

# Anticipating Futures: How Artificial Intelligence acts as an Amplifier of Inequity

Somya Joshi, Björn Nykvist

Senior Research Fellows, Stockholm Environment Institute, Sweden (somya.joshi@sei.org)

## Abstract

The accelerating advancement of Artificial Intelligence (AI) brings disruptions, not simply more powerful or efficient than current information technology, but a qualitatively different, and non-linear progression. Policy institutions simultaneously grapple with the governance and decision-making challenges in realizing the SDGs as well as with the sheer pace of information flows and technology shifts. In this policy brief we discuss the inequity in distribution of risks and benefits when it comes to AI and implications this might have for the implementation of the SDGs. Left unattended, we argue, AI advancements risk amplifying inequalities both between and within nations (SDG 10) and thus demands special attention from governance and policy circles globally.

## Introduction

Although the evolution of Artificial Intelligence (AI) has followed familiar socio-technical patterns of hype bursts followed by long winters, the wide release in the recent past of tools based on generative AI, such as the large language model ChatGPT, arguably represent a turning point. With lightning speed this has brought disruption and discontinuity outside of niche academic communities and technology companies, into wider society, with a promise to upend our assumptions regarding human development.

What sets AI apart from broader technological evolution is that we might be experiencing, or be close to reaching, a tipping point, where we no longer explicitly tell very capable systems what to do. Instead, the AI system is given a goal, a vast set of examples, and it figures out the patterns or rules itself, with emergent and surprising capabilities & capacity as a result. This is done via training sets of human generated data, that mirror our real-world prejudices, prioritization and power hierarchies, and thus risk amplifying current inequities in both representation and outcomes. But it is also the case that AI progress is currently unfathomably fast, the complexity and inner workings of these models are incomprehensible or inaccessible, and according to AI researchers themselves, the outcomes are unpredictable, whilst still weakly governed (Klein 2023). The ways in which these technologies are integrating allow for new possibilities hitherto unimagined. This clearly has impacts on human development and sustainability, enshrined in the Sustainable Development Goals (SDGs).

## Global Goals and AI

SDG 10 explicitly articulates the need for reducing inequalities and ensuring no one is left behind as an integral part of achieving the SDGs. It highlights inequality within and among countries as a persistent cause for concern, exacerbated by the recent pandemics and geopolitical rifts. What role will AI play in widening or reducing these gaps, then becomes a mission critical question and one central to this brief.

AI holds the potential of addressing climate change (Covels et al. 2021; Leal Filho et al. 2022) and biodiversity loss (Eastwood et al. 2022), whilst contributing to more effective monitoring and applications of natural resources and emissions (Emami Javanmard and Ghaderi 2022). In doing so, accelerating progress towards the achievement of the SDGs (Goralski and Tan 2020). Recent scholarly work has gone as far as to term this technological progression a “vector of hope” (Del Río Castro, González Fernández, and Uruburu Colsa 2021). Countering this optimism somewhat, Vinuesa et al. 2020; Galaz et al. (2021) highlight how all of the 17 SDGs are believed to be moderately or strongly affected by AI and automation, but caution that 59 or more of the sustainable development targets might be inhibited by AI. If the SDGs are to be employed as a framework for AI ethics, what is needed is that we consider closely the context and the broader sociotechnical system (Saetra et al 2021).

## Inequity Dimensions

The field of AI ethics so far has been predominated by studies raising awareness about how livelihoods will be impacted by automation, the qualitative shift of jobs available based on geography and income, as well as ethical concerns related to bias and prejudice built into training data and the models that follow.

Regarding economic opportunity, these concerns stem from the observation that AI continues a long-standing trend of increasing automation. AI will impact workers differently. New tools like Chat-GPT are also generally expected to drive unevenly distributed global economic growth (Bughin et al. 2018). Multinationals or large tech firms based in the Global North, with access to data and the resources to develop AI, might receive the majority of these benefits, driving inequalities further.

The ethics and privacy concerns about data collection, the embedding of bias and prejudices into AI models when using historical data to train said models, and the lack of transparency of decision rules generated is also well recognised. Automation and inequity in particular are drivers to flag in the discussion of AI impacts more broadly in society (Crawford, 2021; Eubanks, 2017; Zuboff, 2018, Zajko, 2022).

Adding nuance to the discussion on inequity, is the body of scholarly work underpinning the materiality of AI and its supply chains, as well as the life cycle analysis and footprints of such large complex systems. van Wynsberghe et al (2021) call for a “movement to foster change in the entire lifecycle of AI products (from idea generation, training, re-tuning, implementation, governance) towards greater ecological integrity and social justice”. These aspects of justice appear hard to attain when we look at how developments in AI are critically linked to non-universal access to increasingly large data sets and the computing infrastructure required to process them. As a tiny fraction of nations and companies control the development and application of AI, this raises important questions regarding the geographical disjuncture between where AI development takes place and where rewards of GDP are captured.

## Governing AI

To make decisions or design policies based on the best available information requires effective tool support. AI is an enabling technology that could allow decision makers to significantly improve their capabilities to make decisions based on larger and more complex information. However, whilst AI in this way can drive better decision making (Di Vaio et al. 2020), it can also be used for misinformation, deep fakes, and algorithmic bias increasing risks of systematic deviation from equality that emerges in the outputs of an algorithm (Kordzadeh and Ghasemaghahi 2022).

This calls for a pressing need for policy making institutions to take responsibility and act in

consideration of an ethical framework such that AI systems can be trusted (Theodorou and Dignum 2020). Policy institutions are just beginning to grapple with the speed of this shift, with many calling for urgent governance, regulation and ethics frameworks to be built into the nascent stages of its development. But there exists relatively sparse discussion on “how” this can be achieved. One refreshing signal against the noise of AI hype, is that of Bender and Gebru (2021), who describe how large language models absorb the hegemonic worldview from their training data. They call for a shift in mindset to one of careful planning, before starting to build either datasets or systems trained on datasets. As an antidote to the hysteria and hype of AI, they call for research to be mindful of people who stand to be adversely affected by this technology. This can only be materialised by realigning the research process for considering environmental impacts, for doing careful data curation and documentation, for engaging with stakeholders early in the design process, for exploring multiple possible paths towards long-term goals.

## Policy recommendations / conclusions

In this policy brief we have argued that the equity implications of AI go beyond the politics of data collection, learning and use. AI risks acting as an amplification tool, that is fraught with design considerations and inherent unknowns. This calls for the need to recognize growing evidence of algorithmic bias, the risks of deep fakes and misinformation, wherein AI acts as a driver of further political polarization as well as exacerbation of inequalities.

To counteract these trends, we recommend the following steps as design guidelines bridging the world of policy making and that of emerging technology such as AI:

1. Urgent need for Global Governance: Stepping up efforts on AI governance, that transcend national boundaries and engage both public and private actors at the forefront of AI development. OpenAI, the creators of ChatGPT, for example recently called for global standards in training AI systems (Altman 2023). This is needed to address the alignment problem of ensuring AI works in the best interest of humans, but in the context of SDG 10, we also urgently need to elevate the discussion on risks of amplified distributional inequity as raised earlier on the risks and benefits of AI.
2. Holding up a new mirror: rather than replicating the structural and societal inequities embedded

in our historic training data, as well as in opaque or proprietary processes of how weights, rewards, and priorities are applied in reinforcement machine learning, we need to bake into the design of AI, ethical frameworks that break out of such patterns of bias and equity gaps. Adopting regulations general enough to cover emerging risk with rapidly growing capability of AI systems is critical (Tegmark 2022).

3. “Sustainability Budgets” in the development of AI: As Raper, Coeckelbergh et al (2022) have suggested, there is a need to design budgets similar to the ones we currently have for ensuring privacy of data, to avoid sustainability deficits in the way AI is unleashed upon wider society. This recommendation in the context of AI and global agenda setting, would enable environments that prioritise human and ecosystem development over short term profits or cyclical industry & political agendas.

What the three recommendations all have in common is an architecture that favours principles for sustainability, global governance and human development, rather than short term visions of progress or growth for growth’s sake. However, they beg the question on who would comprise such committees or global task forces to ensure the needs of the most vulnerable segments of society are balanced with that of innovation and technological progress. Here we recommend a global coalition not of first-movers but rather including those that are the most vulnerable and have the highest stakes in the outcomes of this megatrend. Efforts to bridge AI and policy particularly within the context of “leaving no one behind” as enshrined in the SDGs, need to consider downstream effects in order to block foreseeable harm to society and different social groups. We thus call on collaborative design approaches with representatives of affected communities, as well as policy experts and scientific scholars to strengthen global governance of AI.

## References

- Altman, Sam. 2023. “Planning for AGI and Beyond.” OpenAI. <https://openai.com/blog/planning-for-agi-and-beyond>.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (pp. 610-623).
- Bughin, Jacques, Jeongmin Seong, James Manyika, Michael Chui, and Raoul Joshi. 2018. “Notes from the AI Frontier, Nodeling the Impact of AI on the World Economy. Discussion Paper September 2018.” McKinsey Global Institute.
- Cowls, Josh, Andreas Tsamados, Mariarosaria Taddeo, and Luciano Floridi. 2021. “The AI Gambit: Leveraging Artificial Intelligence to Combat Climate Change—Opportunities, Challenges, and Recommendations.” *AI & SOCIETY*, October. <https://doi.org/10.1007/s00146-021-01294-x>.
- Del Río Castro, Gema, María Camino González Fernández, and Ángel Uruburu Colsa. 2021. “Unleashing the Convergence amid Digitalization and Sustainability towards Pursuing the Sustainable Development Goals (SDGs): A Holistic Review.” *Journal of Cleaner Production* 280 (January): 122204. <https://doi.org/10.1016/j.jclepro.2020.122204>.
- Eastwood, Niamh, William A. Stubbings, Mohamed A. Abou-Elwafa Abdallah, Isabelle Durance, Jouni Paavola, Martin Dallimer, Jelena H. Pantel, et al. 2022. “The Time Machine Framework: Monitoring and Prediction of Biodiversity Loss.” *Trends in Ecology & Evolution* 37 (2): 138–46. <https://doi.org/10.1016/j.tree.2021.09.008>.
- Emami Javanmard, Majid, and S. F. Ghaderi. 2022. “A Hybrid Model with Applying Machine Learning Algorithms and Optimization Model to Forecast Greenhouse Gas Emissions with Energy Market Data.” *Sustainable Cities and Society* 82 (July): 103886. <https://doi.org/10.1016/j.scs.2022.103886>.
- Galaz, Victor, Miguel A. Centeno, Peter W. Callahan, Amar Causevic, Thayer Patterson, Irina Brass, Seth Baum, et al. 2021. “Artificial Intelligence, Systemic Risks, and Sustainability.” *Technology in Society* 67 (November): 101741. <https://doi.org/10.1016/j.techsoc.2021.101741>.
- Goralski, Margaret A., and Tay Keong Tan. 2020. “Artificial Intelligence and Sustainable Development.” *The International Journal of Management Education* 18 (1): 100330. <https://doi.org/10.1016/j.ijme.2019.100330>.
- Klein, Ezra. 2023. “This Changes Everything.” *The New York Times*, March 12, 2023. <https://www.nytimes.com/2023/03/12/opinion/c-hatbots-artificial-intelligence-future-weirdness.html>.
- Leal Filho, Walter, Tony Wall, Serafino Afonso Rui Mucova, Gustavo J. Nagy, Abdul-Lateef Balogun, Johannes M. Luetz, Artie W. Ng, et al. 2022. “Deploying Artificial Intelligence for Climate Change Adaptation.” *Technological Forecasting and Social Change* 180 (July): 121662. <https://doi.org/10.1016/j.techfore.2022.121662>.
- Tegmark, Max. 2022. “FLI Position on the Proposal for a Regulation Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).” Future of Life Institute. <https://futureoflife.org/wp->

content/uploads/2021/08/FLI-Position-Paper-on-the-EU-AI-Act.pdf?x76795.

- Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark, and Francesco Fuso Nerini. 2020. "The Role of Artificial Intelligence in Achieving the Sustainable Development Goals." *Nature Communications* 11 (1): 233. <https://doi.org/10.1038/s41467-019-14108-y>.
- Wynsberghe, Aimee van. 2021. "Sustainable AI: AI for Sustainability and the Sustainability of AI." *AI and Ethics* 1 (3): 213–18. <https://doi.org/10.1007/s43681-021-00043-6>.
- Theodorou, A., & Dignum, V. (2020). Towards ethical and socio-legal governance in AI. *Nature Machine Intelligence*, 2(1), 10-12.
- Nerini, F.F.; Sovacool, B.; Hughes, N.; Cozzi, L.; Cosgrave, E.; Howells, M.; Tavoni, M.; Tomei, J.; Zerriffi, H.; Milligan, B. Connecting climate action with other Sustainable Development Goals. *Nat. Sustain.* 2019, 2, 674–680
- Raper R, Boeddinghaus J, Coeckelbergh M, Gross W, Campigotto P, Lincoln CN. Sustainability Budgets: A Practical Management and Governance Method for Achieving Goal 13 of the Sustainable Development Goals for AI Development. *Sustainability.* 2022; 14(7):4019. <https://doi.org/10.3390/su14074019>
- Sarewitz, D. and Nelson R., 2008, Three rules for technology fixes, *Nature*, 456