

TECHNOLOGY

The Dark Side of Generative AI: Automating Inequality by Design

by Josh Entsminger, Mark Esposito, Terence Tse, and Aurelie Jean



Image Credit | Tim Mossholder

Some applications of generative AI come with the unintended consequence of a rise in inequality.

✔ **INSIGHT** | OPINION 26 Jun 2023

OpenAI's launch of ChatGPT has accelerated the race towards developing general-purpose artificial intelligence systems (GPAIs) (OpenAI, n.d.). This trend started with Google's acquisition of DeepMind in 2014 (Lohr, 2015) and has since been fueled by investments in large language models (LLMs) by companies such as Microsoft (OpenAI, 2019) and Google (Vincent, 2021). This has led to a rebranding of LLMs as generative AI, which has the potential to provide novel outputs with prompts. However, this collective societal input powering generative AI has also led to a radical centralization of collective intelligence, often under a single interface and single company.

The investments and acceleration of LLMs have driven a global push among well-funded actors, splitting the AI industry into two: the narrow AI industry, focused on small and individual solutions, and the general AI industry. The winner of the race for GPAI will be able to internalize and compete with all other niche AI firms or firms. PwC's recent billion-dollar investment in generative AI has entrenched this push through Generative AI (Wang, 2021), and whoever is first to reach general-purpose AI will be the ultimate winner.

However, there are concerns about the governance of GPAIs. Italy has raised privacy-based concerns (Zorz, 2021), and the EU is discussing whether GPAI falls under the remit of current AI legislation (Lomas, 2021). In the US, there is a debate on how to take advantage of the unique national position as the leading provider of GPAI solutions (Zimmer, 2021). There is a risk that if the US stops regulating, they will legitimize the position of upstream actors, such as OpenAI and Anthropic, while enabling downstream harms.

There are also concerns about the potential harms of generative AI, including biases within generative AI outputs (Geburu, 2018) and the potential to reduce the cost for downstream malicious actors, such as scammers and propaganda (Buolamwini, 2018). These concerns have provoked a global debate around the harms of these systems, calling for a slow-down of AI innovation and citing existential risk (Hinton, 2021).

While some argue that AI systems have real, full harms now and that political attention needs to focus on extant harms, others argue that the implicit vision and narrative underlying the acceptance of generative AI enables the kind of value and society that is, in

fact, desirable (Altman, 2018). All efforts and firms in the pursuit of AI are a function of how global physical and legal infrastructure for data management and ownership are designed.

Therefore, the question is not simply whether to slow down or speed up the development of AI, but what direction AI development is going in. A collective validation of the current era of generative AI, under the guise of pursuing GPAI, risks entrenching a new political economy of AI, one built on the increasing centralization of collective intelligence and inputs, the centralization of decisions on what kinds of downstream harms are permitted, and the privatization of decisions on existential risks. This could lead to a new generation of inequality by design.

There is a need for a deliberative democratic movement on how to govern these systems, placing these actors at the front of the table. It is not enough to have a democratic exchange about risk and governance without addressing how to reorient the political economy of AI. This political economy is one of global “data extractivism” powering surveillance capitalism, a version that will be accelerated, not disrupted, by new advancements and diffusions of LLMs for conversational services (Zuboff, 2019).

As generative AI continues to gain popularity and momentum, concerns over its potential dark side are becoming increasingly prevalent. While there is no denying the incredible potential that these systems hold, it is essential to recognize and address the risks they pose.

One of the most significant risks associated with generative AI is the centralization of collective intelligence. With companies like OpenAI and Anthropic driving the development of GPAIs, there is a growing concern that these entities could become the gatekeepers of all AI development, effectively monopolizing the industry. This would not only limit the diversity of perspectives and ideas but also restrict the availability and accessibility of these technologies.

Another significant concern surrounding generative AI is the potential for bias in outputs. As these systems rely heavily on pre-existing data, they are vulnerable to perpetuating existing biases and discrimination. This could lead to significant social, economic, and

political consequences, further entrenching existing inequalities and injustices.

There is also the risk that generative AI could make it easier and more affordable for malicious actors to spread disinformation and propaganda. By automating the creation of convincing and persuasive content, these actors could manipulate public opinion, sow discord, and cause significant harm.

Despite these risks, there is a growing push towards the development of GPAIs, with significant investments being made in these technologies by companies and governments worldwide. However, there is a need to ensure that these developments are guided by ethical principles and values, and that they serve the best interests of society as a whole.

To achieve this, a deliberative democratic movement is needed, one that places the actors at the front of the table, and that addresses the underlying political economy of AI. This movement must consider the ethical implications of these technologies, address issues of bias and discrimination, and ensure that they are developed and deployed in a manner that promotes social justice, equality, and human rights.

Ultimately, the question is not whether to slow down or speed up the development of AI but what direction that development is taking. By recognizing and addressing the dark side of generative AI, we can ensure that these technologies are used for the greater good, and that they contribute to a more just, equitable, and sustainable future.



Josh Entsminger [Follow](#)

Josh Entsminger is a PhD candidate in Innovation and Public Policy at the University College London Institute for Innovation and Public Purpose (IIPP).



Mark Esposito [Follow](#)

Mark Esposito is Professor at Hult Int'l Business School and Harvard University's Division of Continuing Education and works in public policy at the Mohammed Bin Rashid School of Government. He directs the Hult Futures Impact Lab. He co-founded Nexus FrontierTech and the Circular Economy Alliance. He has written over 150 articles and edited/authored 13 books. His next book, "The Great Remobilization" will be published by MIT University Press in the course of 2023.



Terence Tse [Follow](#)

Terence Tse is Professor of Finance at Hult International Business School and a co-founder and executive director of Nexus FrontierTech, an AI company. He is also a co-founder of Excellere, a think tank with the goal to help people explore and release their potential through new technologies. He has worked with more than thirty corporate clients and intergovernmental organisations in advisory and training capacities. He has written over 110 articles and three books including The AI Republic: Building the Nexus Between Humans and Intelligent Automation (2019). His next book, The Great Remobilization, will be published by MIT University Press, in the course of 2023.



Aurelie Jean

Aurelie Jean is a computational scientist, entrepreneur, and author, who shares her time between the United States and France. She has founded a development and consulting agency on computer & data modeling and has co-founded an AI deep tech startup on early detection of breast cancer.