Predicted benefits, proven harms

How AI's algorithmic violence emerged from our own social matrix

Dan McQuillan
6th June 2023

Artificial intelligence (AI) finally seems to be living up to the hype. We can chat with it, ask it to write an essay or use it to generate a photo-realistic image of anything we can imagine. At the same time, these capacities seem to confirm the dystopian potentials of machine intelligence. If AI can already "pass" law [https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4335905] and medical licence exams [https://www.medrxiv.org/content/10.1101/2022.12.19.22283643v2.full] , how long before it turns into Skynet, the self-aware AI that turns on its creators in the *Terminator* films?

The discourse surrounding artificial intelligence is, however, a spectacle. It acts as a smokescreen that diverts our attention from both the far-reaching impacts of actual AI, as it exists today, and the disturbing politics it incubates. Grasping the sociological import of AI means engaging with what's going on under the hood. Understanding how the machinery works is a precondition for realising that in fact, there is nothing that even borders on "intelligence" here, and for reading across to its social effects and their repercussions.

Indeed, lifting the lid on AI allows us to reveal its deep entanglement with social relations, rather than synaptic ones. There is a double illusion at play: AI is misrepresented as a form of cognition, underpinned by the misunderstanding of mind as computational. But AI doesn't emerge from a sci-fi future brain: it originates within the social matrix we already inhabit. The operational concepts are not intelligence or even learning, but optimisation (the fundamental mechanism), layers (stages in the internal operations), vectors (the internal representation of words, images or anything else) and loss function (the construction of the difference between current and desired output).

While there is much work to be done in uncovering the many social consequences of these apparent abstractions, there is a single social thread running through contemporary AI from end to end: algorithmic violence, which encompasses both symbolic violence and predictive interventions.

Algorithmic violence behind the scenes

AI increases what historian and philosopher Hannah Arendt called "institutional thoughtlessness" – the inability to critique instructions or reflect on consequences. Its objective devaluations interface all too readily with existing bureaucratic cruelties, scaling administrative violence in ways that intensify structures of inequality, such as when AI is used to aid decisions about which patients are prioritised in healthcare or which prisoners are at risk of reoffending.

AI also relies on extractive violence because the demand for low-paid workers to label data or massage outputs [https://time.com/6247678/openai-chatgpt-kenya-workers/] maps onto colonial relations of power, while its calculations demand eye-watering levels of computation and the consequent carbon emissions, with a minimum baseline for training that is equivalent to planeloads of transatlantic air passengers [https://arxiv.org/abs/1906.02243] .

The compulsion to show balance by always referring to AI's alleged potential for good should be dropped; we must acknowledge that the social benefits are still speculative, but the harms have been empirically demonstrated. We must recognise that this algorithmic violence is legitimised by AI's claims to reveal a statistical order in the world that is superior in scale and insight to our direct experience. This also constitutes – among other things – a logical fallacy and a misuse of statistics.

Statistics, even Bayesian, does not extrapolate to individuals or future situations in a linear-causal fashion, and completely leaves out outliers. Ported to the sociological settings of everyday life, this results in epistemic injustice [https://www.lighthousereports.com/investigation/suspicion-machines/] , where the subject's own voice and experience is devalued in relation to the algorithm's predictive judgements.

While on the face of it AI may seem to produce predictions or emulations, I would argue that its real product is precaritisation and the erosion of social security; that is, AI introduces vulnerability and uncertainty for the rest of us, whether because our job is under threat of being replaced or our benefits application depends on an automated decision. Instead of sci-fi futures, what we get is the return of 19th-century industrial relations and the dissolution of post-war social contracts.

Analysing the apparatus

Feminist philosopher Karen Barad describes the dualisms of the world we experience as the consequence of particular exclusions (or "cuts") that result from specific arrangements of "apparatuses". These exclusions are neither social nor material, but they create new distinctions between the social and the material. A sociological analysis of AI needs to be grounded in its materiality and approach AI as an apparatus of this kind.

From this perspective, AI is not a way of representing the world but an intervention that helps to produce the world that it claims to represent. Setting it up one way or another changes what becomes

naturalised and what becomes problematised. Who gets to set up the AI becomes a crucial question of power.

Approaching AI as a non-dualistic apparatus, in Karen Barad's sense, enables us to stay with the trouble of, for example, so-called Large Language Models (LLMs) such as ChatGPT. It allows us to assert that, yes, LLMs are "bullshit engines", and to mean that in both technical and social senses at the same time. They are bullshit engines in the technical sense because they optimise on the generation of plausible natural language text based not on a model of causal relations and real meanings but blindly, based on preprogrammed statistical patterns in a vast (but never full) corpus of existing language. They are also bullshit engines in the social sense because, like marketing and middle management, their apparently reasoned arguments are actually completely detached from lived experience.

The hype over LLMs is stress-testing the possibility of real critique, as academia is submerged by a swirling tide of self-doubt about whether to reluctantly assimilate them and public narratives are dominated by a focus on existential risk that sucks attention from everyday harms.

Dangerous as this is, there is an even more pressing issue: AI is intensifying existing necropolitics, the term used by Cameroonian political theorist Achille Mbembe to refer to the use of social and political power to dictate how some people may live and how some must die, and we urgently need to resist that.

Necropolitics, an invisible force

While deep learning isn't on the way to civilisation-threatening superintelligence, it is all too resonant with the re-emergence of reactionary social and political agendas. For example, AI creates partial states of exception; that is, situations where our rights are removed in some sense without recourse to any due process. This occurs through a largely invisible redlining of everyday lives across spaces as diverse as the Airbnb platform, the Amazon warehouse or the immigration visa system.

Post-structuralist theorists Gilles Deleuze and Felix Guattari warned us that fascism creeps up well before it becomes obvious in political movements, for example as small instances of "microfascism" that lurk imperceptibly in our everyday life. AI concentrates and condenses these microfascisms through the way it enacts states of exception, transformative violence and essentialised othering.

The net effect of neural networks acting in the world is "necropolitical"; in other words, AI-assisted decisions and exclusions aggregate life-preserving pathways for some, while subjecting others to iterative neglect. An algorithmic condensation of social conditions results in the acceleration of differences such as racialised deprivation or disability. To put it crudely: AI can decide which humans are disposable.

These lifeboat logics are already at play at national borders that ensure migrants die in oceans or

deserts and in the reduction of welfare systems to below the bare survival minimum, and were starkly visible during the COVID-19 [https://www.ft.com/content/0e63541a-8b6d-4bec-8b59-b391bf44a492] pandemic. Peering into the core of contemporary AI with the stereo vision of the technical and the sociological shows its ongoing entanglement with eugenics; AI is great at ranking people into superior and inferior and it can bring to life the commitment of so many in Silicon Valley to ideas such as Long Termism [https://www.currentaffairs.org/2021/07/the-dangerous-ideas-of-longtermism-and-existential-risk] .

The very concept of artificial general intelligence (AGI), which is the ultimate goal of companies such as OpenAI and DeepMind, but at this point is still far from being attainable, is rooted in particular types of applications of mathematics and statistics to social matters such the history of IQ and the eugenics movement.

The real risk presented by AI is not imminent superintelligence, but the way it will appeal to increasingly desperate systems of governance casting about for populist solutionism in the face of the overlapping crises of neoliberalism. What needs to be headed off is the apparatus that will emerge from a convergence between Silicon Valley (underpinned by ideologies such as Long Termism) and the surge in far-right politics pretty much everywhere.

Social resistance and marginalised people

In this work of resistance, the most relevant perspectives are not those of "AI [https://www.theverge.com/2019/3/27/18280665/ai-godfathers-turing-award-2018-yoshua-bengio-geoffrey-hinton-yann-lecun] godfathers [https://www.theverge.com/2019/3/27/18280665/ai-godfathers-turing-award-2018-yoshua-bengio-geoffrey-hinton-yann-lecun] " but of the marginalised people whose lives are most impacted. The standpoints from which a constructive learning against the machine can arise are exactly those minoritised experiences to which a considerable body of sociological research is already dedicated. A radical feminist and decolonial critique of AI's approach to the world, which feminist science and technology scholar Donna Haraway would describe as the "view from nowhere", can pose embodied knowing against its epistemic hubris while simultaneously shining a light on the invisibilised social relations which hold up AI and the rest of society.

One way to counter the risk of venture capitalists turning AI into a weapon against the rest of us would be to put *care* at the centre of both research and action. As another feminist science and technology scholar Maria Puig de la Bellacasa put it, we need to see AI as a "matter of care".

While AI's internal models impose thoughtless optimisations, we can push for apparatuses that valorise relationality and embrace difference.

Socially responsible computing is possible

A movement to replace AI will call into question both material and conceptual boundaries and will mobilise care and mutual aid in the face of abstractions that enact actual violence. Researchers and social movements can't restrict themselves to reacting to the machinery of AI but need to generate social laboratories that prefigure alternative socio-technical forms. The need to put down all machinery hurtful to the commonality calls for a structural renewal based on socially useful production and the creation of more convivial technologies.

Necropolitical power is materialised in current forms of advanced computation, but I believe another computing is possible; one that supports the more-than-human solidarities needed for life beyond the climate crisis. Neither capitalist realism nor AI realism are realistic any more. We need to move from cyberpunk to solarpunk [https://thefirethesetimes.com/2023/01/20/lukadowell/] , to radically rework our technopolitical arrangements away from the nihilism of neural networks and towards adaptive and open-ended systems.

A convivial techno-social system is one that can come into balance with its environment, where recursive coordination starts with autonomy from below. A system that doesn't hallucinate a total view of reality but negotiates respectfully with the different and the unexpected. Rather than sociology restricting itself to a sociological account of emerging AI harms, I hope that it can inform the process of developing this alternative techno-social apparatus.

References and further reading

Arendt, H. (2006). *Eichmann in Jerusalem: A Report on the Banality of Evil* (1 edition). Penguin Classics.

Barad, K. (2007). *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press Books.

Bellacasa, M. P. de la. (2017). *Matters of Care* (3rd ed. edition). Univ Of Minnesota Press.

Deleuze, G., & Guattari, F. (2013). *A Thousand Plateaus*. Bloomsbury Academic.

Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575–599. https://doi.org/10.2307/3178066

Mbembe, A., & Corcoran, S. (2019). *Necropolitics*. Duke University Press.

About the author

Dan McQuillan

Dan McQuillan is Lecturer in Creative and Social Computing at Goldsmiths, University of London. After completing a PhD in experimental particle physics, he worked with people with learning disabilities

and mental health issues, created websites with asylum seekers, and worked in digital roles at both Amnesty International and the National Health Service. Dr McQuillan is author of *Resisting AI – An Anti-fascist Approach to Artificial Intelligence* (Bristol University Press, 2022) Twitter: @danmcquillan. Mastodon: @danmcquillan@kolektiva.social

Cite this work

McQuillan, D. (2023, June 6). Predicted benefits, proven harms: How AI's algorithmic violence emerged from our own social matrix [Online]. The Sociological Review Magazine. https://doi.org/10.51428/tsr.ekpj9730

Copyright