

Dr. Matan Shapiro is a Social Anthropologist currently researching synoptic surveillance as part of the ERC-funded SAMCOM Project at the department of Digital Humanities, King's College London.

Title: The HAL 9000 Dilemma

HAL 9000 is the supercomputer in Arthur C. Clarke's and Kubrick's "*2001: A Space Odyssey*". Among other impressive functionalities, HAL is depicted in the movie as the ultimate surveillance machine, which not only monitor its human operators but also takes decisions based on what it sees and hears. At some stage, facing its prospective disconnection following a series of unexplained error in the spacecraft equipment, HAL decides to kill the human crew. What I call the HAL 900 Dilemma is pragmatic, rather than ethical or political – if HAL-like AI monitoring capabilities are becoming ubiquitous, how do we make sure that self-learning machines use their ability to surveill responsibly, in ways that benefit rather than threaten society? The aim of this panel is to think through this question from different points of view, involving both technological perspectives and humanities approaches to surveillance, AI, and the relationships between society and technology.

Prof. Luca Viganò is Professor at the Department of Informatics of King's College London, UK, where he heads the Cybersecurity Group. His research focuses on formal methods and tools for the specification, verification and testing of cybersecurity and privacy. He is particularly interested in formal analysis of socio-technical systems, where security depends intrinsically on human users, and in explainable cybersecurity, where, in addition to more formal approaches, he has been investigating how different kinds of artworks can be used to explain cybersecurity and how telling (i.e., explaining notions in a formal, technical way) can be paired with showing through visual storytelling or other forms of storytelling. Luca is also a playwright and screenwriter. His works have been published and produced in Italy, the UK and Russia. His first short film in English, "The First", explores a future scenario where the rights of sentient beings clash with freedom, identity and ethical judgment.

Title: What do we talk about when we talk about surveillance?

"Show, don't tell" has become the literary commandment for any writer. It applies to all forms of fiction, and to non-fiction, including scientific writing, where it lies at the heart of many scientific communication and storytelling approaches. In this talk, I will discuss how "show and tell" is actually often the best approach when one wants to present, teach or explain complicated ideas such as those underlying notions in cybersecurity. I will discuss how different kinds of artworks can be used to explain "surveillance", "AI", and "surveillance and AI", and I will illustrate how telling (i.e., explaining notions in a formal, technical way) can be paired with showing through different forms of storytelling. As concrete examples, I will discuss how fairy tales, popular films, TV series and other artworks can be used to explain surveillance and AI notions in such a way that they can be understood by non-experts.

Dr. Claire Dungey is a postdoctoral researcher at King's College London. She has a PhD in anthropology from Aarhus University in Denmark, and has conducted research about schooling, family life and children's journeys to school in various regions globally – such as in Uganda, Lesotho

and Germany. Her research interests include children and young people's mobilities, family monitoring and surveillance, as well as children's aspirations and engagement with education.

Title: 'When the product is free, you are the product': Surveillance, family life and critical perspectives on AI in Germany.

In Germany, data safety is often discussed in a national perspective, such as when smart dolls were seen as illegal espionage software in 2017. Based on fieldwork in Germany this paper explores how parents (often fathers) were critical of tracking technologies and preferred to pay money for specific devices seen as more secure rather than free apps that relied on selling personal data to third parties. They warned their children about artificial intelligence online, e.g. when their children received website recommendations. However, they still used AI such as Amazon Alexa to play music or communicate with family members in their homes at a distance, and asserted that this data stayed 'in the family'. The paper argues that parents' use of surveillance technologies despite their privacy concerns is closely related to ideas about convenience, family negotiations, and watching over as a form of care. I discuss how future anxieties about AI and data safety need to be seen in the context of the past, but also anticipations for an uncertain future, in which complete avoidance of smart technologies was not desired.

Dr Jennifer Cearns is an Affiliate Research Fellow of the Centre of Digital Anthropology at UCL and an Associate Researcher of the Alan Turing Institute, whose research addresses AI in the realm of the (inter)subjective. Her current project examines the impact of AI on human relations, and considers how AI might simulate empathy within the context of digital mental healthcare. She is author of *Circulating Culture* (University of Florida Press, 2023) and co-editor of *Contraband Cultures* (UCL Press, 2024).

Title: 'Through the Looking Glass?'

With the development of increasingly complex Large Language Models (LLMs) and their utilisation within an abundance of everyday technologies, users of these technologies (and the words they produce) are now placed within vast matrices of data that interpolate the Self with the Other. This short paper looks at how new forms of surveillance are emerging alongside LLMs, which move us beyond Foucault's classic formulation of the panopticon. Drawing on ethnographic research amongst users and developers of technologies that use LLMs, the paper proposes new ways of thinking about the relationship between Self and Other, and addresses the ways in which these emerging intelligent technologies disrupt ideologies of self and personhood that date back to the Enlightenment.

Sanja Mrkšić Kovačević is a PhD candidate at the University of Stavanger, Norway. She has background in Economics, Engineering Management and Risk Management and she is currently working on developing a risk science approach to managing and communicating uncertainty in the Big Data age. She is a PhD board member of the Society for Risk Analysis Europe - Nordic Chapter.

Title: Exploring the intersection of Artificial Intelligence and Online Surveillance: Insights from Architecture, Engineering and Construction sector

Artificial intelligence is finding greater applications, with varying effects depending on advancements in specific fields. However, the everyday application of AI raises a number of questions from a risk science perspective. In this regard, one of the major concerns is online surveillance and privacy. This concern encompasses the collection and use of data, privacy concerns, and the potential misuse of surveillance technologies. All this is even more amplified in light of the recently introduced European Union's Artificial Intelligence Act, which is praised as pioneering legislation worldwide. In order to obtain insight into how and where artificial intelligence may lead us in the future, we decided to examine the Architecture, Engineering and Construction (AEC) sector more closely. AEC constitutes one of the largest industries on a global scale, but it is infamous for its slow digitalization. We conducted interviews with professionals working in the AEC sector, focusing on the pace of AI adoption and taking a deeper look at the risks and uncertainties involved. Our findings show AI can be seen as a two-folded instrument, used to both prevent, but nevertheless to create, crises.

Matt Adams is the Head of Security Enablement at Citi, where he is responsible for enabling the secure adoption of disruptive technologies such as GenAI, DLT/Blockchain, and Quantum Cryptography. In this role, he also leads a team dedicated to identifying and developing use cases for leveraging these technologies to address cybersecurity challenges. Matt's expertise lies in the intersection of disruptive technologies and cybersecurity, and he is committed to driving innovation and enhancing security practices in the financial services industry.

Title: Online Privacy in an Era of AI Surveillance and AI-Powered Deception Techniques: Challenges and Opportunities for Responsible Innovation

Artificial intelligence is enabling powerful new forms of online surveillance by governments and companies. Machine learning allows the analysis of massive amounts of data to profile individuals based on their digital footprint across multiple online services and devices. AI can infer sensitive details about a person's life by analysing the content of their online activity. In response, individuals are starting to use AI agents to generate fake online activity and obscure their true digital footprint. Examples include AI chatbots that carry out realistic online conversations, AI-generated social media content to create misleading profiles, and automated web browsing to leave false interest trails. These AI deception techniques are an emerging form of privacy-enhancing technology to counter AI surveillance. However, they could also undermine the accuracy of legitimate AI systems and enable malicious deception if misused. An escalating arms race dynamic between AI surveillance and counter-surveillance might emerge, similar to cybersecurity. Distinguishing real from AI-generated actions online could become a significant challenge. Thoughtful regulation, oversight and technology governance will be needed to manage the far-reaching societal impacts as these technologies rapidly evolve. Public awareness and proactive policy approaches are essential to balance the benefits and risks.

Rob Morgan is a writer, designer, consultant and researcher across VR, AR, mixed reality and immersive theatre, and is founder and Creative Director of AR design studio Playlines. Rob writes and

narrative-designs award-winning immersive experiences, and has helped create story worlds and XR installations for some of the world's largest licenses, attractions and cultural institutions. He is a pioneer of Mixed Reality theatre, and as a game writer has helped create some of the most critically-acclaimed games in VR and interactive narrative. Rob is currently a Visiting Fellow at King's College London, researching and prototyping location-based digital narratives, and he is a founding director of the XRchiving conference on immersive technology for galleries, libraries, museums and archives. His upcoming book 'Augmenting Imagination: Storytelling for Mixed Reality and Spatial Computing' is due Sept 2024 from Routledge.

Title: Embodied surveillance, augmented ethics and AI

With Apple's entry into spatial computing, more and more people are experiencing digitised or augmented versions of reality. These users aren't just influenced by data on handheld devices: they're experiencing re-mediated realities through perception-altering overlays. The Augmented Reality market alone is estimated to be worth \$7.5BN by 2028. Our capacity to digitally customise, recontextualise, and even reskin reality is about to dramatically expand.

But AR/spatial technologies don't just change how we SEE: they're also WATCHING US as a fundamental part of their functionality. Recent studies show that immersive technologies can surveil, de-anonymise and emotionally profile their users – and nearby bystanders. Coupled with machine learning these technologies could be used by bad actors, as well as by the world's largest corporations, to build models of people and places with an unprecedented degree of accuracy and insight.

Not only that, but reality-augmenting technologies have immense potential for influencing user behaviour: coupled with AI content generation it's now easy to imagine some users choosing to live within divergent 'reality bubbles' which curate their reality in ideologically- or commercially-driven ways. Based on research for his upcoming book 'Augmenting Imagination', Rob will present an overview of the potential dangers of AI-powered reality-augmentation, and present some opportunities and mitigation strategies from a digital storytelling perspective.