand if and/or how popular images of “intelligent” machines influence public opinion in relation to the regulation of AI and autonomy in warfare.

**Author bios:** Ingvild Bode is Associate Professor of International Relations at the Center for War Studies, University of Southern Denmark. She is the Principal Investigator of the European Research Council funded project AutoNorms, which investigates how autonomous weapon systems may change international norms. Ingvild is principally interested in analysing processes of policy and normative change, especially in the areas of weaponised artificial intelligence, the use of force, and United Nations peacekeeping. Ingvild’s research has been published in journals such as European Journal of International Relations, Review of International Studies, and Chinese Journal of International Politics and with leading publishers. Previously, Ingvild was Senior Lecturer in International Relations at the University of Kent and a Japan Society for the Promotion of Science research fellow with joined affiliation at United Nations University and the University of Tokyo.

Dr Hendrik Huelss is Assistant Professor of International Relations at the Center for War Studies, Department of Political Science and Public Management, University of Southern Denmark. He works at the intersection and knowledge frontier of international political sociology and AI with a focus on military technologies. His research and publication activities aim at producing critical knowledge on how AI influences the emergence and function of norms in international relations’ settings. Theoretically, he draws on different insights from critical security studies, STS, Foucault studies and IR theory. Dr Huelss publishes in high-ranking journals such as Journal of European Public Policy, International Theory; International Political Sociology and Review of International Studies.
Anna Nadibaidze is a Ph.D. Research Fellow in International Politics at the Center for War Studies, University of Southern Denmark. She is also a researcher for the European Research Council funded AutoNorms project.

Tom Watts is a Leverhulme Early Career Fellow at Royal Holloway, University of London (RHUL), leading a project on Great Power Competition and Remote Warfare. Previously, he was a Teaching Fellow in War and Security at RHUL (2018-2020) and a Graduate Teaching Assistant at the University of Kent (2014-2018). Tom graduated with a PhD in International Relations at the University of Kent in 2019. Tom’s research interests are in the field of International Security with a particular focus on American foreign policy, “remote warfare” and lethal autonomous weapons systems. His research has been published with Geopolitics, Global Affairs, the Bulletin of the Atomic Scientists, Drone Wars UK, and the Oxford Research Group.

Presentation 3: Responsible AI reporting requires cross-border collaboration

Abstract: AI-powered global platforms impact the lives of billions around the world. Yet the failure of the news media to include diverse perspectives, especially those of the global south, results in the exclusion of those best positioned to produce contextualized and nuanced stories. This also perpetuates the harm in those communities by treating non-Western communities as an afterthought. Unless we bridge this gap between the global reach of AI-powered technologies and the lack of inclusivity and collaborative-mindedness in the journalism industry, the much-needed global scrutiny for these platforms and positive change for the world will always be out of our reach. To illustrate this point, this presentation will highlight the exclusion of global south journalists from accessing documents from the 2021 Facebook leaks. Then we will compare the representation and inclusivity among reporters of the 2021 Facebook Papers consortium to that of other more successful journalistic cross-border collaborations led by the International Consortium for Investigative Journalism (ICIJ), Organized Crime and Corruption Reporting Project (OCCRP), and Forbidden Stories. By comparing the impact of each of these reporting projects, it becomes clear that collaborative and inclusive practices are required not just as a moral imperative but as a practical way to produce impactful, quality journalism on a global issue such as AI.

Author bio: Boyoung Lim is a Senior Editor and AI Network Manager at the Pulitzer Center on Crisis Reporting.

Chair bio:

Dr Jan Voosholz is the Scientific Director of the International Center for Philosophy NRW and Research Associate at the Chair for Epistemology, Modern, and Contemporary Philosophy Associated with the CST.
Presentation 1: Exploring Children’s Rights and Child-Centred AI

Abstract: When considering the responsible design, development, and deployment of AI technologies and the plurality of visions for technological futures, children are not only frequently missing from the conversation but are also on the receiving end of many data harms and injustices. The uncurbed development of AI systems with little consideration of human rights, needs, or preferences disproportionately affects children who have not been considered as part of the decision-making process, resulting in innumerable harms. Cases such as the Ofqual exam grading crisis in the UK, mental health chatbots mishandling children’s reports of sexual abuse, and smart toys selling data including children’s voice recordings illustrate some of the widespread harms to children that have gone unchecked and unaddressed. This important stakeholder group must be meaningfully included in the conversations surrounding the future of technological innovation in order for them and duty bearers to collectively steward a shared future for responsible AI and ensure that such harms do not persist. To address these concerns and operationalise child-centred AI decision-making, we examine whether and how children are considered across the AI project lifecycle. Through desk-based research, our project with UNICEF interviewing UK public sector organisations, and a 2-year research project with Children’s Parliament and the Scottish AI Alliance engaging with over 120 pupils across Scotland, we assess how children have or have not been engaged with AI and identify the challenges to incorporating children’s views regarding digital technologies. Given that many UK public sector organisations aspire to engage children but do not know how or where to begin, in our paper, we introduce various international frameworks involving AI and children and analyse the principles outlined. Sharing preliminary insights from our engagement work, we bridge the gap between those principles to pave a way forward for responsible, child-centred AI development across the globe.

Author bios: Dr Janis Wong is a Research Associate in the Public Policy Programme at the Turing. She has an interdisciplinary PhD in Computer Science from the Centre for Research into Information, Surveillance and Privacy (CRISP), University of St Andrews. Janis is interested in the legal and technological applications in data protection, privacy, and data ethics, where her PhD research aimed to create a socio-technical data stewardship and governance framework that helps data subjects protect their personal data under existing data protection, privacy, and information regulations. She also holds an MSc in Computing and Information Technology from the University of St Andrews and an LLB Bachelor of Laws from the London School of Economics.

Dr Mhairi Aitken is an Ethics Research Fellow in the Public Policy Programme. She is a Sociologist whose research examines social and ethical dimensions of digital innovation particularly relating to uses of data and AI. Mhairi has a particular interest in the role of public engagement in informing ethical data practices. Prior to joining the Turing Institute, Mhairi was a Senior Research Associate at Newcastle University where she worked principally on an EPSRC-funded project exploring the role of machine learning in banking. Between 2009 and 2018 Mhairi was a Research Fellow at the University of Edinburgh where she undertook a programme of research and public engagement to explore social and ethical dimensions of data-intensive health research.
Morgan Briggs Trained as a data scientist, Morgan currently serves as the Research Associate for Data Science and Ethics within the Public Policy Programme. Morgan works on a variety of projects relating to ethical considerations of data science methodologies and digital technologies at the Turing including continued, ongoing work on AI explainability, building upon the Turing and ICO co-badged guidance, Explaining decisions made with AI. She is also a researcher on the UKRI-funded project called ‘PATH–AI: Mapping an Intercultural Path to Privacy, Agency and Trust in Human–AI Ecosystems’ and an international project entitled ‘Advancing Data Justice Research and Practice’ which is funded by grants from the Global Partnership on AI, the Engineering and Physical Sciences Research Council, and BEIS. Morgan has continued to research topics related to children’s rights and AI stemming from research that was conducted with UNICEF.

Sabeehah Mahomed is a Research Assistant in Data Justice and Global Ethical Futures under the Public Policy Programme. Her current work includes researching and analysing the context of children’s rights as they relate to AI through a series of engagements and research into policy frameworks both nationally and internationally. Sabeehah holds a MSc in Digital Humanities from the Department of Information Science/Studies at University College London (UCL) with distinction. Her master’s thesis focused on Racial Bias in Artificial Intelligence (AI) and public perception, a portion of which was presented at the ‘Ei4Ai’ (Ethical Innovation for Artificial Intelligence) Conference in July 2020, hosted by UCL and University of Toronto.

**Presentation 2: Human First Innovation for AI ethics? : a Cross-cultural Perspective on Youth and AI**

**Abstract:** New and emerging technologies such as AI and robots introduce a range of risks and opportunities, locally and globally. Narratives surrounding the development of AI often seem to fall into a dichotomy between utopia and dystopia. The extent to which narratives are utopian or dystopian seems to vary by culture, with Japanese views in particular leaning more towards utopia, focusing on the potential societal benefits of AI, especially to cater to a rapidly aging population. By contrast, European and other Western narratives, exemplified by the image of “The Terminator” are typically dominated by fears; for example, that AI/robots will drive mass unemployment and inequality. In order to maximize new opportunities and minimize risks and create a better AI society, we need to understand AI use globally. Generation Z (GenZ: born between 1996–2010) will be the main beneficiaries and users, nevertheless, there are few studies which focus on youth and AI. This study introduces two of my on-going cross-cultural projects on youth and AI: “a future with AI” project in collaboration with the United Nations and “Project GenZAI” in the Moonshot R&D program. The latter project conducts large surveys and in-depth interview studies in six countries (Japan, China, Singapore, US, UK and Chile). Theoretically, this study extends the complexity model of communication (Takahashi, 2016) by exploring key dimensions of AI engagement. The aim of this study is to show universalism and
cultural specificities in terms of both opportunities and risks of AI/robots, to global understanding of an AI future where human happiness takes centre stage. Finally this study offers suggestions towards an AI future driven by “Human First Innovation”. AI has to be used towards achieving our sustainable future globally. But to do so, we must move from “AI first” and “nation first” to “human first” innovation.

**Author bio:** Toshie Takahashi is Professor in the School of Culture, Media and Society, as well as the Institute for AI and Robotics, Waseda University, Tokyo. She has been appointed as an Associate Fellow of the CFI, the University of Cambridge. She has held visiting appointments at the University of Oxford, Harvard University and Columbia University. She conducts cross-cultural and trans-disciplinary research on the social impact of robots as well as the potential of AI for Good. She is currently leading two projects on youth and AI. The goal of both projects is to contribute towards a vision of a future where human happiness takes centre stage. The first one is “A Future with AI” project in collaboration with the United Nation. She is also involved in “Moonshot R&D projects” by leading the “Project Gen ZAI”, engaging youths now for a global AI future in collaboration with the CFI, University of Cambridge, Stanford University, University of Chile, Pompeu Fabra University, Nanjing University, the National University of Singapore and others. Finally, Takahashi sits on the advisory committee of the Information and Communication Council, Ministry of the Internal Affairs and Communications, Japan.

**Presentation 3:** (Old) age in the age of artificial intelligence – crossing generational borders in AI research and development

**Abstract:** In the last few years, we have witnessed a surge in scholarly interest and scientific evidence of how algorithms can produce discriminatory outcomes, especially with regard to gender and race. However, the scholarly debate of fairness and bias in artificial intelligence (AI) has paid insufficient attention to the category of age (within the life-course perspective) and older persons as a socio-demographic group. Ageing populations have been largely neglected during the turn to digitality and AI and older persons were identified as potential “vulnerable data subjects” at higher rate of exclusion (Malgieri and Niklas, 2020). Ethical AI needs to cross the generational boarders that are currently ruling in both, the AI research and development of AI products and services. Perspectives of all demographic groups are fundamental to creating desirable AI for the future and older persons should not constitute yet another “dislocated community” in AI ethics. In this presentation, the concept of “AI ageism” is introduced to make a theoretical contribution to how the understanding of inclusion and exclusion within the field of data-driven technologies can be expanded to include the category of age. “AI ageism” can be defined as practices and ideologies operating within the field of AI which exclude, discriminate or neglect the interests, experiences, and needs of older populations and can be manifested in five interconnected forms: (1) age biases in algorithms and datasets (technical level), (2) age stereotypes, prejudices and ideologies of actors in AI (individual level), (3) invisibility of old age in discourses on AI (discourse level), (4) discriminatory effects of use of AI technology on different age groups (group level),
(5) exclusion as users of AI technology, services, and products (user level). Additionally, the paper provides empirical illustrations of the way ageism operates in these five forms.

**Author bio:** Justyna Stypińska completed her PhD on the topic of “Age discrimination in the labour market. A sociological–legal analysis”. Her research focuses on multiple forms of age discrimination and age inequalities in contemporary societies, especially in their most recent digital forms of late capitalism. In her newest project, starting beginning of 2023 and funded by the Volkswagen Foundation Germany “AI Ageism: new forms of age discrimination and exclusion in the era of algorithms and artificial intelligence” she will analyse with an international team (UK, Spain, Poland) the effects of use of artificial intelligence technology on the ageing populations in Europe.

**Chair bio:**

**Dr Rune Nyrup** is a Senior Research Fellow on the Science, Value, and the Future of Intelligence project. Before joining CFI, Rune did a PhD in philosophy of science at Durham University and has previously studied and taught at the University of Copenhagen. His research covers the role of different kinds of reasoning—e.g. analogical, explanatory, diagnostic—across the sciences, especially in relation to the question of pursuit worthiness, i.e., which hypotheses or models to prioritise for further investigation, and the proper role of ethical and political values in science. At the CFI, he is exploring these issues in relation to artificial intelligence, both as a field of research and as a technology.
Abstract: Representation in AI is intrinsically linked to value – what is represented is what is valued – and an expression of the world we want reflected in our technologies. But what does “representation” practically mean for AI systems? Similar to other terms in the responsible machine learning realm, e.g. AI “fairness” and “transparency”, we have a sense of it but the term is generally not well-defined enough to clearly operationalize best practices. We see representation in AI called for in news and opinion pieces, literature, AI guidelines, product aspirations, and proposed policy in both the EU and UK. There’s some consensus that broad representation is a net good (even necessary) element of responsible AI, yet the devil is in the details. For example, the notion that representative sampling – the creation of a set with instances proportional to that of a larger population – does not typically take into account the diversity within each group, applicability to the task at hand, or the inherently socio-political and intercultural nature of full representation. As practitioners, it is incumbent upon us to engage in participatory and justice-oriented techniques with distributed power mechanisms to allow for better representation of the diverse array of voices and needs in the design of our systems. In this workshop we will untangle notions of representation in evaluation datasets, and walk through case studies to gather hints towards best practice. Through discussion, we will uncover limitations of representation, e.g. the manners in which a mathematical or primarily technical notion of “representation” in sampling can be useful, but ultimately leave much to be desired. Together, we will engage in alternate visioning to understand what an equitable conception of representation could look like. In such a vision, how do the current practices for dataset curation and development need to change to shift power?

Author bios: Stevie (she/her) conducts sociotechnical research on DeepMind’s Ethics Research Team. She has a global, human rights focus in her work and typically investigates questions as to the impacts of technology on marginalised and in-conflict communities outside the US and Europe. Her current work is on the topics of data governance and representation, meaningful participation, and the effective alignment and evaluation of AI systems. Stevie’s research often has direct implications for both ML practitioners and policy.

Boxi (they/them) is a Responsible Development & Innovation Manager at DeepMind. They lead applied ethics projects in partnership with AI research projects / programs with a focus on language and multimodal generative models. This work involves working with teams to identifying sociotechnical harms and mitigations, developing appropriate ethics tools and frameworks and facilitating ethical review and governance. Boxi is also completing an MSc in Social Science of the Internet at the Oxford Internet Institute.
Abstract: Critical challenges arise when translating AI ethics principles into practice and intercultural contexts. With a crucial one originating around value tensions. Evidently not all desired values can be embedded in AI, as some might be conflicting like fairness and accuracy, individual benefit and collective benefits, transparency and privacy. These tensions need to be resolved in context and culture sensitive manners. However, how to identify which values to prioritize in AI development, and how to meaningfully speak to these differences? Simultaneously the design perspective for AI ethics is advancing. Ethics literature recognizes designers as important professionals as they cannot only provide technical means but also address values of people and society and create ways how to express them in material culture and technology (Van den Hoven, Vermaas, & Van de Poel, 2015). Also, studies show that design approaches can deal well with satisfying conflicting demands (Dorst, & Royakkers, 2006). This is a proposal for a workshop, in which we will use this design perspective to create provotypes as an artifact to meaningfully discuss differences across cultures in resolving value tensions for AI. These are artifacts/pictures that embody tensions in a certain context to explore new design opportunities (Boer & Donovan, 2012). With this workshop of ca. 1 to 1,5 hours, we will collaboratively explore the creation of provotypes that can embody value-tensions present between commonly used AI ethics principles and more local ones in an interactive manner. This will be guided and prepared with examples from industry by Dasha Simons (Managing Consultant Trustworthy AI at IBM). Real-time polling and a physical workshop format will be used. We invite all to join this explorative workshop!

Author bio: Dasha is passionate about bringing the human heartbeat into technology development by creating more trustworthy AI by design. She uses her creativity and human-centred perspective to find new ways how we can make AI more trustworthy, by making it more explainable, transparent and fairer. At IBM her role is twofold. On one hand internal, by enabling internal teams in creating trusted AI. Examples include: setting-up the global CoE for Trustworthy AI at IBM, leading the training initiatives on Trustworthy AI for consultants working in EMEA. On the other hand, she is advising various industries ranging from financial institutions, public sector and consumer goods from operational to C-level on trustworthy AI development. Dasha strongly believes a design perspective can support the current technological and policy feats in AI ethics. Her close collaboration with Delft University of Technology is focused on this, by exploring new methods and tools, providing educational support and brought to live in her role as Advisory Board member at the AI Futures Lab. She is a frequent speaker at events and conferences and co-authored various publications in the design field. For more information, please have a look at: https://www.designfortrustworthyai.com/about-dasha-simons
Presentation 1: Post-modern dance performance and a group conversation about responsible design and social impact of AI

Abstract: This is an artistic intervention. Building on the dance work that I am doing as a Fellow with the Edgelands Institute, I will present a post-modern dance, and I will guide a group conversation about responsible design and social impact of AI. The choreography has been built from the gestures of ordinary people (in the US, Europe, and South America) as they consider the impact of AI on daily life as well as the gestures of technologists as they strive to enact responsible design. (For more on post-modern dance, see Cunningham’s “50 Looks”.) The moves of the dance are suitable for every body; no special dance skills are needed and everyone regardless of age or ability can participate. Because dance is not a verbal means of communication, it transcends language differences. It provides an embodied means for people to experience the perspectives of other people. Dance – especially post-modern dance -- provides a unifying platform to discover different priorities and perceptions. Dance can alert people to other ways of knowing and sensitize them to other people’s concerns. Improving people’s awareness of experiences beyond their own is a foundational step towards more socially just AI design practices. After watching the dance, audience members in pairs/small groups will be asked to try on one or more of the moves from the dance. After doing the move, they will be asked to discuss what the gesture expresses and why someone might be feeling that way in conjunction with AI. Drawing on the theories and methods of dance anthropology, this work integrates concepts such as embodiment and kinesthesia in order to support more inclusive design practices in AI. As such, the project offers a new approach to addressing issues of social justice in AI design practices.

Author bio: An ethnomethodologist and artist, Dr. Betsy Campbell studies the practices of innovation teams. She is a Fulbright Specialist and the winner of two Academy of Management awards for her research and teaching related to social justice in innovation contexts. In 2022–23, she is a Fellow with the Edgelands Institute where she is using the arts to advance research on the responsible design and social impact of AI. Earlier in her career, Campbell founded high-tech companies, helped create a business unit within an established company (acquired by Lucent for $1.5B), and launched a 501(c)(3) (Harvard Alumni Entrepreneurs) which she grew to over 5,000 members worldwide before leaving to get her PhD. She also was the co-director for the MIT Community Innovation Lab. She is the author of several books -- including Practice Theory in Action and The Innovator’s Discussion -- as well as plays for the stage. She is an affiliate of the Penn State Arts and Design Research Incubator, an appointed director to the Harvard Graduate School of Education’s Alumni Council, and an active member of the Explorers Club. In the past, she has had a residency at Jacob’s Pillow, served as a visiting scholar at the Hastings Center, and been an invited participant in the Smithsonian Apollo Dialogues workshop.
Abstract: ΑΠΟαποικιοΠΟΙΗΣΗ (ΑΠΟαποικιοΠΟΙΗΣΗ), a word-play between the Greek words “decolonisation” (αποικιοΠΟΙΗΣΗ) and “poetry” (ΠΟΙΗΣΗ), is an artistic investigation on the topics of colonialism/decolonisation in Cyprus, in both historical and digital context. It is concerned with the compilation of a community-created corpus of written Cypriot Greek (CyGr), and the subsequent publication of a poetry collection, through the collaboration of 9 writers with a custom-trained Natural Language Processing (NLP) AI model, GPT-2. This presentation will provide an overview of the project’s process, challenges, and results, from consenting data collection practices, to co-authoring poetry with the CyGr NLP model. An important challenge was that CyGr is a non-standard language variety with no officially codified writing system and with limited language resources; to reduce “noise” in our data, we had to homogenise the orthography of the created corpus. Another challenge was the ethical dilemma of whether or not to allow open access to the corpus of a language variety that is globally spoken by a small population (therefore, not profitable) and that has been stigmatised by hegemonic narratives. The project aims to contribute to the de-stigmatisation of CyGr and the spreading of a more systematic homogeneous way of spelling. It also aims to provide technological agency to an island on the peripheries of the global AI infrastructure, where the legacies of colonialism are still evident. While Cyprus is currently at the early stages of adopting AI technologies, ΑΠΟαποικιοΠΟΙΗΣΗ attempts to sow the first seeds towards a framework of what a decolonised Cypriot AI could look like from a local perspective, and open a dialogue about the role and influence of AI in our lives. It also provides an opportunity for CyGr speakers to interact with technology in the language variety they use in their daily lives.

Author bios: Alexia Achilleos is an artist, with a background in fine art, archaeology and cultural studies. She is currently a PhD Fellow at CYENS - Centre of Excellence, undertaking artistic research on colonialism and AI, as well as a Research Associate at the Media Art & Design Research Lab (MADLab) at Cyprus University of Technology. Alexia is interested in the social, cultural, and political issues that impact narrative and power dynamics. By investigating data and utilising machine learning processes, often in a historical context, she aims to re-examine such issues found within history and society, but also within AI technology itself, particularly challenging the idea of technology as universal and objective. Her work has been exhibited at institutions and conferences such as NeMe Arts Centre, World Intellectual Property Organisation, Forking room/Post Territory Ujeongguk and NeurIPS.

Spyros Armostis is a Lecturer in Linguistics at the Department of English Studies, University of Cyprus. He holds an MPhil and PhD in Linguistics (Phonetics) from the University of Cambridge and a BA in Classics from the University of Cyprus. His publications lie mainly in the fields of phonetics, phonology, variational sociolinguistics, and clinical linguistics. His research activities have also expanded to other areas, such as grapholinguistics, language documentation and revitalisation and second language acquisition. Focal points of his work have been the study of Cypriot Greek and Cypriot Arabic as well as of other varieties, such as Standard Modern Greek, Pontic Greek, English, and French. Finally, he is the Coordinator of the Cy[Φ] Lab (Cyprus Phonetics Laboratory), Department of English Studies, University of Cyprus.
The NGO “Ipogia Skini”, began its activity in 2011 with nomadic type of cultural activities until 2014, when it found shelter in a rural space in the centre of Limassol. This space is called Synergeio and it serves as a basis for creating as well as producing the cultural projects of the NGO. Since its establishment, Synergeio has hosted all the theatre productions of the NGO Ipogeia Skini and its Performing Arts Festival “Sinaxis” for the past four years. Beyond its own Productions, Synergeio provides shelter to local and international artists/art groups who have a vision to create beyond the given structures, to research and experiment with new ideas and artistic trends. In addition, Synergeio is today a meeting point in the city centre. A space which is open to citizens’ initiatives for cultural and social action that promotes a form of art that evolves from solidarity as well as through the reflection of everyday life.

Presentation 3: Cultural Memory: Artistic Experiments in AI

Abstract: A series of artistic research projects exploring cultural (dis)connections across time and geography using AI and anti-colonial methods. As an Algerian and a creative technologist, Yasmine explores how she and other displaced peoples could reconnect with intangible heritage, reviving and contextualising shared cultural memories as well as building new collective approaches to AI practice. This presentation spotlights three of Yasmine’s current projects: 1) An Algerian Techno-Ritual How can auto-ethnographic visual art be made without reproducing the colonial gaze? Using AI as a digital witness, producing meaning from aesthetics while rejecting biometric data capture and surveillance culture through a set of AI-generated face filters that echo the aesthetic of Algerian face tattoos. 2) Mediterranean Hand Gestures This interactive project uses a camera to detect hand gestures and audio to produce corresponding verbal sounds. Yasmine was inspired to create a system that can preserve and demonstrate common Mediterranean hand gestures, capturing the meanings of these non-verbal expressions of emotion before they are forgotten. 3) AI Justice Matrix: The Futility of Policy Craft An online platform and collaborative authorship project that invites the perspectives of practitioners concerned with our relationship with technology, developed as part of Yasmine’s JUST AI fellowship at the Ada Lovelace Institute. Fundamentally, this project refutes the notion that effective policy making, as it relates to AI ethics, is at all possible. It acts as a critique of Euro-centric knowledge processes and the way they manifest as curated information flows passing though sanctioned knowledge keepers. It treats all sources and expression of knowledge as valid. It offers issues to consider when contemplating AI practice without necessarily offering an answer.
**Author bio:** Yasmine is a researcher and creative technologist. She is a fellow of the Royal Society of Arts and the Ada Lovelace Institute, researching anti-colonial ethics for artificial intelligence. She was named as one of '100 Brilliant Women in AI Ethics 2022'. She is currently collaborating with the CSNI at South Bank University and The Photographer’s Gallery on an Alan Turing Institute funded project on visual cultures and computer vision and has recently published a paper on ethnic disparities in higher education in England, which was featured in the Times Higher Education. She has researched and taught at universities in the UK and Sweden. Her creative and consultancy projects are listed on her website: https://yasmine-boudiaf.com/.

**Chair bio:**

**Dr Stephen Cave** is Director of the Leverhulme Centre for the Future of Intelligence, Senior Research Associate in the Faculty of Philosophy, and Fellow of Hughes Hall, all at the University of Cambridge. Stephen earned a PhD in philosophy from Cambridge. He then joined the British Foreign Office, where he served for nearly a decade as a policy advisor and diplomat before returning to writing and academia. He now holds the grade of Director of Research at the University of Cambridge (full professor equivalent). His research is mostly in the philosophy of technology, with two strands. First, he writes about the ethics of AI and robotics. His publications on this include the volumes AI Narratives (Oxford University Press, 2020), Imagining AI (Oxford University Press, forthcoming 2023), and Feminist AI (Oxford University Press, forthcoming 2023). Second, he writes about the ethics of life-extension and the philosophy of (im)mortality. His publications on this topic include Immortality: The Quest to Live Forever and How it Drives Civilization (Crown, Penguin Random House, 2012), a New Scientist book of the year now available in many other languages, and Should We Want To Live Forever (Routledge, forthcoming 2022). He has also written and spoken on a wide range of philosophical and scientific subjects, including in the New York Times, The Atlantic, The Guardian, and on television and radio around the world.
For further queries:
desirableai@lcfi.cam.ac.uk