

The Hidden Costs of AI

Decolonization from Practice back to Theory

Oumaima Hajri

This paper is dedicated to the ethical and societal aspects of artificial intelligence (AI) and focuses on two main topics: the hidden costs of AI and the importance of taking recourse to theory. Investigating the hidden costs of AI is crucial because, while there is a positive narrative surrounding its potential benefits, we must also consider its impact. Who benefits from it? And who is further marginalized? Additionally, it is essential to understand the underlying logic(s) and theories before jumping to technological innovations, in order to prevent the reduction of complex societal problems to mere technological solutions.

AI is a multidisciplinary field informed by various knowledge systems and disciplines. While technical disciplines have played a significant role in the development of AI, other areas such as the natural sciences (biology, physics), literature, linguistics, philosophy, and psychology have also made substantial contributions. The works of Jacques Loeb, a German-American biologist, and Nicolas Rachevsky, a Ukrainian-American physicist, exemplify the multidisciplinary influences on AI (Watson 1913; Piccinini/Bahar 2013). Loeb's concept of treating organisms as chemical machines highlights the biological perspective, thus emphasizing the reduction of living organisms to mechanistic aspects. Rachevsky's idea of using theoretical mathematics to simplify complex biological systems demonstrates the integration of mathematical and computational thinking into studying life.

Given the diverse streams of knowledge that have contributed to AI, it can be regarded as a multifaceted field that cannot be easily classified in one single category. It is not strictly a technical field, although technical disciplines are central to its development. AI is also not solely a vision for a utopian society or a social construct, even though it has implications for both. AI can therefore instead be understood as an umbrella concept encompassing a wide range of approaches and perspectives. To comprehend the various facets of AI, it is helpful to break it down into different categories. Based on the current scholarly literature, one possible taxonomy includes approaching AI as an instrument, AI as an infrastructure, and AI as an ideology. These categories highlight different aspects and roles of AI, such as its use as a tool

for knowledge amplification, its role in shaping technological infrastructure, and its broader philosophical implications.

AI as an Instrument

AI technologies have emerged as vital sources of knowledge, enabling us to perceive intricate features, patterns, and correlations beyond human capacities. We find ourselves captivated by the allure of vast datasets, believing that the larger the dataset, the better the outcome. AI is, however, more than merely a tool for observing and representing; it has begun shaping our perception of reality and influencing how we deal with data subjects. We are entranced by the notion that algorithms can self-learn and self-evaluate, thus bestowing an almost unquestionable authority upon them. But we must nonetheless ask ourselves why we readily accept algorithmic decisions without critically assessing their context. For example, why do we trust an algorithm that labels specific individuals as fraudulent or implies that people of a particular skin colour are more prone to criminality without considering these decisions?

Moreover, our society has become 'datafied', leading to an obsession with numbers and quantification (Van Es/Schäfer 2017). We have transitioned from understanding the immaterial aspects of the natural world to relying on calculations for every facet of our existence. In this mechanized worldview, human beings are reduced to mere datapoints. The social context, which is far too complex to be encapsulated in a mathematical formula, is thus cast aside. We have embraced the belief that numbers do not lie and that they alone can reveal the complete truth. Yet, we must recognize that technology is far from neutral. AI-based decisions are prone to inaccuracies, discriminatory outcomes, and embedded biases. These issues stem from the assumptions embedded within the design of algorithms, thus shaping AI systems to embody specific forms of power and authority driven by the values of those who create them.

Hence, every choice in the development and deployment of AI must be well-considered, with ethical considerations at the forefront. Within the realm of AI, biases manifest themselves in various forms. Historical biases are perpetuated as AI systems reinforce existing social inequalities, often limiting the roles available to women or perpetuating stereotypes (Benjamin 2019). Dataset biases arise due to the classification frameworks ingrained within our data. Outdated and distorted taxonomies misrepresent social diversities and exacerbate social hierarchies, warping our worldview (Noble 2018). Additionally, algorithmic biases and statistical biases emerge due to algorithms' efficient but compressive nature, leading to information resolution issues and a loss of social context (O'Neil 2017).

In summary, AI's role as an instrument for knowledge magnification presents both opportunities as well as challenges. While AI enables us to uncover valuable insights, we must approach its use cautiously. Technology is not neutral, and AI systems are susceptible to biases and inaccuracies. It is therefore imperative for us to critically evaluate algorithmic decisions, consider the broader social context, and ensure that AI aligns with ethical principles. Let us not forget that behind the algorithms, numbers, and datapoints lie the intricacies of humanity, which cannot be reduced to a mathematical formula.

AI as an Infrastructure

While we tend to perceive AI as intangible and non-physical, recognizing its profound impact on our physical world is crucial. AI is not invisible; it is highly physical and tangible, and its presence reverberates through our society. When we delve into the physicality of AI, we realize the extensive resources it requires. The construction of AI involves data centres, chips, computers, and various physical components, and thus involves a series of invisible factories (Crawford 2021). One aspect of AI's physicality that warrants attention is its carbon footprint and its implications for climate change (Okafor-Yarwood/Adewumi 2020). Amid discussions of environmental sustainability, we often overlook the role of our technologies and the escalating digitization accompanying them. Consider data centres' energy and water consumption, the computing power needed to sustain AI systems, and the consequential environmental impacts.

A study at the University of Massachusetts, Amherst, revealed a startling truth: a lifecycle assessment for training several standard large AI models can emit more than 626,000 pounds of carbon dioxide (Strubell/Ganesh/McCallum 2019). This staggering figure is equivalent to five times the lifetime emissions of an average car in the United States. It also highlights the substantial environmental costs associated with AI. What AI requires are thus not solely physical resources, but also planetary resources. This brings us to a critical point of reflection: the extraction of materials required for AI infrastructure. For instance, lithium, often referred to as 'grey gold', is predominantly found in Bolivia and is essential for batteries and chips. Similarly, cobalt, mainly sourced from the Democratic Republic of Congo, is vital in producing AI components (Crawford 2021).

These countries in the Global South are rich in raw materials, making them targets for resource extraction. We must, however, recognize the social implications of these technological advancements. The inequities that emerge are reminiscent of contemporary colonialism (Ensmenger 2018). For example, the chips within our smartphones, indispensable components of AI, are likely mined by children in Congo under deplorable conditions. We exploit their resources, subject them to

harsh working conditions, benefit from the technology, and, to add insult to injury, dump our electronic waste in these countries as well (Young 2020). What was once a treasure for us has now become trash for them. Consequently, the divide between the Global North and the Global South continues to widen.

This form of technological exploitation exacerbates the socioeconomic disparities between nations. Acknowledging AI as an infrastructure, as a physical entity with profound consequences, is thus vital. The carbon footprint it leaves behind and the planetary resources it exploits are pressing concerns to be addressed. We must confront the realities of contemporary colonialism and address the unjust distribution of benefits and burdens associated with AI. As we forge into the future, we must ensure that technological progress is accompanied by ethical and equitable practices, fostering a world where everybody shares the benefits of AI.

AI as an Ideology

Examining the ideologies that drive innovation, knowledge production, and beliefs within this field is essential to understanding the path towards decolonization. As mentioned above, AI has become a ubiquitous term eagerly embraced by all sectors. Attaching the label 'AI' to any technology brings a certain prestige and allure. Everything seems to revolve around AI, and everyone claims to be an AI expert. A so-called 'air of magic' is associated with AI for marketing purposes. Unfortunately, the desire to jump on the AI bandwagon has turned it into an end in itself, as evident in the race for dominance and hegemony between China and the United States (Stilgoe 2020).

This ties into the ideology of 'data as the new gold'. But, as Gina Neff points out, we treat gold as a commodity that can be found and extracted relatively effortlessly (Neff 2013). Data, on the other hand, is not simply mined; it is constructed and shaped by human actions. Another crucial aspect is the prevailing ideology of tech solutionism, which insists that innovation inherently relies on digitization (Lanier/Weyl 2020). As a result, traditional approaches to problem-solving should be considered before technology. Furthermore, the knowledge informing this field is heavily concentrated in Silicon Valley, primarily among data scientists and AI engineers. So, we must ask ourselves: Who shapes the understanding within this field? (Stilgoe 2020).

Why do we trust technical experts to make far-reaching decisions that profoundly impact our societies? Why do we grant such extensive power to privately owned companies? It is time to reclaim power and give it back to the people. The dominant engineering perspective within this sphere raises questions about whether we are using technology because we believe it to be a solution or if we are creating problems to implement our predetermined solutions. To address the

decolonization of AI concretely, two crucial factors come to the fore: the language surrounding AI and the imaginaries we associate with it.

The language used to represent AI often masks the accountability of the human beings who develop the technology. AI is portrayed as ‘intelligent’ and ‘self-learning’. This raises important questions about human intelligence and creativity. Moreover, the ‘artificial general intelligence’ concept prompts us to ponder the potential implications for humanity. Additionally, while discussions about robot rights are ongoing, it is essential to acknowledge that some human beings lack fundamental human rights (De Graaf/Hindriks/Hindriks 2021). The language we use to discuss AI erodes our understanding of what it means to be human. From a historical perspective, language has been a tool used in processes of colonization to convey certain beliefs and values. Deconstructing the concept of AI allows us to critically examine what it truly means to be intelligent. Looking back, we can observe that intelligence was historically linked to theories of racial improvement and eugenics, thus legitimizing racial hierarchies and justifying the enslavement of certain groups (Cave 2020).

If our association with intelligence is rooted in these historical contexts, it is unsurprising that these ‘intelligent’ systems we are striving to build inadvertently perpetuate harm towards particular groups. Why is it that these systems consistently harm specific communities? Why is it that those in the Global South who risk their lives extracting raw materials for our technologies do not benefit from them? Our word choices and language are generally never neutral, but instead inherently value-laden and political. In conclusion, the decolonization of AI requires a critical examination of the ideologies driving its development. We must challenge the notion that AI is a magical solution, the belief that data is readily available for extraction, as well as the concentration of knowledge in specific regions and disciplines. Let us reshape the language surrounding AI, redefine intelligence, and work towards equitable and just technologies that uplift all of humanity.

As we delve further into the discussion on decolonizing AI, we must address another crucial aspect: the whiteness of AI. It is vital to examine the representation of AI in imagery and language and to understand the need for a cultural shift within the field. First, let us consider the prevalence of predominantly white stock images and the portrayal of white robots. This raises questions about the realistic depiction of technology and the people behind it. We must aim for images that accurately represent the technology, its strengths, weaknesses, context, and applications (Cave/Dihal 2020). Similarly, speech recognition systems should be designed to recognize and accommodate various accents and dialects. The notion of what constitutes an accent itself deserves critical examination (Cave/Dihal 2020). It is evident that AI requires a cultural change, and we must acknowledge that everyone plays a role in driving this change. The cultural sector, in particular, has a vital role in creating new AI cultures. Culture and art can help us reimagine new futures, engage with the subject matter, and make the challenges of our digitizing society tangible.

Before concluding, I would like to share two interesting best practices that can contribute to the decolonization of AI. Firstly, using clear language around AI is crucial. Initiatives promoting accurate reporting about advanced technologies such as robots and AI have been introduced. Guidelines, like those initiated by Ben Shneiderman for journalists, emphasize the importance of clarifying human initiative and control (Shneiderman 2022). Instead of suggesting that computers act independently, we should highlight the human programming behind these actions. Furthermore, it is essential to avoid using of human verbs such as ‘understand’, ‘learn’, or ‘think’ to describe computers in human-like ways. We must emphasize the human labour involved in operating AI systems.

Secondly, we need better images of AI. The current dominant images perpetuate misconceptions and limit public understanding of AI systems. Initiatives like the non-profit collaboration ‘Better Images of AI’ are researching, curating, and providing more accurate representations of AI. They aim to showcase a broader range of humans and cultures, highlight AI’s human, social, and environmental impacts, and depict the human labour involved. They promote authenticity, inclusivity, and creativity by encouraging a shift in imagery.¹

Conclusion

I started with the importance of taking recourse to theory by understanding the logic(s) behind certain beliefs and reflecting on the language we use. We must acknowledge that these logics are primarily driven by capitalist ideologies prioritize profit generation over human wellbeing and rights. Moreover, they are often underpinned by discriminatory and racist ideologies that reinforce hierarchies and superiority, privileging ‘Whiteness’ in particular. To effect meaningful change, we must resist the temptation to jump to solutions without understanding the theory. Simple ethical principles like ‘transparency’ or ‘accountability’ can lead to ethics-washing and divert attention from the core issues at stake.

From a historical perspective, we have seen that transparency as an accountability mechanism has not been helpful—because it has been proven that it is where it starts and ends (Ananny/Crawford 2018). Take, for example, the case of Islamophobia, it is evidently transparent that this minority group has been excluded for decades and still is. It is also increasingly recognized that such forms of exclusion and biases are being perpetuated through technologies optimized for a specific type of group (Samuel 2021). But why are things not changing?

Let me end with a great quote by bell hooks and her reflection on the importance of theory:

1 <https://betterimagesofai.org/> (accessed in August 2023).

I came to theory because I was hurting—the pain within me was so intense that I could not go on living. I came to theory desperate, wanting to comprehend—to grasp what was happening around and within me. Most importantly, I wanted to make the hurt go away. I saw in theory then a location for healing (hooks 2014).

Theory can be confronting because it showcases our biases and the dark side of humanity, but it also enables us to recognize the injustices present in this world; it allows us to understand living conditions that should not be tolerated. As bell hooks argues: theory can be liberating—and should always come first, prior to interventions.

References

- Ananny, Mike/Crawford, Kate (2018). Seeing without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability. *New Media & Society* 20 (3), 973–89. <https://doi.org/10.1177/1461444816676645> (all URLs here accessed in August 2023).
- Benjamin, Ruha (2019). *Race after Technology: Abolitionist Tools for the New Jim Code*. Cambridge/Medford, Polity.
- Cave, Steven (2020). The Problem with Intelligence: Its Value-Laden History and the Future of AI. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 29–35. <https://doi.org/10.1145/3375627.3375813>.
- Cave, Steven/Dihal, Kanta (2020). The Whiteness of AI. *Philosophy & Technology* 33 (4), 685–703. <https://doi.org/10.1007/s13347-020-00415-6>.
- Crawford, Kate (2021). *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven/London, Yale University Press. <https://doi.org/10.12987/9780300252392>.
- De Graaf, Maartje M. A./Hindriks, Frank A./Hindriks, Koen V. (2021). Who Wants to Grant Robots Rights? Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, 38–46. <https://doi.org/10.1145/3434074.3446911>.
- Ensmenger, N. (2018). The environmental history of computing. *Technology and Culture* 59 (4), 7–33.
- hooks, bell (2014). *Teaching to Transgress*. New York, Routledge. <https://doi.org/10.4324/9780203700280>.
- Lanier, Jaron/Weyl, Glen (2020). AI is An Ideology, Not A Technology. *Wired*. Available online at <https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/>.
- Neff, Gina (2013). Why Big Data Won't Cure Us. *Big Data* 1 (3), 117–23. <https://doi.org/10.1089/big.2013.0029>.

- Noble, Safiya Umoja (2018). *Algorithms of Oppression*. New York, NYU Press. <https://doi.org/10.2307/j.ctt1pwt9w5>.
- Okafor-Yarwood, Ifesinachi/Adewumi, Ibukun Jacob (2020). Toxic Waste Dumping in the Global South as a Form of Environmental Racism: Evidence from the Gulf of Guinea. *African Studies* 79 (3), 285–304. <https://doi.org/10.1080/00020184.2020.1827947>.
- O’Neil, Cathy (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, Crown.
- Piccinini, Gualtiero/Bahar, Sonya (2013). Neural Computation and the Computational Theory of Cognition. *Cognitive Science* 37 (3), 453–88. <https://doi.org/10.1111/cogs.12012>.
- Samuel, Sigal (2021). AI’s Anti-Muslim Bias Problem. *Vox*. Available online at <https://www.vox.com/future-perfect/22672414/ai-artificial-intelligence-gpt-3-bias-muslim>.
- Shneiderman, Ben (2022). Guidelines for Journalists and Editors about Reporting on Robots, AI, and Computers. Medium. Available online at <https://medium.com/@ben.shneiderman/guidelines-for-journalists-and-editors-about-reporting-on-robots-ai-and-computers-6a69c3b813cd>.
- Stilgoe, Jack (2020). *Who’s Driving Innovation: New Technologies and the Collaborative State*. Cham, CH, Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-32320-2>.
- Strubell, Emma/Ganesh, Aananya/McCallum, Andrew (2019). Energy and Policy Considerations for Deep Learning in NLP. arXiv:1906.02243. <https://doi.org/10.48550/arXiv.1906.02243>.
- Van Es, Karin/Schäfer, Mirko T. (Eds.) (2017). *The Datafied Society: Studying Culture Through Data*. Amsterdam, Amsterdam University Press. <https://doi.org/10.151/9789048531011>.
- Watson, John B. (1913). Psychology as the Behaviorist Views It. *Psychological Review* 20 (2). <https://doi.org/10.1037/h0074428>.
- Young, Jason C. (2020). Environmental Colonialism, Digital Indigeneity, and the Politicization of Resilience. *Environment and Planning E: Nature and Space* 4 (2), 230–51. <https://doi.org/10.1177/2514848619898098>.