Introduction

In 2023, we witnessed the publication of ChatGPT, an AI Large Language Model (LLM) that seems to herald a new phase of human civilisation. This recent breakthrough in AI technology has compelled most of our traditional human-based system to adapt to the changes ushered in by it.

From car assembly lines to self-driving cars (Knight, 2021); From receiving fast food orders to preparing meals (Thompson, 2023). Any discussion about anything in recent years revolves around AI, or more precisely – increasing efficiency and productivity using AI. Unavoidably, the quandary of whether to integrate AI technologies into our education system, the backbone of human civilisation, has become the target of public scrutiny and topic of hot debate.

Ever since our hunter-gathering days, our education system has always been human-based. Human teachers play a vital role in a person's development – from transmitting knowledge about toolmaking to inculcating cultural values (Boyette and Hewlett, 2017). The system is long-established, and through hundreds of thousands of years of trial and error, empirically validated. Through human teachers, homo sapiens have produced countless great leaders, inventors, thinkers, content-creators, such as Plato, Nikola Tesla, and Stephen King, just to name a few, who often became teachers themselves to pass down their exceptional knowledge to the next generations. Throughout our existence as a species, this has always been the case. However, this soon might not be. The social landscape that we were once so familiar with might soon undergo a transformation that renders it beyond recognition (Yuval Noah Harari, 2019).

At the rate at which AI technology is advancing, it seems to be inevitable that AI will play an increasingly crucial role in the education system. However, this raises an important question, to what degree exactly? Should we replace all human teachers with AI robots? It is impossible to predict the future so we can only make inferences with the current state of AI and regulations related to it. Based on this premise, then, the answer will be no - we should not replace all human teachers with AI robots, at least, not with what we have right now.

The immaturity of AI

In an interview with the ABC News, Sam Altman, the CEO of OpenAI, the company that developed ChatGPT, admitted that he is “A little bit scared of AI” and urges the public to be more cautious with their own product (Ordonez, Dunn and Noll, 2023). In another interview on the podcast show Lex Fridman Podcast, he admitted that AI technology is still in its infancy. Using GPT-3.5, an LLM developed by OpenAI, as an example: the LLM is prone to errors, and oftentimes entirely confident about its absurd and factually wrong statements. This also happens to human teachers, but with one key difference: Humans are capable of correcting themselves, whereas LLMs are completely dependent on the data they have been trained on and the update packages distributed by their manufacturers.

One of the biggest benefits of AI robots is their connectivity and updateability (Yuval Noah Harari, 2019). However, what if the information is wrong? Suppose that we were to replace all human teachers with AI robots powered by any LLM we have right now. With how immature the technology still is, hundreds of millions of children across the world could all be imparted with factually inaccurate information all at the same time, by the same LLM. Human teachers, on the other hand, are individuals who have the ability to think and to correct their own mistakes independently by simply fact-checking from books or from the internet, and therefore do not have such problem.

Furthermore, it is almost guaranteed that most teachers would not share dangerous, extreme knowledge with their students, even when their students specifically asked them to. LLM, however, is far from bullet-proof when it comes to censoring such information. A recent study has found a universal way to “jailbreak” the safety mechanism of various LLMs and forced them to reveal potentially destructive information to anyone, even children (Zou et al., 2023). What if the students of AI robot teachers were to learn these tricks and used it on their “teachers” in order to learn how to make explosives, for example?

The Danger of AI generated contents

Even if we were to ignore all the technical difficulties that AI technologies are facing at this moment, they would not be free of problems either. One such problem is the ethical issues that arise with respect to AI. For example, who gets to decide what is wrong and what is right?

Different perspectives and cultural values lead to distinct notions of right and wrong. How many people did Alexander the Great slaughter; how many tribes exterminated; how many families brutally torn apart during his campaign in the middle east, and before he was bestowed with the title "the Great"? From a 21st century humanist perspective, the great military general Alexander was far from being "Great" - he and his subordinates were war criminals and should be put on trials in the name of justice. From a military tactician's perspective, however, Alexander was a military genius who totally lived up to his name (Worthington, 2016). It is therefore difficult to decide that which perspective should the AI teachers adopt, and who gets to hold such unparalleled power to programme the brains of the next generations?

One could certainly argue that the same problem already exists in our current human-teacher-based education system. However, with how controllable and distributable AI robots are, the situation could be worse if we replaced all the human teachers. It is difficult to ensure that all teachers align with the political ideologies of a regime. The brainwashing techniques are not perfect, and the amount of resource and time required to produce a "model-citizen" are expensive - Human beings could think. They would not necessarily believe and act as the authority demands (Hunter, 2015). AI robots, at least as of today, could not formulate their own perspectives. LLMs are parrots that simply regurgitate the information provided to them (Karthik Valmeekam et al., 2022). It would hence be much easier to design effective AI propaganda machines than human ones.

Decline in Internet Freedom

Exacerbated by AI, in 2023, we witnessed the 13th consecutive year of decline in internet freedom (Ryan-Mosley, 2023). Power-hungry politicians have been quick to adopt such powerful technology to serve their questionable purpose: Various Venezuelan state media outlet and pro-Burkinabè media have been discovered of using AI technology such as deepfakes to spread pro-government propaganda (Janowitz, 2023). The result would be even more destructive, if this were to happen on a larger scale with an entire generation of innocent school children being brainwashed with political agendas whose only purpose, in the name of education, was to keep the regime in power.

Conclusion

AI technology is a double edge sword: With how powerful it is, it could greatly improve the living standards of many around the world. It could also, however, create chaos. We should certainly, and perhaps inevitably, cultivate such power to fuel our advancement as a species - with moderation and patience, however. Our understanding of AI is still too immature for us to adopt it on a wider scale. We need to make the technology itself more robust. We need to update our laws and regulations. We need to have better back-up plans if things go wrong. Once we have achieved these, perhaps then, we would be able to fully harness the unmatched potential of AI to improve our daily lives, including designing the perfect teacher that would help us mold the next generations of scientists, artists, philosophers, and leaders.

Reference list

Bowman, S.R. (2023). Eight Things to Know about Large Language Models. *arXiv (Cornell University)*. doi:https://doi.org/10.48550/arxiv.2304.00612.

Boyette, A.H. and Hewlett, B.S. (2017). Teaching in Hunter-Gatherers. *Review of Philosophy and Psychology*, [online] 9(4), pp.771–797. doi:https://doi.org/10.1007/s13164-017-0347-2.

CNBC News (2017). *Putin: Leader in artificial intelligence will rule world*. [online] CNBC. Available at: https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html [Accessed 6 Oct. 2023].

Fridman, L. (2023). *#367 – Sam Altman: OpenAI CEO on GPT-4, ChatGPT, and the Future of AI | Lex Fridman Podcast*. [online] Lex Fridman Podcast. Available at: https://lexfridman.com/sam-altman/ [Accessed 6 Oct. 2023].

Hunter, J. (2015). *Brainwashing in a Large Group Awareness training? : the Classical Conditioning Hypothesis of brainwashing.* [online] Available at: http://hdl.handle.net/10413/14468 [Accessed 8 Oct. 2023].

Janowitz, N. (2023). *Venezuela Is Using Fake AI American Newscasters to Spread Disinformation*. [online] Vice News. Available at: https://www.vice.com/en/article/z34jge/venezuela-ai-newscaster-disinformation [Accessed 8 Oct. 2023].

Karthik Valmeekam, Olmo, A., Sarath Sreedharan and Subbarao Kambhampati (2022). Large Language Models Still Can’t Plan (A Benchmark for LLMs on Planning and Reasoning about Change). *arXiv (Cornell University)*. doi:https://doi.org/10.48550/arxiv.2206.10498.

Knight, W. (2021). *Ford’s ever-smarter robots are speeding up the assembly line*. [online] Wired. Available at: https://www.wired.com/story/fords-smarter-robots-speeding-assembly-line/ [Accessed 6 Oct. 2023].

Ordonez, V., Dunn, T. and Noll, E. eds., (2023). *OpenAI CEO Sam Altman says AI will reshape society, acknowledges risks: ‘A little bit scared of this’*. [online] ABC News. Available at: https://abcnews.go.com/Technology/openai-ceo-sam-altman-ai-reshape-society-acknowledges/story?id=97897122 [Accessed 6 Oct. 2023].

Ryan-Mosley, T. (2023). *How generative AI is boosting the spread of disinformation and propaganda*. [online] MIT Technology Review. Available at:

https://www.technologyreview.com/2023/10/04/1080801/generative-ai-boosting-disinformation-and-propaganda-freedom-house/ [Accessed 8 Oct. 2023].

Thompson, B. (2023). *This AI chef has learned to make you a salad, but do you want it to?* [online] New Atlas. Available at: https://newatlas.com/technology/ai-chef-video-learning-salad/ [Accessed 6 Oct. 2023].

Worthington, I. (2016). *Alexander the Great*. [online] Serious Science. Available at: https://serious-science.org/alexander-the-great-6490 [Accessed 6 Oct. 2023].

Yuval Noah Harari (2017). *Homo deus a brief history of tomorrow.* London Vintage.

Yuval Noah Harari (2019). *21 Lessons For The 21St Century.* Spiegel & Grau.

Zou, A., Wang, Z., Kolter, J.Z. and Fredrikson, M. (2023). *Universal and Transferable Adversarial Attacks on Aligned Language Models*. [online] arXiv.org. doi:https://doi.org/10.48550/arXiv.2307.15043.