

# SHAPING AI

## Shifting AI Controversies

### Prompts, Provocations, and Problematisations for Society-Centered AI

Statement and Call for Contributions to the Final Conference of the International Project “Shaping AI”  
[shapingai.org](http://shapingai.org)

### Project Partners

Centre for Media, Communication and Information Research (ZeMKI), University of Bremen / Alexander von Humboldt Institute for Internet and Society (HIIG) (PI: Christian Katzenbach)

médialab, Sciences Po, Paris (PI: Donato Ricci)

Centre for Interdisciplinary Methods (CIM), University of Warwick (PIs: Noortje Marres, Michael Castelle)

NENIC Lab, INRS Montreal/Quebec (PI: Jonathan Roberge) / Algorithmic Media Observatory, Concordia University (PI: Fenwick McKelvey)

### Date and Venue of the Conference

Social Science Research Centre Berlin (WZB)  
Reichpietschufer 50, 10785 Berlin

29 & 30 January 2024

### Statement

**Controversies about AI abound**, especially since ChatGPT took over the Internet by storm, becoming the most popular application in the Web’s history within only a few months. The current excitement about the perils and prospects of general-purpose AI applications like ChatGPT is only the most recent wave of public interest in the long history of “artificial intelligence” (AI). With its metaphysical imaginaries of human-machine symbiosis, anthropomorphic robots and machine thinking, arguably oversized scientific claims and technological developments in this field have always raised concerns. What the current debate makes much more visible than previous attention cycles, though, is that contemporary AI companies and scientists dominate not only the discourse promoting AI’s prospects but also that on AI’s perils. From engineers at OpenAI to research pioneer Geoffrey Hinton, technologists and industry-based scientists increasingly articulate warnings that AI might cause serious and fundamental damage to societies. With this move, the already dominant players are now also occupying the space of public critique, yielding

the risk that activism, social science, critical journalism and the arts are pushed even further to the margins of public and expert debates. Are we currently having the public controversies on AI that we should have, or is AI panic derailing us from actual and relevant concerns? How do we get to the controversies that we need and to the exploration and articulation of society-centered AI?

**The ChatGPT Moment.** The introduction of OpenAI's ChatGPT application in December 2022 brought to the public fore a set of controversies that had been brewing within machine learning research communities for years about the potential limitations and risks of large language models (LLMs). By fine-tuning and thus recasting these models into a dialogical mode, AI products again captured public attention in diverse fields of business, education, and communication. But to what extent can their outputs be trusted? Hundreds of thousands of non-specialists suddenly became immersed in what had previously been open research questions regarding the subtleties of how these multibillion-parameter models' outputs depend strongly, if unpredictably, on their input prompts as well as the nature of the underlying training data, and what their environmental, legal, and social impacts might be. At the same time, users' continuous prompting of ChatGPT constitutes a proverbial "firehose" of interactional data that might inform the development and marketing of future models, helping OpenAI maintain market dominance. How does this vastly increased public participation in the processes of AI research affect the controversiality of AI, and how?

**The prompt as provocation and problematisation.** In these debates on the proper use (or not) of dialogical LLMs, the prompt has risen as a core theme. As much as we prompt AI when using interactive user interfaces, AI has prompted *us* to reconfigure societies around its promised value, one that remains vague and contested nonetheless essential with nations at risk of falling behind. The BBC's recent headline "UK needs culture shift to become AI superpower" sharply exemplifies how the discourse of AI is shaping society - far beyond the implications of the actual technologies. Much like prompt engineering itself where the task is transformed to meet an effective prompt in the interface, we are witnessing a prompting of society. How can we repurpose "the prompt" as a provocation and problematisation for addressing the actual concerns about AI?

**Shaping AI is an international research collaboration that has examined the trajectories and triggers of controversies on artificial intelligence in the past 10 years.** Across investigations in media, research, policy and engagement activities we have learned that the controversiality of AI is not something that is always already there and that can be taken for granted. Of course, applications such as facial recognition systems or autonomous weapon systems are and should be subject of public controversies. Yet the controversial character of AI technologies and applications is also mobilised and can be strategically employed by different actors and stakeholders.

**Media debate on AI is strongly dominated by technologists, scientists and entrepreneurs.** "Who is the best person to talk about AI other than the one who is actually making it?" A journalistic routine to foreground products and promises has translated into a concentration of few voices, often computer scientists who have become both entrepreneurs promoting the technologies and evangelists of certain theories of human cognition. In comparison, social scientists and critical voices struggle to make their way into public discourse on AI. For example, while expert critics had warned for years about large language models' issues with discrimination and misleading statements, the media only started to prominently cover such stories when tech leaders began to sign open letters themselves calling for temporal bans and articulating threats [to democracy](#) and [world order](#). These framings risk closing [the debate on actual harms](#) generated by current generations of AI and to allow for the ever same influential players to perform their significance and maintain AI's hype.

**Researchers are amongst the most concerned by, and invested in AI hype.** Yet, does this practice of hyping represent a unanimous discourse or a particular set of actors or spokespersons? In this context, it makes a difference how controversiality and reflexivity

play out *within* the field of AI, namely to locate the internal ambivalences and criticisms, which may or may not struggle to find their way into the public sphere. Are deeper concerns currently emerging that would be epistemological as well as socio-political in nature, ranging for instance from the question of scaling and data quality to AI deployment in (dis)information systems? What in turn could prompt the “AI social science deficit” to be reduced and could research in and of itself become a society-centric issue?

**Public policies privilege national economic prosperity over public inquiry into societal risks and harms.** The great excitement around AI technologies has elicited keen interest from policymakers in shaping and integrating these systems into their respective national and multinational policy contexts. While the European Union and other international organisations, including the United Nations, actively engage in policy dialogues and policy-making, the vast majority of endeavours remain centred on advancing the prosperity of domestic industries and research clusters. Given the pervasiveness of AI technologies across national boundaries and legal jurisdictions, however, the focus on the nation-state fails to acknowledge the global impact experienced by individuals in the face of unregulated and partially concealed deployments of AI systems. This is reinforced by the fact that the majority of government policies distinctly focus on the economic advantages and developments of AI systems, creating brands such as “AI made in Germany”. How do we make sure that AI policy ensures human rights for individuals, and that regulations prioritise society-centered development with regard to AI?

**These developments demand new ways of engagement: of defining AI problems while remaining grounded in the different practices that these technologies call into question.** Various attempts at novel interactions with computational technologies have sprouted: from reverse engineering to critical making; interfaces have been crafted to enable comprehension and transparency; from interactive computing notebooks to prompts are but a few examples of the myriad endeavors undertaken to build new interfaces between AI technologies and their *users*. Yet, these approaches run the risk of constraining themselves to addressing problems defined within technology-centered thinking. What we need is establishing ways to interact and intervene in defining, structuring, and transforming the problems these technologies pose to the *public*.

**We need more diverse prompts, provocations, and problematisations of AI if publics are to acquire agency in the process of defining the benefits and harms of AI for society, i.e. to realize Society-Centered AI.** The current moment features peak attention for issues of AI and society, and final negotiations on key AI regulations are underway. Even the most high-profile proponents of AI consider these technologies as disruptive, problematic, potentially harmful or indeed dangerous. But this rather recent cultivation of AI panic by technologists might rather derail the discussion away from infrastructural inequalities and other actual concerns about AI. At the same time, this development constitutes a moment of opportunity to counter technology-centered discussions with provocations and problematisations for society-centered AI. What kind of prompts do we need now? How do we elicit the controversies that we need?

## **Interventions and Contributions are invited from Scholars, Civil Society and Practitioners**

Against this background, we welcome contributions that address these issues from scholars of diverse disciplines as well as interventions from civil society, practitioners and developers. Your submissions should engage with the questions and provocations posed in the text above.

**Extended abstracts** of approximately 4,000 to 6,000 characters in length (excl. references) should be submitted **no later than 30 October 2023** via our submission form. Speakers will be notified by 16 November 2023.