Abstracts: Boston Bergen Forum 2024

1) **Maria Brincker (UMass Boston)**

"Responsible agency and the nature of surveillance effects"

Hans Jonas proposed in his 1973 essay "Technology & Responsibility: Reflections on the new task of ethics" that we need a new ethics that takes responsibility beyond local face-to-face encounters to the larger effects of our actions as we expand our powers through various new technologies. Looking at Jonas argument in 2024 I suggest that we should also consider how we can salvage spaces of more locally understandable action that better suits the nature of volitional action. Using Aristotle's thoughts on volitional action I try to make this point in the context of the data flows of digital surveillance.

2) **Eugenia Stamboliev (Vienna)**

"Trustworthy AI as a politicised conflict"

'Trustworthy artificial intelligence' (TAI) is contested; philosophically and politically. As Big Tech grows in power, we face a struggle to reconcile ethical and economic demands in AI development. In 2019, the European Commission published the Ethics Guidelines for Trustworthy AI (EGTAI), aiming to strengthen the ethical authority and find common ground among the AI industry, ethicists, and legal regulators. By employing Laclau's work on empty signifiers and critical discourse analysis, I will present work (with Tim Christiaens) on how TAI unites diverse stakeholders, but also enables a hegemony of AI industry demands over ethics. This politicised struggle leaves AI ethics with the uncomfortable choice of affirming the industry's hegemonic position, undermining the purpose of ethics guidelines, or contesting industry hegemony.

3) **Gloria Origgi (CNRS Paris)**

"How algorithms deconstruct collective hermeneutical resources"

In her influential book, Epistemic Injustice, the philosopher Miranda Fricker defines "collective hermeneutical resources" as the shared tools of social interpretation that are accessible to most people in a given society. How are these resources changing with AI? In a technological world where algorithms generate new collective hermeneutical resources and limit the access to the old ones, it becomes difficult to the users to understand how their behavior is manipulated and to make sense of their experience of the informational environment. This has political consequences on the sense of a "shared reality and experience" that is needed to act as a responsible citizen. This talk will clarify the notion of collective hermeneutical resources and its application to algorithmically generated informational environments.

4) **Filipe Campello (Universidade Federal de Pernambuco):**

"Artificial Affects: Making AI a tool for Social Freedom"

In this talk, I propose the concept of artificial affects as a model for understanding the circulation of political emotions in digital networks. Using Brazil as a case study, I will demonstrate how the circulation of affective polarization within echo chambers has been crucial to the rise of the far right and the spread of hate speech. From a normative perspective, I propose a model of algorithmic pluralization and decolonization. This model can draw inspiration for instance from 'quilombola' and Amerindian technologies to challenge the neoliberal framework and move towards a model of social freedom.

5) **Anna Maria Lorusso (Bologna)**

"*Facts, fictions, gossips and other truth claims*"

How has the category of truth changed in our digital society, made up of social media and A.I.? Despite those who believe that truth is no longer a central category in the new communication landscape, I will try to explain why for me the category of truth is still crucial, in particular in the form of *truth claim*, between fictionality, self-legitimation and aspiration. In a epistemic context of multiplied "truth claims", what changes is both the space of factuality and the space of sociality. Within this media landscape in which claims to truth are multiplied, I would like to focus in particular on three aspects:
-the blurring distinction between realistic and fictional contents
-the perlocutionary force of contents, based on a logic similar to that of gossip
-the interpretative skills required of the contemporary reader.

6) **Carl Öhman (Uppsala)**

"Gods of Data: Language Models as the personified Authority of the Past"

What are Large Language Models? This talk proposes that they are best interpreted as *gods*. In a theological sense, gods are supernatural beings beyond time and space, which is obviously nothing like any form of AI. In an *anthropological* sense, however, gods are understood as the personified authority of a group over time—a collective of ancestors synthesized into a unified voice. And *this* is exactly what a language model is—an amalgamation of our digital ancestors. This reinterpretation casts new light on the nature of LLMs, but more importantly, it allows us to formulate new ways of critiquing their political implications.

7) Natali Helberger (Amsterdam)

"ChatGPT: what is the impact of LLMs on our democracy?"

With the role out of ChatGPT, Open AI and Co launched maybe one of the largest societal experiments ever: what happens if we make extremely powerful models to automate language available to essentially anybody on this world who has access to the internet and some form of computer? With the launch of ChatGPT and other Large Language Models the question of "What is responsible use of AI" becomes vital to the continued success of our democracy, because language is what makes our democracy function. In my keynote, I will examine the potential implications of ChatGPT & Co for our democracy, as well as our ability – as a democratic society – to govern Generative AI, as most recently attempted in the European Union's AI Act.

8) Mark Thomas Young (UiB)

"What's Missing from AI Ethics: Economics, Politics and Power"

Despite the rapid proliferation of literature on AI ethics in recent years, the field currently finds itself in a crisis. A growing recognition that the field of AI ethics is unequipped to grapple with the complex social and ethical challenges generated by the rapid spread of this technology has even led a number of recent authors to disassociate their ethical analyses of AI from the moniker of AI ethics altogether. In this paper, I will argue that the current dissatisfaction with AI ethics stems ultimately from a general failure within the field to reckon with the wider socio-economic context within which the technology is produced and sustained. In the first section I will explore how the common tendency to consider the ethical problems of AI to admit of technical solutions reflects an underlying conception of the nature of the technology as algorithms produced through the activity of software design. In contrast to such depictions, I will draw upon recent studies from the social sciences in order to shed light on the nature of AI as an industry which is sustained by a global network of material infrastructure and human labor. Here I will seek to show that by approaching AI as a capitalist industry, this literature encourages us to adopt an economic lens when examining the nature and impacts of the technology. In the second section I will illustrate different ways in which this economic lens holds significance for the ethics of AI. In contrast to current discourse which implies that AI can be whatever we want it to be – if only we can decide on what is the correct value or ethical rule to embed within the technology, I will attempt to show how the current economic context of the technology severely constrains the ways in which it can respond to ethical concerns. Furthermore, I will illustrate how ethical initiatives associated with the democratization of AI, such as open sourcing, can appear differently when viewed from an economic lens. In the final section, I will highlight how the neglect of economic context in AI Ethics connects to wider issues surrounding the proper role of philosophy in society while reflecting deeper challenges concerning the significance of context in moral philosophy itself.

9) Alec Stubbs (UMass Boston)

"AI Friendship: On the Uncontrollability of the Other"

This presentation explores the potential for friendships with artificial intelligence (AI) through Aristotelian ethics and Hartmut Rosa's philosophy of resonance. It argues that AI, lacking mutual recognition and moral virtue, cannot fulfill Aristotle's criteria for virtuous friendships. Furthermore, Rosa's concept of uncontrollable, resonant relationships underscores that AI, inherently predictable and controllable, fails to engage in the transformative dynamics essential to true friendship. Thus, AI relationships may enhance feelings of alienation rather than genuine connection, as they contradict the foundational aspects of interpersonal resonance.

10) João C. Magalhães (Groningen)

"The end of recognition theory?"

Originally conceived in the 1990s, Axel Honneth's work remains the most influential version of recognition theory. By updating Hegel's formulation with insights from the social sciences, Honneth has created a fully secular framework that aims to explain both the social constitution of individual identity and the ideal form of this process. However, can his theory properly account for how digitalization transforms the struggle for recognition? If not, how might we rethink mediated recognition—and the very concept of freedom? This presentation will tentatively explore these problems.

11) Anat Biletzki (Quinnipiac & MIT)

"Israel 2023: A Tale of Two Conflicts"

Two acknowledged political conflicts, one based on governance and the other identified as a military encounter, have beset the state of Israel—attractively termed "Startup Nation"—in 2023. The first, the "Constitutional Revolution," led to protests predicated on design, communication, propaganda, and mobilization highly dependent on digitization (reminiscent of the Arab Spring, more than a decade earlier). The second, typically named the "Israel-Hamas War," has provided opportunity for AI to serve the military goals of Israel (in pursuing what has been termed Humanitarian Crimes). These will be interrogated with a view to giving a "perspicuous representation" (Wittgenstein) of AI as an ethical-political policy; in other words, as failing to be a merely neutral tool in political conflict.

12) Samia Touileb (UiB)

"Unleashing digital companions: societal, political, and ethical implications of LLMs in a pre-post-human world"

Regardless of one's stance on transhumanism or high P(doom) values, it is evident that Large Language Models (LLMs) exert a profound influence on our world. While they can be viewed as positive tools with the potential to revolutionize our interactions with AI, it remains crucial to acknowledge their limitations and, more significantly, their societal impacts. In this presentation, I will delve into the societal, political, and ethical implications of LLMs. Additionally, I will explore ongoing discussions regarding the next developmental steps, including speculative futures in a post-human world.

13) Leif Hemming Pedersen (Roskilde):

"Recognition struggles in the (in)visibilization society"

This talk draws on Axel Honneth's (1995, 2014) influential recognition-theoretical framework to argue that mediated recognition struggles have undergone a process termed 'from accessibility to notability' in times of deep mediatization. This is exemplified with the use of ethnographic material from a case study with a young Danish football freestyler, who uses TikTok to counter negative stereotypes about (Muslim) women. The talk seeks to demonstrate that while easier access to public sphere(s) have in some ways strengthened the potential for more voices and transformational recognition struggles, remaining issues of accessibility have been further complicated by the manipulative and opaque circumstances of platforms that subjects these struggles to an imperative of struggling for commercialized notability.

14) Jacob Burley (UMass Boston & Harvard)

"Algorithmic Agency!?"

Given the explosion of progress in natural language processing (NLP) and artificial intelligence (AI) over the last few years, the attribution of agency to computer systems has become increasingly tempting and increasingly controversial. AIs like OpenAI's Chat-GPT, Anthropic's Claude, and Google's Gemini are capable of highly sophisticated cognitive tasks including natural language, audio and visual processing. In addition, such models have enabled a wave of artificial companions such as Replika and Nomi and mental health coaches such as Wysa and Woebot. Furthermore, AI is beginning to show capabilities in areas once considered the sole domain of human agents such emotional sensitivity, contextualization, novelty, and creativity. Such capabilities have caused some to consider whether agency is a unique property of human beings (and other organisms) or whether current AI systems should be considered legitimately agential. The aim of this paper is to argue narrowly that contemporary AIs cannot be considered epistemic agents -

specially, that their outputs lack justificatory status because they are not produced on the basis of reasons. The goal of this analysis is not to throw stones at transformative technologies but, rather, to ensure their responsible implementation. I raise two concerns in this direction. First, as AIs become more intelligent, reliable, and capable they are likely to be relied upon more and more for decision-making. As such, it is critical to clarify the epistemic status of AIs. Mistakenly attributing justificatory properties to AIs risks misattributing blame and praise to artificial agents instead of human agents. Second, mistakenly conflating the agential status of AIs with the agential status of human beings forces us either to overestimate the agential status of AI or downplay the agential status of human beings. The first option would seem to require the extension of moral, epistemic, and legal rights and duties to AI. The second risks diminishing the basis of the rights and duties in human agents. As we consider a future in which interactions with apparently agential artificial systems are likely to be a norm, we must carefully consider the consequences of attributing agency to such systems.