Symposium on Information Controls

May 16th 2025, 8:30 am – 5 pm

MB4.2 Macadam building, Strand campus, King's College London

Hosted by the Department of Digital Humanities, King's College, London, with support from the London Arts and Humanities Partnership (LAHP).

PROGRAMME

8:30 – 9 am: **Registration**

9 – 9:15 am: **Welcome and opening remarks**

9:15 – 10:45 am: **Panel 1– Misinformation, Disinformation, Malinformation**

1. *Deception Analysis with Artificial Intelligence: An Interdisciplinary Perspective*
– Stefan Sarkadi (King's College, London)

Humans and machines interact more frequently than ever and our societies are becoming increasingly hybrid. A consequence of this hybridisation is the degradation of societal trust due to the prevalence of AI-enabled deception. Yet, despite our understanding of the role of trust in AI in the recent years, we still do not have a computational theory to be able to fully understand and explain the role deception plays in this context. This is a problem because while our ability to explain deception in hybrid societies is delayed, the design of AI agents may keep advancing towards fully autonomous deceptive machines, which would pose new challenges to dealing with deception. In this presentation, I build a timely and meaningful interdisciplinary perspective on deceptive AI and reinforce a 20 year old socio-cognitive perspective on trust and deception, by proposing the development of DAMAS -- a holistic Multi-Agent Systems (MAS) framework for the socio-cognitive modelling and analysis of deception. In a nutshell this presentation covers the topic of modelling and explaining deception using AI approaches from the perspectives of Computer Science, Philosophy, Psychology, and Intelligence Analysis.

2. *Information Control and Disinformation in East Africa: An Analysis of Digital* Dynamics in Burundi **– Steve Karake** (Decent Work for All Burundi)

Information control in the digital age is a central issue in East Africa, where states, digital platforms, and non-state actors shape the informational landscape. This paper explores the case of Burundi, highlighting the mechanisms of censorship, disinformation, and digital surveillance that influence access to information and citizen participation. The study relies on an analysis of digital media regulation policies, online surveillance practices, and strategies employed by citizens and journalists to circumvent censorship. It also examines how disinformation spreads through social networks and its socio-political consequences. Methodologically, the research combines a qualitative approach, including interviews with journalists and civil society actors, with an analysis of public policies and digital trends. This approach provides a better understanding of the tensions between information regulation and fundamental rights to freedom of expression and access to information. By focusing on the Burundian case and regional dynamics, this contribution enriches the reflection on underexplored geographies of information control and offers avenues for more balanced governance of digital spaces in Africa.

3. *THICK FAKES: Malinformation and the Future of Information Warfare –*
Hossein Derakhshan (King's College, London)

This conceptual article aims to define and expand on the category of 'malinformation,' a neologism coined by the author, as part of information disorder model (Wardle & Derakhshan, 2017): misinformation, disinformation, and malinformation based on two criteria of falseness of the information and the intention of their creators or publishers to harm or undermine another entity. Misinformation refers to false information without intent to harm, disinformation refers to false information with intent to harm, and malinformation (borrowed from French) is genuine information published with intent to harm someone. Notable examples of malinformation include revenge porn, doxing, leaks, etc. Drawing on a recent case of the Iranian state television's use of manipulated CCTV footage, this article proposes an expanded definition of malinformation as the manipulation of the context of an existing genuine piece of information. These contexts are: Space (e.g. cropping photos or videos, selective quotes, manipulated location labels or captions, etc.); Time (e.g. re-editing the order or speed of audio or video clips, altered date labels or caption, etc.); Domain (publicizing private information e.g. leaks, revenge porn, etc.); Agent (labelling genuine information as generated by AI or other); and Proportion (concealing the sample or population size when discussing scientific empirical research) The paper argues that much of what is called disinformation is in fact malinformation and a conceptual distinction between the two is necessary. It suggests that given the low cost and high speed production, as well as wide reach and deep impact, of malinformation, it is much more challenging and costly to tackle malinformation than disinformation. The article concludes with a call for a separation of dis- and malinformation and for more empirical research on its production, reception/resistance, and impact as well as technical and regulatory measures to tackle it.

10:45 am – 11:15 am: **Coffee Break**

11:15 am – 12:45 pm: **Panel 2 – State Controls**

1. *Algorithmic Governance and Postcoloniality: A Case Study of AI Traffic Enforcement Systems in Kerala, India* – **Ashwin Varghese** (Centre of Governance and Human Rights, University of Cambridge)

In 2021, the State of Kerala, in southwest India, embarked on a process of digital overhaul of the governance mechanism under its E-Governance policy, integrating ICT, big data, and AI technologies into essential state services, ushering in an era of Kerala's experiment with algorithmic governance. Drawing from fieldwork in Kerala, in this paper, I take a close look at the indigenously developed AI-powered automated traffic enforcement system. I highlight how states in their specific contexts in the Global South are adopting AI technologies to enable the development of particularistic information control systems. Launched in 2023, the technology is deemed successful in reducing road accidents, as well as increasing the state's revenue through automatic processing of fines for traffic violations. The E-Governance paradigm outlines a desire to reduce dependency on private enterprises for the implementation of AI systems in the State. In this paper, I critically analyse these claims by unpacking the operational power relations, dominant imaginaries and narratives, by tracing the development of the automated traffic enforcement system. In doing so, I outline the social, political, and economic conditions underlying the creation and implementation of this information control system. In this context, I note how the idea of E-Governance has evolved in the state to respond to concerns about data sovereignty, discretion, agency, and potential harms of AI. For this purpose, I use the framework of postcoloniality as it emerges in critical theory in the Global South to locate contemporary practices. I bring dominant theories of algorithmic governance and empirical realities in conversation with each other to develop critical perspectives on the production and practices of AI technologies and information control systems in the Global South.

2**.** *Information Controls in Sub-Saharan Africa: Digital Repression, State Censorship, and Resistance Strategies* **– Kehinde Adegboyega** (Human Rights Journalists Network Nigeria)

Governments across Sub-Saharan Africa increasingly deploy digital repression tactics—ranging from internet shutdowns to mass surveillance—to suppress dissent and restrict access to information. This paper examines information control mechanisms in Nigeria, Cameroon, and Ethiopia, analyzing how legal frameworks and cybersecurity laws are instrumentalized to silence journalists, activists, and opposition voices. Focusing on key incidents such as Nigeria's #EndSARS protests and Cameroon's Anglophone Crisis, the study highlights the political motivations behind state-driven censorship and its impact on democratic participation, press freedom, and human rights. It also explores the complicity of private sector actors in facilitating these controls. Beyond documentation, this paper assesses resistance strategies, including legal challenges, circumvention tools, and media-led advocacy campaigns. It examines the role of civil society organizations, regional human rights bodies, and international coalitions in countering digital authoritarianism and promoting digital resilience. By bridging academic discourse with practical insights from journalism and activism, this study contributes to the ongoing conversation on digital rights in Africa. It argues for stronger legal protections, coordinated advocacy, and policy reforms to safeguard online freedoms. The findings call for urgent action from scholars, practitioners, and policymakers to address the escalating crisis of information controls in the region.

3. *State-Controlled Typewriter Ownership: On The Poetry of The Unwritten* **– Mattia Natale** (King's College, London)

Whilst much critical attention has been devoted to practices of state control on media and information through the Internet, limited critical attention has been paid to practices of control of physical machines that enable such production as opposed to the tools that enable its divulgation (the Internet, xerox, mimeographs, printed media). By focusing on authoritarian regimes in the Eastern bloc, in which control over means of literary production was exercised by the state (as well as its distribution), a critical space opens up for understanding the dynamics of written-unwritten text in direct relation to the physicality of the means of production of a given text. This paper aims to explore literary practices of concrete poetry in Eastern-bloc countries that enforced a control of typewriter ownership in the second half of the 20th century. Concrete poetry serves as a particularly apt literary genre to analyse in this context as its production is inherently tied to typewriters, as this was one of the main means of composition of concrete poems. As a movement that aimed to re-evaluate the nature of language through its most basic written (and, sometimes, verbal) components, its authors from across the world used mathematical principles of combinatorics (variation, permutation, combination) or explored the newly emerging nature of computational coding, the typewriter offered a typographical 'ground zero' for these linguistic explorations; in an environment where typewriter ownership is under state supervision and control, a machine like the typewriter finds itself to fulfill a different function – in which both what is left on paper as well as what is left untyped acquires a deeply personal meaning.

12:45 pm – 1:45 pm: **Lunch**

1:45 pm – 3:15 pm: **Panel 3 – Infrastructures**

1. *Russian Internet Infrastructure in the Age of Digital Sovereignty and Infrastructural Coercion: The case of TSPU* – **Dmitry Kuznetsov** (University of Amsterdam)

My paper examines Russia's evolving internet governance strategy in the wake of its full-scale invasion of Ukraine, focusing on the implementation of Technical Measures to Combat Threats (TSPU). While the Russian state has long engaged in decentralized control and indirect censorship, the post-2022 period marks a shift toward infrastructural coercion—an explicit, crisis-driven intervention to reshape information flows. Building on Maxigas and ten Oever's (2023) work on infrastructural ideology, I introduce infrastructural coercion as a means by which can states accelerate socio-technical transformations through direct and centralized intervention in technical structures. This framework is applied to analyze Russia's rapid deployment of TSPU, a Deep Packet Inspection based system formalized in the 2019 "Sovereign Internet Law" but rapidly implemented after February 2022. The transition from indirect influence to explicit control reflects the regime's struggle to maintain hegemony, replacing previous strategies with coercive mechanisms that minimize uncertainty and ensure compliance. Using a discourse-historical approach, this study examines legislative texts, government directives, and industry discussions from the Conference of Russian Telecom Operators and Data Centres (КРОС) (2018–2024). The findings reveal how key industry actors, initially resistant to stringent control measures, adapted to new state mandates. While early discussions ridiculed the "sovereign internet" law, post-2022 discourse reframed wartime restrictions as necessary adjustments, highlighting the growing normalization of infrastructural coercion. This research advances scholarship on digital sovereignty by shifting focus from state policy to industry adaptation, emphasizing the role of infrastructural actors in mediating coercive state interventions. The case of Russia underscores how globally available technologies like DPI can facilitate rapid centralization of decentralized networks, raising concerns about the broader implications of infrastructural coercion in both authoritarian and democratic contexts.

2. *Societal Foundations of Cryptography* – **Martin R. Albrecht** (King's College London) and **Rikke Bjerg Jensen** (Royal Holloway University of London)

"Encryption is deeply threatening to power" (Whittaker, 2024) and "Cryptography rearranges power: it configures who can do what, from what" (Rogaway, 2015). These are two examples of a broader, and widely accepted, idea in the field of cryptography: that cryptography is in conflict with power. This assumes that cryptography is, or at least can be, a technology that limits power; a tool in the toolbox of resistance against overreach by an authority. Using concrete examples, we will explain that this is an incorrect characterisation of cryptography. Rather, cryptography -- a central pillar of privacy guarantees -- fundamentally relies on and presumes power. We show how this premise of power is built into fundamental definitions of the field, not just in its practice. Put succinctly: to speak of a security notion such as cryptographic confidentiality means to speak of power and indeed violence in an immediate, non-metaphorical sense. This does not mean cryptography is wrong: its assumptions about power hold but they mark cryptography as belonging to a particular kind of society.

Concretely, we will cover (1) the presumption of astronomical computational budgets and the absence of violence; (2) self-restraining nation-state adversaries (cryptography is legal); (3) what explains the presence of such adversaries. Depending on the time available, we would cover a subset of these.

3. Examining Organised Breakdowns of the Internet as a Means of Information Controls – **Gowhar Farooq** (King's College, London)

In this paper, I introduce the concept of 'organised breakdowns' -- the deliberate and systematic disruption of infrastructures as a form of information control. I focus on infrastructural breakdowns beyond technical failures and accidents and engage with them as a well-thought-out means of information regulation and control. Using the Internet as a case study, I demonstrate how securitisation logics are often used to justify and enable organised breakdowns and, in turn, controls over information circulations. Unlike spectacular attacks or disruptions that catch attention – and make infrastructures significantly visible – organised breakdowns of infrastructures, I argue, also involve mundane governmental techniques and procedures. These methods, operating through technological control and socio-political means, can be normalised through repetition and bureaucratic procedures to make them a part of everyday life for those at the receiving end of such mechanisms. Apart from challenging the perceived nature of infrastructural breakdowns in science and technology studies (STS), this paper contributes to our understanding of 'control societies,' where information controls become key instruments of governance and domination.

3:15 pm – 3:45 pm: **Coffee break**

3:45 pm – 4:45 pm: **Civil society roundtable discussion on information controls**

4:45 pm – 5 pm: **Closing remarks**

Registration: There is no registration fee for this event, but advance registration is required. Please register by clicking here

For any inquiries, please contact the organisers:

Dr Ashwin Mathew: ashwin.mathew@kcl.ac.uk

Gowhar Farooq: m.g.farooq@kcl.ac.uk