Review Article

Generative Artificial Intelligence: The Black Box of Modernity

Myke Healy*

Department of Education, Werklund School of Education, University of Calgary, Calgary, Canada

ABSTRACT

This paper examines the intersection of generative Artificial Intelligence (AI) with modernity, focusing on its colonial aspects and educational implications. By analyzing AI's role in high-profile military operations, the study underscores AI's expanding influence beyond technological domains. The research argues that AI, as a product of modernity, reinforces colonial structures through its development and deployment processes, critiquing AI's training methodologies that exploit labor in the Global South and perpetuate biases through skewed datasets. It addresses the environmental concerns associated with AI's energy-intensive nature and the opacity of its decision-making processes, metaphorically described as a "black box." In education, the study explores AI's integration, emphasizing potential academic integrity issues and the reinforcement of neoliberal paradigms. The research concludes by advocating for a critical, relational approach to AI in education, balancing technological advancements with human connection and ethical considerations, thus contributing to the ongoing discourse on AI ethics and offering insights into the complex interplay between technology, power dynamics, and educational practices in late modernity.

Keywords: Generative artificial intelligence; Academic integrity; Modernity; Coloniality; AI in warfare; AI and education; Large language models; AI bias; Data colonialism; Neoliberalism; Environmental impact of AI; Automation bias; Relationality in education; AI empire; Radical tenderness

INTRODUCTION

On Friday, November 27, 2020, Mohsen Fakhrizadeh, Iran's top nuclear scientist, was driving with his wife to their country house outside of Tehran when a remote-controlled Belgian-made FN MAG machine gun fired 15 bullets, killing him while leaving his wife uninjured. The New York Times revealed the intricacies behind the Israeli operation in September 2021, noting how Artificial Intelligence (AI) was essential in its execution. The operation's complexity lay in the remote relay: The physical gun was 1000 miles from the sniper, requiring AI to account for weapon recoil, the approaching car's speed, and facial recognition to confirm the targets.

The assassination foreshadowed the growing integration of AI in warfare. Today, the US Department of Defense is spending $500 million a year on the Replicator Initiative, aiming to deploy thousands of lethal autonomous drones across air, land, and sea, a strategy ominously dubbed "hellscape". Similarly, in the current Israel-Hamas war, the IDF's Directorate of Targets has used an AI platform called "The Gospel" to assist in selecting 12,000 targets in the Gaza Strip. As AI's role in warfare expands, concerns mount. Dr. Marta Bo of the Stockholm International Peace Research Institute warns of potential "automation bias" and over-reliance "on systems which come to have too much influence over complex human decisions". As a species generally so proud of our agency, outsourcing decisions of life and death to artificial intelligence is a stunning example of technology in late modernity.

For the majority of us in Vanessa Machado de Oliveira's "North of the North," the threat to our person of hells cape autonomous AI drones is non-existent. We may consider AI war tools as flotsam in our online feeds, an interesting tech story divorced from the trauma, misery, and death such systems are purpose-built to inflict with ruthless efficiency. Such examples serve as valuable signposts on our path of modernity, particularly related to artificial intelligence. As money, attention, human capital,

and voluminous resources are poured into AI across the globe, the impact of the technology is already being felt across multiple domains of human endeavor.

This paper will query the colonial aspects of generative artificial intelligence and how the technology is but the long extension of modernity at work in education. The inquiry begins by establishing the researcher's positionality, examines AI through the perspective of modernity, explores the colonialist embedded in AI training and utilization, and concludes by connecting these themes to the crisis of AI and academic integrity.

LITERATURE REVIEW

Positionality and place

Notwithstanding AI-aided assassinations and hells cape drones, I self-identify as an AI optimist and an early adopter, continuously interested in exploring new technologies. I am aware of how my positionality aligns with the stereotypes that have come to epitomize Silicon Valley, what Sasha Costanza-Chock troubles as "White male geek culture, replete with heteropatriarchal cultural structures, forms of humor, and mechanisms for normalizing white cis male standpoints". I have unlearned many of these ways of being, which has taken decades due to my margination in the fetid stew of toxic masculinity during my formative years at an elitist all-boys school and summer camps.

In researching this paper, I questioned how my position within colonial structures influences my comfort with artificial intelligence tools. Does my positionality align me with the creators of these technologies, inherently making me more at ease in this space? I further question how my role as a senior administrator at a technology-rich independent boarding school skews my perception of tech accessibility. My professional 'place' for the past 20 years has been a privileged bubble of technological abundance and connectivity. With this awareness, I understand that I am on the pointy end of the jagged frontier of AI adoption, and making pronouncements about generative AI's impact on education is pre-mature, problematic, and potentially myopic. When chatgpt came online on November 30, it was just one more tool to play with in a long line of previously tried software and hardware innovations. Within 24 hours, I could not help but feel this was something different.

The Black Box of modernity

Frontier Large Language Models (LLMs) and the chatbots like ChatGPT, Perplexity, Claude, and Gemini we use to interact with them are built on massive data sets of material scraped from the publicly available internet, licensed sources, and users. If we agree with the contention that the sum content of the internet is modernity in digital form, LLMs are but the offspring of this dataset, the distillation of modernity. Specifically, what this looks like is when we ask a chatbot a question, the algorithm queries the large language model, and the most statistically relevant answer is returned, built word after word. For example, if I ask a generative AI chatbot: "Complete the following phrase: that the cat sat on the…", the model will use deep neural net learning to find the most likely next word,

potentially choosing between "mat", "floor", "chair", and other statistically relevant words. While a facile example, it becomes startling when the same model is asked to comment on the intersections of modernity, coloniality, place, loss, and love: Topics that feel intuitively impossible to generate meaningful answers using statistical patterns. Indeed, the models are "black box processes understood only in terms of its inputs and outputs, rather than in terms of its internal processes". Given the popularity of generative AI tools, the opacity of their inner workings seems a suitable metaphor for modernity and how we interact with and benefit from the complex systems around us without fully understanding how they work [1].

What is not opaque about generative AI responses is the surety with which they are issued, whether verifiably accurate, biased, or hallucinated. The underlying training corpus of many LLMs use data from users on "Wikipedia, Reddit and YouTube who skew young, white, male and American; which means that overtly racist, sexist, and ageist perspectives are overrepresented in the data and many other viewpoints are excluded altogether".

Bias in data collection for LLMs has far-reaching implications. Indigenous voices, which traditionally unite "past and present in memory" through oral traditions, are inherently excluded from digital training corpora due to their non-textual nature. This exclusion not only skews the AI's knowledge base and output but also exemplifies a broader issue in late modernity: an increasing tendency to outsource human cognition and decision-making to digital tools, despite the questionable quality, omissions, and veracity of vast swaths of their internet-derived knowledge base [2].

In such a modernity scenario, "genuine teachings will be distorted by this market (for both sellers and customers) and the truth in it will be dissolved without people even noticing it". The full wrath and potential of the distortive power of generative AI are breathtaking, from the macro impacts of election interference to the micro terror of deep fake revenge porn and praying on the elderly through phishing scams using the cloned voices of loved ones. As Machado de Olivera notes, "A lot of time and energy will be wasted" in this time of dissolved and distorted truth [3].

The energy being wasted is not just human capital. Training large language models and running generative AI chatbots raises significant concerns about their electricity consumption and carbon emissions. The training alone of the first widely available LLM, Open AI's breakout GPT-3, consumed 1,287 megawatt-hours of energy, with estimated carbon emissions of 552 metric tons of CO_2, equivalent to a single person taking 550 roundtrips between New York and San Francisco.

The Electric Power Research Institute in Palo Alto estimates that a ChatGPT request requires 2.9 watt-hours to process, compared to 0.3 watt-hours for a Google search, making ChatGPT close to 10 times more electricity-intensive. As a point of comparison, a text-based ChatGPT query consumes as much energy as 2-3 hours of smartphone use. The computational demands of emerging AI capabilities in image, audio, and video generation have "no precedent" regarding projected future energy consumption [4,5].

Globally, data centers already account for around one percent of the world's greenhouse gas emissions and projections suggest that AI-intensive data centers could consume upwards of 9.1% of U.S. electricity generation annually by 2030. Modernity obscures this intensive energy use, as mundane-looking data centres in banal industrial parks process 1.1232 sextillion operations per hour.

Energy costs and intense neoliberal competition are driving researchers to focus on energy efficiency, including "model quantization, distillation, sparsification" and a constant drive to create more energy-efficient chips and training methods. Such efforts further obfuscate our understanding of the complexities of the black box between AI inputs and outputs.

As generative AI mirrors the ambiguities of modernity, our challenge particularly as educators-is to critically evaluate and understand the profound implications of relying on these opaque, energy-intensive systems that shape our perceptions, decisions, and realities

The coloniality of generative AI

The imprint of colonialism can be seen tangibly in the physical training of large language models and their underlying structure, use, and ownership. In her paper Ethics for the majority world: AI and the question of scale, Paola Ricaurte notes that the "political economy of AI cannot be separated from the effects of its design, development, use, deployment and disposal in an ethical framework for the majority world". The training of large language models through the outsourcing of Reinforcement Learning from Human Feedback (RLHF) is a case study in modern colonialism.

Tech companies in the global north are outsourcing human AI model training "for poverty wages and at great psychological detriment by data workers in majority-world countries". Data workers in Kenya were given poverty wages for piece work reviewing "graphic scenes of violence, self-harm, murder, rape, necrophilia, child abuse, bestiality and incest," causing "serious trauma". Despite the singular focus on equating AI with Silicon Valley, major Chinese companies use the same subcontractors [6].

In their impeccably researched article AI Empire: Unraveling the interlocking systems of oppression in generative AI's global order, Jasmina Tacheva and Srividya

Ramasubramanian offer the following striking analysis Generative AI cannot dissociate itself from the processes of extractivism, automation, essentialism, surveillance, and containment, which perpetuate historic structures of heteropatriarchal, colonial, racist, white supremacist, and capitalist oppression it has always already been built on them.

The authors note how Big Tech capitalism (now synonymous with 'AI Empire') is operating under the theory of "extractivism", where majority world human labour is exploited and appropriated in service of maximizing profitability. Apart from extracting human labour to train large language models, the AI Empire is accelerating what Thatcher et al. call data colonialism, characterized by the "asymmetrical power relation

between the individuals whose actions generate individual datums and those who come to own and profit from the big data they become".

Through the daily use of online tools that require us to opt-in to data collection prior to use, we are unwittingly part of 'big data' coloniality. The insatiable need for LLM training data is intensifying what Mejias and Couldry call dataficiation, which combines "the transformation of human life into data through processes of quantification, and the generation of different kinds of value from data". When using 'free' online services like Gmail, Snapchat, Google Maps, LinkedIn,

TikTok, and ChatGPT on a daily basis, the average user does not stop to think about how they are paying for these services by allowing their digital footprint to be datafied and sold on the open market.

DISCUSSION

In contexts where using all of this data contravenes national privacy laws, companies beta-test their wares in jurisdictions that lack such safeguards. For example, Mohamed et al. Note the algorithmic coloniality of ethics dumping, epitomized by Cambridge Analytica beta-testing its "algorithmic tools for the 2017 Kenyan and 2015.

Nigerian elections, with the intention to later deploy these tools in US and UK elections". The authors note that "these systems were later found to have actively interfered in electoral processes and worked against social cohesion". Summarized brilliantly, Tacheva and Ramasubramanian offer the following analysis of these "hidden conduits of AP Empire".

With invisible ubiquity, managerial efficiency, fake accuracy, and simulated objectivity, they orchestrate a chilling symphony of oppression, with marginalized communities worldwide bearing the brunt of AI Empire's automated violence [7-9].

The coloniality of generative AI emerges as a complex web of exploitation, extractivism, and asymmetrical power relations, perpetuating historical injustices under the guise of technological progress and innovation.

Future paths

In a secondary school setting, generative AI use by students is running headlong into the strictures of curriculum, assessment policies, and the domain-specific procedural knowledge of teachers. Today's high school students could be forgiven for wanting to forgo the labour and effort of writing a boilerplate school essay when a quick "plastic" version can be generated easily. As Vanessa Machado de Oliveira writes so poignantly.

Hospicing modernity

Because of the way modernity has socialized us to find security in certainty, stability, and predictability, as things start to fall apart we will seek out ways to fulfill these desires. We will seek plastic intellectuality, plastic sensuality, plastic artistry. We will shop for plastic certainties that can be consumed in soundbites.

Much has been written on social media, journalistic sources, and emerging research related to academic integrity concerns in an educational landscape permeated with generative AI tools. While the potential utility of AI tools in education is widely acknowledged, the

International Center for Academic Integrity (n.d.) notes that "inappropriate reliance can be a barrier to student learning and may lead to academic integrity breaches." Widening the lens, what is troubling is how generative AI accelerates the processes of modernity and neoliberalism in schools, further entrenching patterns of coloniality in education.

The enthusiasm for the potential of AI tutors and AI study aids is driven in part by the perceived gold rush for subscription services, and, more charitably, by the promise of potential universal access to inexpensive one-to-one tutoring. Modern education is awash in neoliberal programmatic 'fixes' in the form of training programs, software purchases, learning efficiencies, and the ever-expanding educational technology complex. These purported fixes often serve to reinforce existing power structures rather than address fundamental educational challenges.

AI tutoring is a compelling case study in this field. Proponents tout AI tutoring as a potential solution to Benjamin Bloom's long-elusive\2-sigma problem, referring to the gap in student achievement between one-to-one tutoring and traditional group instruction. Bloom observed "great differences in student cognitive achievement, attitudes, and academic self-concept under tutoring as compared with the group methods of instruction". Instead of educators struggling to adjust classroom methods to match the efficacy of individual tutoring, AI systems could theoretically provide personalized tutoring to all students at scale.

However, this promise of AI tutoring as a universal educational panacea is met with justified skepticism. Critics like Meyer argue that algorithmic tutoring is simply "not how good tutoring works." Indeed, any parent who watched their child struggle to learn online during the pandemic would be justified in their skepticism of purveyors of online learning solutions. AI tutoring exemplifies modern techno-solutionism, echoing colonial-era beliefs in universal, Western-centric educational panaceas (Supplementary File).

The relationality of learning between teacher and student is the candle in this darkness and a key component in mitigating academic dishonesty. The words of Dwayne Donald shine in this space, noting, "It is an ethical imperative to recognize the significance of the relationships we have with others, how our histories and experiences are layered and position us in relation to each other, and how our futures as people similarly are tied together". Through care, attention, and connection with students, teachers can hold modernity at bay in their classrooms and be better equipped to navigate academic integrity in the age of generative AI. Indeed, it is more important than ever to consider the "ethical imperative to see that, despite our varied place-based cultures and knowledge systems, we live in the world together with others and must constantly think and act with reference to these relationships".

With all its coloniality and problematic future potentials, the black box of generative AI and modernity is best faced with such relationality.

CONCLUSION

In the end, the black box outputs of generative AI are a mirror, reflecting our values through how, and on what, large language models have been trained. The immediacy of the challenge of academic integrity overshadows the subtler and pervasive ways AI systems will continue to reinforce colonial structures in our schools. While there are compelling use cases for AI in education, it also threatens to accelerate neoliberal trends and undermine essential relational aspects of learning. This is a time for human connection and Vanessa Andreotti's "radical tenderness" so educators can assist "with the birth of something new, which is potentially, but not necessarily, wiser".

AUTHOR'S CONTRIBUTION

I am the sole author of this paper and am 100% responsible for all content, errors, and omissions. All sources used in the preparation of this essay are cited in the reference list.

Use of generative artificial intelligence disclosure

The following generative AI tools provided writing and research assistance:

- Ongoing dialogue around rephrasing sentences for structure and clarity through Chat GPT-4o. OpenAI. (2024). Chat GPT4o (Large language model).
- Ongoing dialogue around rephrasing sentences for structure and clarity through Claude. Anthropic. (2024). Claude. (3.5 Sonnet version) (Large language model).
- Ongoing dialogue about the research topics through Perplexity.ai (subscription version) (2024). Perplexity.ai (AI Chatbot) (Large language model).
- General proofreading assistance through Grammarly (subscription version) (2024). Grammarly. (July 31 version) (Large language model).
- Additional general proofreading assistance through Quillbot. (July 31 version) (Large language model).

FUNDING

While I receive partial employer funding, there is no editorial impact, real or assumed.

CONFLICT OF INTEREST

I have no conflict of interest, financial or otherwise, related to this paper.

DIVERSITY STATEMENT

I have actively sought to include and highlight the contributions of women and other minority scholars. Recognizing the historical underrepresentation of female-identifying scholars in artificial intelligence, I prioritize ensuring that women-authored

sources are well-represented in my work. By deliberately including sources authored by women and scholars of colour, I aim to contribute to a more balanced and inclusive academic discourse.

REFERENCES

1. Aljbour J, Wilson T, Patel P. Powering intelligence: Analyzing artificial intelligence and data center energy consumption. Electric Power Research Institute (EPRI); 2024.

2. Bergman R, Fassihi F. The scientist and the A.I.-assisted, remote-control killing machine. The New York Times; 2021.

3. Bertram Gallant T. The other AI: The impact of artificial intelligence on academic integrity. 2023.

4. Bloom BS. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educ Res. 1984;13:1–15.

5. Challenge Success. Cheat or be cheated? What we know about academic integrity in middle and high schools and what we can do about it. 2012.

6. Clark J. Defense innovation official says Replicator initiative remains on track. U.S. Department of Defense. 2024.

7. Costanza-Chock S. Design justice: community-led practices to build the world's we need. 2020. [Google Scholar]

8. Dale R. GPT-3: What's it good for? 2020.

9. Davies H, McKernan B, Sabbagh D. 'The Gospel': How Israel uses AI to select bombing targets in Gaza. The Guardian. 2023.