Beyond the Veil of Algorithms: The Dual Faces of Large Language Models and Generative AI in Shaping Public Knowledge

Our knowledge does not end at the four walls of academia, or even when we reach our age of majority. It is what we reach for, have access to, and is inflicted upon us. Thus, knowledge is both a lifelong odyssey and a defence from disinformation or persuasion, an ever-evolving narrative written by the diverse sources of information that surround and engage us daily.

In the era of rapid technological advancement, Large-Language Models (LLMs)[1] like GPT 4, Q*, and the possibility of Artificial General Intelligence (AGIs) have emerged as transformative tools with the potential to shape the future of education, information dissemination, and even propaganda. With their new ability to process and generate text and images with the multimodal[2] approach, I will explore a student's perspective on a world increasingly driven by artificial intelligence.

Influence on Academia and Education System

---

[1] "A large language model (LLM) is a deep learning algorithm that can perform a variety of natural language processing (NLP) tasks. Large language models use transformer models and are trained using massive datasets — hence, large. This enables them to recognize, translate, predict, or generate text or other content." Elastic, *What is a Large Language Model (LLM),* accessed 30 December 2023 <https://www.elastic.co/what-is/large-language-models>

[2] "…various data types (image, text, speech, numerical data) are combined with multiple intelligence processing algorithms to achieve higher performances…" Rouse, M, 4 July 2023, *Multimodal AI (Multimodal Artificial Intelligence)*, Techopedia, accessed 30 December 2023 < https://www.techopedia.com/definition/multimodal-ai-multimodal-artificial-intelligence>

Dang Lin

Large-Language Models revolutionise the landscape of education. Text generation and review are the most laborious tasks in education, and leveraging LLMs can lead to the analysis and generation of unprecedented amounts of textual content and the completion of complex natural language processing tasks.[3]

According to Yan, L, et. al., LMM's natural language processing capabilities enable interactive and adaptive learning experiences. Through personalised tutoring and content creation (generating flash cards or custom learning modules), LLMs such as ChatGPT have the potential to cater to individual learning styles, fostering a more effective and engaging educational environment[4], allowing teachers to "promote autonomy, competence, and relatedness, significantly enhancing student support and building a more effective learning environment."[5] Steps to integrate this have begun with Khan Academy's Khanmigo, a "Tutor for learners. Assistant for teachers."[6], an educational AI tool.

This technology can support learners with special needs, alleviating strain on our overextended education system and providing opportunities by improving student support services and sustaining early intervention for high-risk students. Ways ChatGPT may advance student support include 24/7 availability and flexibility, seamless integration with

[3] Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). *Practical and ethical challenges of large language models in education: A systematic scoping review.* British Journal of Educational Technology. <https://doi.org/10.1111/bjet.13370>

[4] Chung Kwan Lo, 18 April 2023, *What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature* <https://www.mdpi.com/2227-7102/13/4/410> [4.1]

[5] Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., Alowais, S. A., Alshaya, O. A., Rahman, I., Al Yami, M. S., & Albekairy, A. M. (2023). *The emergent role of Artificial Intelligence, natural learning processing, and large language models in higher education and research.* Research in Social and Administrative Pharmacy. <https://doi.org/10.1016/j.sapharm.2023.05.016>

[6] Khan Academy. (2023). Accessed 25 December 2023, < https://www.khanacademy.org/khan-labs>

Dang Lin

assistive technology, development of 'soft skills' such as social skills and time management, and monitoring and intervention.[7]

However, AI cannot replace human intervention in serious at-risk scenarios, such as a student threatening physical harm to others or themselves; we must use this tool collaboratively with educators, counsellors, and support staff to provide comprehensive assistance. At this stage, we must view LLMs as a complementary tool used in conjunction with appropriate educational practices and interventions to promote inclusivity and accessibility.

Opportunities in Translation and Globalisation

The Italian saying 'Traduttore, Traditore'[8] belies the loss of nuance and preservation of the original meaning of the text through translation. Students find that poor translation can undermine their learning from foreign sources and those who speak overlooked languages can find themselves disconnected from key texts and the wider academic community. Language can lose its cultural, functional, and scientific meaning leading to poorer assessed performance[9] in critical scenarios such as test-taking where poor and unnuanced 'literal translation'[10] of questions leads to unintentional bias, and discrimination, against foreign language students even whilst taking tests in their native language.

---

[7] OpenAI. (2023). ChatGPT [Large language model]. Accessed 25 December 2023, <https://chat.openai.com/share/b5cf86ad-5478-48f4-878c-fb9951900f7e>

[8] "Translator, Traitor", part of the conversation around how translation is a betrayal. There comes the possibility that some text cannot be good faith translated into a different language, and how attempts to do say create entirely new texts that lost the author's true meaning.

[9] Ercikan, K. *Translation effects in international assessments,* International Journal of Educational Research 29 (1998) 543—553

[10] where phrases are translated with word-for-word, without contextual, grammatical or cultural understanding.

Dang Lin

According to OpenAI, LLMs can provide real-time translation. Overall, AI has made machine translation more dependable; rather than relying on word-to-word translation they use a larger database to analyse language patterns and comparisons to generate more fluent text that preserves the intended meaning,[11] enabling students to access information in their native languages without years of study to understand nuances and academic terminology. Those in poorer countries have less time and money to dedicate to learning new languages, therefore this inclusivity enhances the educational experience for individuals and promotes cross-cultural understanding and cooperation on a global scale.

As Alqahtani et. al. concludes, "Ultimately, AI has the potential to transform how educators deliver personalized learning and support, improving outcomes for all students."

Shifting Information Analysis and Synthesis

In other realms of information, LLMs contribute to a paradigm shift in how we access and interact with knowledge. Their capacity to analyse vast amounts of data allows for more efficient information retrieval and synthesis. Users can interact with these models to obtain insights, answers, and analyses, transforming the way information is disseminated and consumed.

This increased accessibility, however, comes with challenges. As LLMs become integral to information retrieval, the risk of misinformation and bias becomes more pronounced as those with poor AI literacy may misuse these tools. The algorithms powering language models are trained on large datasets, and if these datasets contain biases (racism, sexism, etc.), the

---

[11] Habash, F. *How is Artificial Intelligence Changing the Translation Services Industry?* Accessed 25 December 2023, < https://www.getblend.com/blog/artificial-intelligence-changing-the-translation-services-industry/>

Dang Lin

models will inadvertently perpetuate and amplify systemic issues. For example, Amazon paused its use of AI hiring. The 10-year training data was taken from mostly male candidates, leading to the model going against Amazon's stated policies of diversity and inclusion by favouring male applicants. [12] Striking a balance between accessibility and accuracy is crucial to harnessing the full potential of LLMs in the information domain; as the use of AI spreads to critical areas of law and medicine, where matters concerning human welfare and justice, conscientious and educated usage of AI become paramount to prevent user error mistakes that may lead to bias, which, in medicine, leads to poor treatment and delayed or inaccurate diagnoses and, in law, influence reasoning in decisions in uncertainty and distortion of the truth.

Rise of Propaganda and Disinformation

While LLMs offer tremendous benefits, they also raise concerns, particularly in the realm of propaganda and disinformation. The ability of these models to generate human-like text raises the spectre of information warfare, where false narratives can be propagated with unprecedented ease. As seen in the rise of deepfake technology, the line between authentic and manipulated content becomes increasingly blurred. This not only poses a threat to the integrity of information but also challenges the foundations of trust in our digital society. Entire disinformation campaigns flooding our social media platforms can become automated, at a scale and speed beyond human capacity. Fake articles coupled with false images can sway the masses, spark division, and incite violence.[13]

---

[12] Reuters, 10 October 2018, *Amazon scraps secret AI recruiting tool that showed bias against women*. Accessed 30 November 2023, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/>
[13] Anthony Mesnier, 29 May 2023, *LLMs, new Weapons of Mass Disinformation?,* accessed 29 November 2023, <https://towardsdatascience.com/llms-weapons-of-mass-disinformation-4def0dc3dc7>

Dang Lin

One of the world's most disseminated attempts at disinformation via AI-generated media was the notorious claim that President Joe Biden would open the U.S. Selective Service draft to women, supported by a deepfake video created in February that surfaced amid the Israel-Palestine conflict. "Remember," Biden says, "you are not sending your sons and daughters to war. You are sending them to freedom."[14]

Confronting Ethics and Possible Regulation

As we navigate the implications of Large-Language Models on education, information, and propaganda, we must address ethical considerations. The responsible development and deployment of these technologies require transparent practices, accountability, and a commitment to mitigating biases.[15]

In the context of education, ethical considerations involve ensuring that personalized learning experiences do not perpetuate inequality and that student data is managed with the utmost care. Balancing customization and fairness is crucial to building an educational landscape that benefits all. Additionally, as Joseph Weisenbaum argues, ethical concerns extend to potentially replacing educators and support staff with machines, which cannot currently emulate empathy. Students may feel devalued and alienated, as they require authentic

---

[14] Reuters, 19 October 2023, *Fact Check: Video of Joe Biden calling for a military draft was created with AI,* accessed 30 December 2023, <https://www.reuters.com/fact-check/video-joe-biden-calling-military-draft-was-created-with-ai-2023-10-19/>

[15] Müller, Vincent C., *Ethics of Artificial Intelligence and Robotics*, The Stanford Encyclopedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/fall2023/entries/ethics-ai>.

empathy from the people in these positions.[16]

There is debate surrounding AI liability in of health and justice sectors, cautious to experiment with AI technology for the risk of human endangerment. A UK-convened panel revised Asimov's Laws[17] in 2010, clarifying that AI is the responsibility of manufacturers and/or the operator in the Engineering and Physical Sciences Research Council (EPSRC)'s new Principles of Robotics[18] which work to protect 'human beings' to ensure that robotics research is "integrated into our society to the maximum benefit of all of its citizens" (Boden, M, et al., 2011). However, more likely sought to ensure citizen and consumer confidence in the legislation of robots as a product and the adequacy of the preexisting laws to protect consumer interests.[19] "The Principles of Robotics do not seek to determine what is possible; they seek to communicate advisable practices for integrating autonomous robotics into the law for the land." (Bryson, J, 2017)

The potential for biased algorithms and the intentional spread of misinformation is great. Establishing ethical guidelines for the training of LLMs, regularly auditing their outputs for biases, and incorporating human oversight in decision-making processes are steps toward responsible use.

---

[16] Joseph Weizenbaum (1976), *Computer Power and Human Reason*, San Francisco: W.H. Freeman & Company, ISBN 978-0-7167-0464-5.

[17] Isaac Asimov (1964). "Introduction*". The Rest of the Robots*. Doubleday. ISBN 978-0-385-09041-4.

[18] Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., … Winfield, A. (2011, April). *Principles of Robotics.* The United Kingdom's Engineering and Physical Sciences Research Council (EPSRC). <https://webarchive.nationalarchives.gov.uk/ukgwa/20210701125353/https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics>

[19] Joanna J. Bryson (2017) *The meaning of the EPSRC principles of robotics,* Connection Science, 29:2, 130-136, DOI: 10.1080/09540091.2017.1313817

Dang Lin

Regulation is necessary to curb the malicious use of LLMs for spreading disinformation. Stricter controls on the development and deployment of AI technologies in the context of information warfare are essential to safeguarding the integrity of public discourse.

Despite the barriers to discerning the fine line of discrimination, and if national interests may defend injustices, The European Commission's Artificial Intelligence Act2021)[20] represents a significant global regulatory effort. It proposes that AI of different assessed risk levels will have different rules around production and usage, allowing more dynamic regulation of AI. However, more comprehensive legislation is needed to address current limitations in fairness, including the intrinsic ambiguity of 'discrimination' as a concept philosophically and legally.

Open-sourced tools to detect and bring awareness to bias in AI have begun to appear in various civil societies.[21] Their use is limited, however, as companies are not legally incentivised to use or disclose these tools when using AI models.

Navigating the Future

In conclusion, LLMs possess substantial transformative potential for our education. Their capacity to revolutionize personalized learning, support students on a small scale, dismantle language barriers, and reshape the landscape of information access is indeed remarkable.

---

[20] European Parliament, 14 June 2023, *EU AI Act: first regulation on artificial intelligence,* accessed 1 December 2023, <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

[21] Google Developers. *Machine Learning Fairness | ML Fairness* Archived from the original on 2019-08-10. Accessed 25 December 2023, <https://web.archive.org/web/20190810004754/https://developers.google.com/machine-learning/fairness-overview/>

Dang Lin

Nevertheless, we must navigate the deployment of these models with a steadfast commitment to ethical considerations, accountability, and the judicious management of their capabilities. By doing so, we can harness the global power of AI while safeguarding the moral and philosophical foundations that underpin our responsibility to our people and the democratic exchange of ideas.

Dang Lin

Bibliography

Elastic, *What is a Large Language Model (LLM)*, accessed 30 December 2023

<https://www.elastic.co/what-is/large-language-models>

Rouse, M, 4 July 2023, *Multimodal AI (Multimodal Artificial Intelligence),* Techopedia,

accessed 30 December 2023 < https://www.techopedia.com/definition/multimodal-ai-

multimodal-artificial-intelligence>

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., &

Gašević, D. (2023). *Practical and ethical challenges of large language models in education:*

A systematic scoping review. British Journal of Educational Technology,

<https://doi.org/10.1111/bjet.13370>

Lo, C. K., 18 April 2023, *What Is the Impact of ChatGPT on Education? A Rapid Review of*

the Literature <https://www.mdpi.com/2227-7102/13/4/410> [4.1]

Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K.,

Alowais, S. A., Alshaya, O. A., Rahman, I., Al Yami, M. S., & Albekairy, A. M. (2023). *The*

emergent role of Artificial Intelligence, natural learning processing, and large language

models in higher education and research. Research in Social and Administrative Pharmacy.

<https://doi.org/10.1016/j.sapharm.2023.05.016>

Khan Academy. (2023). Accessed 25 December 2023, <

https://www.khanacademy.org/khan-labs>

Dang Lin

OpenAI. (2023). ChatGPT [Large language model]. Accessed 25 December 2023, <https://chat.openai.com/share/b5cf86ad-5478-48f4-878c-fb9951900f7e>

Ercikan, K. Translation effects in international assessments, International Journal of Educational Research 29 (1998) 543—553

Habash, F. *How is Artificial Intelligence Changing the Translation Services Industry?* Accessed 25 December 2023, < https://www.getblend.com/blog/artificial-intelligence-changing-the-translation-services-industry/>

Reuters, 10 October 2018, *Amazon scraps secret AI recruiting tool that showed bias against women.* Accessed 30 November 2023, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/>

Mesnier, A., 29 May 2023, *LLMs, new Weapons of Mass Disinformation?,* accessed 29 November 2023, <https://towardsdatascience.com/llms-weapons-of-mass-disinformation-4def0dc3dc7>

Müller, Vincent C., *Ethics of Artificial Intelligence and Robotics*, The Stanford Encyclopedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), <https://plato.stanford.edu/archives/fall2023/entries/ethics-ai>

Dang Lin

Reuters, 19 October 2023, *Fact Check: Video of Joe Biden calling for a military draft was created with AI,* accessed 30 December 2023, <https://www.reuters.com/fact-check/video-joe-biden-calling-military-draft-was-created-with-ai-2023-10-19/>

Joseph Weizenbaum (1976)*, Computer Power and Human Reason,* San Francisco: W.H. Freeman & Company, ISBN 978-0-7167-0464-5.

Isaac Asimov (1964). "Introduction". The Rest of the Robots. Doubleday. ISBN 978-0-385-09041-4

Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., … Winfield, A. (2011, April). *Principles of Robotics*. The United Kingdom's Engineering and Physical Sciences Research Council (EPSRC). <https://webarchive.nationalarchives.gov.uk/ukgwa/20210701125353/https://epsrc.ukri.org/research/ourportfolio/themes/engineering/activities/principlesofrobotics>

Joanna J. Bryson (2017) *The meaning of the EPSRC principles of robotics,* Connection Science, 29:2, 130-136, DOI: 10.1080/09540091.2017.1313817

European Parliament, 14 June 2023, *EU AI Act: first regulation on artificial intelligence*, accessed 1 December 2023, <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Google Developers. *Machine Learning Fairness | ML Fairness* Archived from the original on 2019-08-10. Accessed 25 December 2023,

Dang Lin

<https://web.archive.org/web/20190810004754/https://developers.google.com/machine-learning/fairness-overview/>

Dang Lin