

Navigating Europe's Artificial Intelligence Act: Application of LLMs in classrooms

Upasana DASGUPTA^{a*} & Rwitajit MAJUMDAR^{b+}

^a*Graduate School of International Studies, Laval University, Canada*

^b*Research and Education Institute for Semiconductors and Informatics,
Kumamoto University, Japan*

*upasana.dasgupta.1@ulaval.ca, +majumdar@kumamoto-u.ac.jp

Abstract: In 2018, OpenAI introduced the first version of the Generative Pre-trained Transformer (GPT), revolutionizing the future of Large Language Models (LLMs). This model demonstrated the potential of pretraining large-scale models with vast text data and then fine-tuning them for specific tasks to the public. LLMs have quickly penetrated educational environments, aiding students from various disciplines in tasks ranging from initiating research to drafting essays. While the latter may breach academic integrity, the former is highly beneficial, especially for exploring new areas or ideas. Comparatively, the user interactions with GPT might resemble initially with that of search engines, despite technological differences, as both provide answers to queries, often reflecting archived as well as mainstream views. The historical evolution of search engines, from Archie's database matching to Google's relevance-based ranking, highlights similar ethical considerations faced by both technologies. The development of search engines underscored the importance of accessible information, a principle equally relevant to GPT and LLMs today. This paper is written in light of recent coming into force of European Union's legal framework on artificial intelligence for the purpose of examining adoption of LLMs in classrooms, and argues for balanced regulations across jurisdictions that acknowledge both the immense educational potential of LLMs and the need for adherence to legal and ethical standards.

Keywords: AI Governance, AI in Education, Education Technology Design

1. Introduction

In 2018, OpenAI launched the first version of the Generative Pre-trained Transformer (GPT), revolutionizing Large Language Models (LLMs) by showing that pretraining on vast text data and fine-tuning for specific tasks is effective. Other companies like Google, Meta, and Anthropic, and even international ones like France's Mistral AI, are also advancing similar LLM AI models. Especially since December 2022 since launch of Open AI's GPT 3.5 which allowed easy user interface, LLMs have quickly penetrated classrooms across disciplines, aiding students from initiating research to writing assignments, though the latter raises academic integrity issues. LLMs can serve as valuable tools, supplementing search engines, despite technological differences. Both GPT and search engines offer answers to queries, prioritize mainstream views, and have raised ethical considerations. On May 21, 2024, the European Council approved the Artificial Intelligence Act to regulate AI systems in the EU which came into force on 1st August 2024 and is a step ahead of many countries including USA and Canada, which are still formulating comprehensive AI legislation. These regulations emphasize prioritizing AI research while ensuring safeguards such as upholding human rights and environmental protections. While recent studies (Ifenthaler et al. 2024) elicited educational technology experts' perspectives on use and impact of AI in education, an interdisciplinary perspective including recent global AI governance aspects which aims to safeguard the use and application is still nascent.

2. Recent Europe's Artificial Intelligence Act

Role of law and governance cannot be undermined as they provide the necessary order and certainty for several actors to exercise their freedom and carry on operations. Absent law and institutions, any realm would be subject to might being right. Law-makers around the world have responded to the remarkable progress of AI, particularly LLMs, ranging from blanket bans to wait-and-see approaches. While enforcing blanket bans until all the effects of a technology are fully understood may be an extreme stance, "wait-and-see" approach where the technology is left unregulated may not be ideal too.

Europe's Artificial Intelligence Act (the "AI Act") came into force on 1st August 2024 but the provisions come into application later, such as, rules on transparency for general-purpose AI systems will apply after 12 months and high-risk systems will have 36 months to comply with the requirements. The AI Act recognizes the importance to AI literacy, noting that "[p]roviders and deployers of AI systems shall take measures to ensure, to their best extent, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons on whom the AI systems are to be used." Among other things, the AI Act prohibits placing in market manipulative or deceptive techniques or placing in market systems that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation with the objective, or the effect, of materially distorting the behaviour to cause significant harm to any person. It may be noted that the AI Act classifies certain AI systems as high risk, including

- a) "AI systems intended to be used to determine access or admission or to assign natural persons to educational and vocational training institutions at all levels;
- b) AI systems intended to be used to evaluate learning outcomes, including when those outcomes are used to steer the learning process of natural persons in educational and vocational training institutions at all levels;
- c) AI systems intended to be used for the purpose of assessing the appropriate level of education that an individual will receive or will be able to access, in the context of or within educational and vocational training institutions at all levels;
- d) AI systems intended to be used for monitoring and detecting prohibited behaviour of students during tests in the context of or within educational and vocational training institutions at all levels."

While the said law classifies use of AI in education as high-risk, it also notes that the important role of AI systems in education especially for promoting high-quality digital education and training and to allow all learners and teachers to acquire and share the necessary digital skills and competences, including media literacy, and critical thinking, in order to take an active part in the economy, society, and in democratic processes. To put it simply, the AI Act does not ban use of AI in education but requires mitigating the risk from the high-risk AI systems by introducing certain checks, such as proper design and use and a robust risk-management system. The said law emphasizes that providers of general-purpose AI models such as generative AI have a special role and responsibility for transparency, since the models they provide may form the basis for a range of downstream systems. Hence, providers for generative AI as well as high-risk systems should maintain up to date documentation, which may be required to be furnished by competent public authorities. In addition, the providers of LLMs should keep in mind that any use of copyright protected content requires the authorization by the concerned rightsholder, unless relevant copyright exceptions and limitations apply. Moreover, general-purpose AI models can pose systemic risks which include, but are not limited to, any actual or reasonably foreseeable negative effects such as serious consequences to public health and safety; the dissemination of illegal, false, or discriminatory content; the facilitation of disinformation or harming privacy, democratic values and human rights; and perpetrate information that may negatively effect that could affect up certain communities, etc. General purpose AI models developers should, therefore, take steps to mitigate the systemic risks by adhering codes of practices and if despite efforts, an untoward

incident occurs, to keep record of such incidents. The AI Act also provide certain penalties for non-compliance with its provisions.

Earlier, in 2022, European Union had published a non-binding “Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators”. These guidelines are designed to help educators understand the potential impact of AI in education and raise awareness of possible risks, so that educators can engage with AI in classroom positively, critically and ethically. The 2022 document suggests that before introducing new AI systems in school, pilot projects should be run and contact should be maintained with AI system providers and there should be community engagement, regular monitoring and evaluation of the risks. It should be clear to the educator when AI can take over and when it cannot and the educators should be aware of the impact of using AI in education. Also there should be procedures in place to minimize the adverse effects of AI in education and accordingly, there should be opt-out procedures too.

3. Implications for LLM-driven services for Education

LLMs offer transformative potential in education by analyzing students' learning patterns and providing personalized materials tailored to individual needs (Holmes et al., 2023). They are being trialed for automatic grading and instant feedback on open-ended assignments, enhancing efficiency for teachers and timely responses for students (Gao, Thomas & Srinivasa, 2023; Nakamoto et al., 2023). Teachers can leverage LLMs to create customized lesson plans and practice questions, enriching the learning experience (Hu et al., 2024; Kabir and Lin, 2023). Additionally, LLMs' multilingual capabilities support inclusivity by translating materials for diverse linguistic backgrounds. However, addressing technical limitations like "hallucinations," along with ensuring data privacy, mitigating bias, and upholding ethical standards, is crucial for responsible integration into education.

The latest European law on AI advises against a blanket ban on LLMs while urging developers and educational service providers to adhere to the rule of law and is a welcome step in the right direction which may be emulated in other jurisdictions. Although AI laws are being implemented globally, they remain inconsistent and unharmonized, creating confusion due to the borderless nature of technology. Common aspects in these laws include calls for greater transparency and reduction of algorithmic bias, but no country or organization has yet defined the necessary technical standards. A 2024 UN General Assembly Resolution (UN Doc A/78/L.49) emphasizes the importance of safe and trustworthy AI for sustainable development, reflecting a shared global perspective. However, it remains recommendatory and lacks detailed technical standards, which need to be definitely laid out while being adaptable, given the rapid advancements in AI.

Acknowledgements

This work is partially funded by Fonds de recherche Société et culture and JSPS KAKENHI JP22H03902

References

- European Union, Artificial Intelligence Act 2024 read with Annex III.
- European Commission, Directorate-General for Education, Youth, Sport and Culture, (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/153756>
- Gao, R., Thomas, N., & Srinivasa, A. (2023, October). Work in Progress: Large Language Model Based Automatic Grading Study. In 2023 IEEE Frontiers in Education Conference (FIE) (pp. 1-4). IEEE.

- Holmes, W. et al. (2023). AI and Education. A View Through the Lens of Human Rights, Democracy and the Rule of Law. Legal and Organizational Requirements. AIED 2023. vol 1831. Springer, Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development, 2024, UN Doc A/78/L.49
- Hu, B., Zheng, L., Zhu, J., Ding, L., Wang, Y., & Gu, X. (2024). Teaching Plan Generation and Evaluation With GPT-4: Unleashing the Potential of LLM in Instructional Design. IEEE Transactions on Learning Technologies.
- Ifenthaler, D., Majumdar, R., Gorissen, P., Judge, M., Mishra, S., Raffaghelli, J., & Shimada, A. (2024). Artificial intelligence in education: implications for policymakers, researchers, and practitioners. Technology, Knowledge and Learning.
- Kabir, M. R., & Lin, F. (2023, July). An LLM-Powered Adaptive Practicing System. In AIED 2023 workshop on Empowering Education with LLMs-the Next-Gen Interface and Content Generation.
- Nakamoto, R., Flanagan, B., Yamauchi, T., Dai, Y., Takami, K., & Ogata, H. (2023). Enhancing Automated Scoring of Math Self-Explanation Quality Using LLM-Generated Datasets: A Semi-Supervised Approach. Computers, 12(11), 217.