Volume 9, Issue 8, August – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24AUG1297

Ethical Considerations for Companies Implementing LLMs in Education Software

Mekam Kontche Steve
IUT (university institute of technology) of Douala

Technology will never Replace Great Teachers, but Technology in the Hands of Great Teachers is Transformational
— George Couros

Abstract:- Large Language Models (LLMs) have revolutionized natural language processing, offering significant advancements in educational software through applications like personalized learning and virtual tutoring. This position paper investigates the ethical considerations for companies integrating LLMs into educational tools. Key issues include data privacy, with a focus on safeguarding sensitive student information against breaches while ensuring transparency and consent. The paper highlights the risk of misinformation, as LLMs might generate incorrect or misleading content that could affect students' learning. It also addresses concerns about algorithmic bias, which can lead to unfair treatment of students from diverse backgrounds, and the potential over-reliance on AI, which may undermine critical thinking and human oversight. Additionally, the paper explores the challenge of equitable access to LLM-based technologies, particularly in underserved communities. The analysis concludes with practical recommendations for companies, including robust data protection measures, balanced AI integration with human oversight, and strategies to enhance access for all students. By emphasizing these ethical challenges, the paper aims to guide responsible AI implementation in education, ensuring that technological advancements benefit all learners fairly and effectively.

Keywords:- *Large Language Models — Artificial Intelligence — Education — Educational Technologies — Responsible AI.*

I. INTRODUCTION

LLMs an acronym for large language models are models that have been trained on a vast corpora of text data, enabling them to predict the next word in a sentence and generate coherent, contextually relevant text across a wide variety of tasks [1].The rapid advancement of Large Language Models (LLMs) in AI has ushered in transformative changes across various sectors, The GPT (Generative Pre-trained Transformer) [1] model developed by OpenAI was the first large language model that was publicly released in 2018 , since then new startups have begun sprouting like mushrooms, introducing innovative ideas across various domains, particularly in education with functionality like personalized learning, automates administrative tasks, virtual tutors,and various AI-driven tools, However, alongside these promising innovations come significant ethical challenges that companies must carefully address to ensure responsible deployment an Ai integration in their offerings for educational settings.it therefore becomes increasingly important to critically examine the ethical considerations of using LLMs in education software. The purpose of this paper is to discuss the potential impacts of this new generation of software and to explore the ethical challenges that arise from integrating AI into the education sector.

➢ *In light of these Challenges, we Address the Questions:*

- How can companies ethically implement LLMs in education software while mitigating risks and ensuring equitable access?
- What safeguards can companies implement to prevent over-reliance on AI in education, ensuring human oversight and critical thinking?
- What professional development opportunities should be provided to educators to help them effectively incorporate LLMs software into their teaching practices?
- How should educational leaders assess the effectiveness of LLMs software in the classroom, including their impact on student achievement and engagement?
- To what extent should LLMs be integrated in school activities and curriculum?

By critically analyzing the ethical implications of LLMs in education, this study aims to contribute to the ongoing conversation on responsible AI deployment.

II. OVERVIEW

➢ *Brief History and Development of LL and LLMs*

The research of LM has received extensive attention in the literature, which can be divided into four major development stages:

- **Statistical Language Models (SLMs)** in the 1990s. These models were based on statistical learning methods and utilized the Markov assumption to predict the next word in a sentence using recent context. Known as n-gram models, such as bigrams and trigrams, SLMs were widely used in tasks like information retrieval and natural

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24AUG1297

language processing. However, they suffered from the curse of dimensionality due to the vast number of transition probabilities required for higher-order models, leading to data sparsity issues. To address this, techniques like back-off estimation and Good-Turing estimation were introduced to smooth data and improve accuracy.

- **Neural Language Models (NLMs)**, which leveraged neural networks such as multilayer perceptrons (MLPs) and recurrent neural networks (RNNs) to model word probabilities. A key advancement was the introduction of **distributed word representations**, which represented words as vectors based on aggregated contextual features. This innovation led to powerful new techniques like **word2vec**, which significantly improved performance across NLP tasks by embedding words into a shared vector space. This shift marked a broader application of language models in representation learning.

- **Pre-trained Language Models (PLMs)** emerged, beginning with models like **ELMo**, which pre-trained bidirectional LSTM networks to capture context-aware word representations. Later, **BERT** introduced the use of the **Transformer architecture** with self-attention mechanisms, setting a new standard for NLP tasks. This "pre-training and fine-tuning" paradigm became dominant, with models like **GPT-2** and **BART** refining the process through different architectures and pre-training strategies.

- **Large Language Models (LLMs)** took pre-trained models to an unprecedented scale. By significantly increasing model sizes—such as GPT-3's 175 billion parameters—LLMs demonstrated emergent abilities, surpassing smaller models in handling complex tasks. LLMs like **GPT-3** showcased capabilities like few-shot learning through in-context adaptation. The release of **ChatGPT**, based on the GPT series, further demonstrated the conversational abilities of LLMs, spurring a sharp rise in research interest and applications in diverse fields.

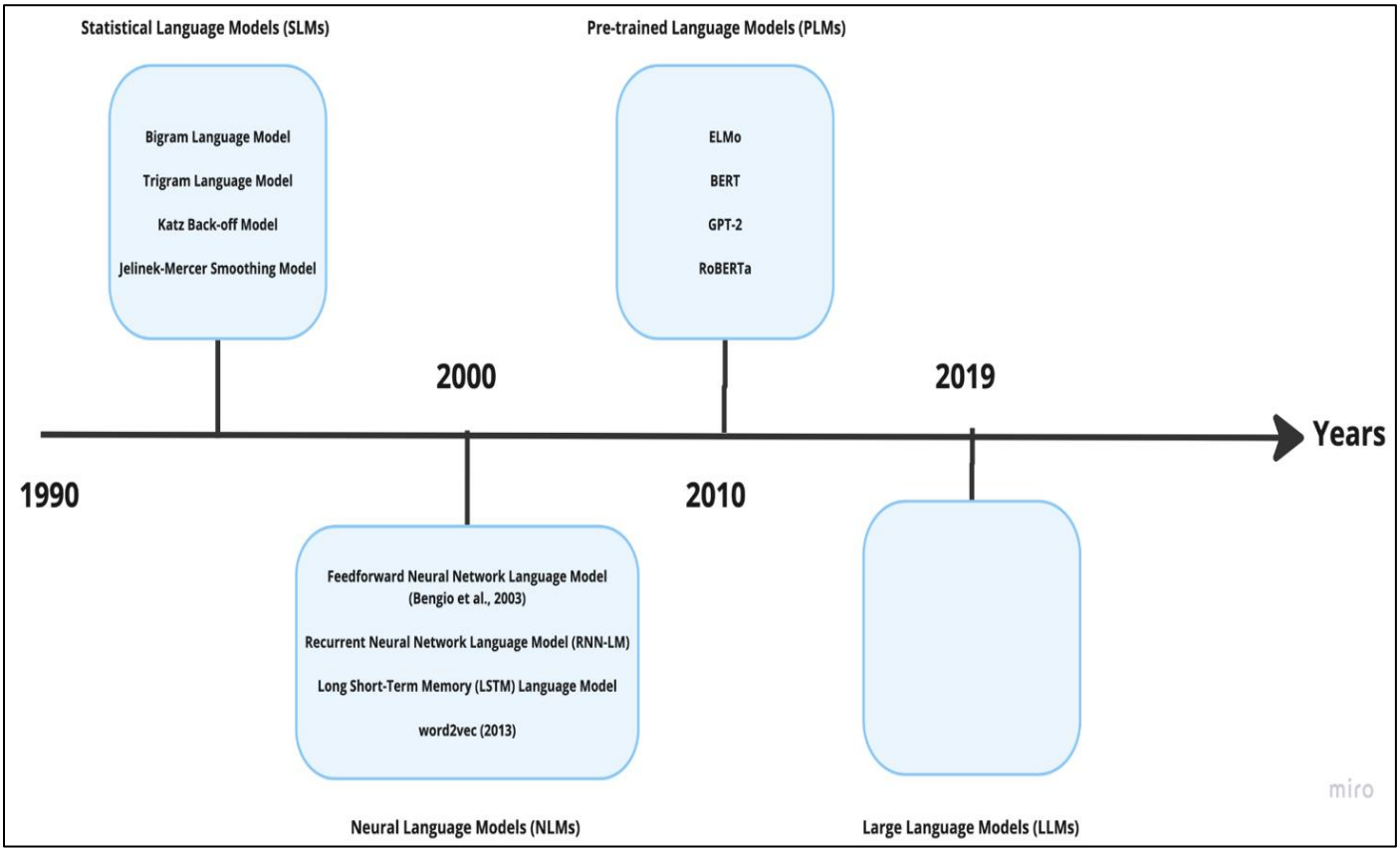

Fig 1 Chronological Display of Language Model Evolution with Examples

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24AUG1297

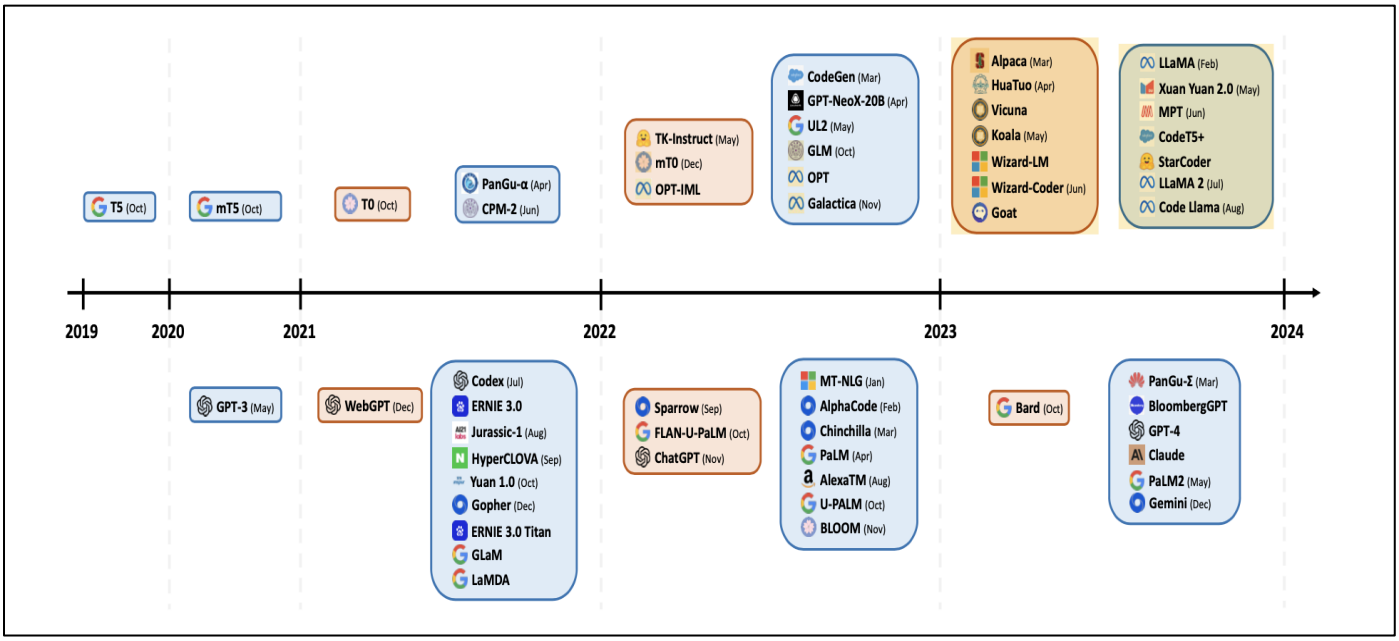

Fig 2 Chronological display of LLM releases: blue cards represent 'pre-trained' models, while orange cards correspond to 'instruction-tuned' models. Models on the upper half signify open-source availability, whereas those on the bottom half are closed-source. The chart illustrates the increasing trend towards instruction- tuned models and open-source models, highlighting the evolving landscape and trends in natural language processing research.[2]

➢ *Current Applications in Education*

Large language models (LLMs) based software have already show many successful application in real life :

- For elementary school students, several LLM-based software applications have already shown many successful implementations, assisting in the development of reading and writing skills. For example, allowing children to read aloud sentences or write with a pen device, the software will then recognize the characters and sentences written or spoken by the children, and evaluate letter formation, correctness, and pronunciation while also providing feedback on pronunciation we can take for example mobile software like Speech Blub by Blub Blub inc or Writing Wizard:Handwriting by L'escapadou. Recent studies of Meridith Lovell and Linda Phillips [3] have even shown successful results after evaluation for several authorized commercial software for teaching reading and writing in the primary grades.

- For middle and high school students, several LLM-based software applications have been implemented in various ways to support the learning process. for example generating creation content like quizzes,flashcards,course resume and even multilingual translation for example ImaginAi by ImaginAi. in a recent work of Dijkstra et al. [4], researchers have used GPT-3 to generate multiple-choice questions and answers for a reading comprehension task and argue that automated generation of quizzes not only reduces the burden of manual quiz design for educators but, above all, provides a helpful tool for students to train and test their knowledge any time while learning from textbooks and other resources.

- For university students, AI-powered software applications have been integrated into the academic experience to assist in more advanced reading, writing, and research tasks. Tools such as Grammarly and ProWritingAid use AI to enhance writing by providing grammar and style suggestions, while Zotero and Mendeley aid in organizing research materials and generating citations. Additionally, Turnitin utilizes AI to detect plagiarism, ensuring academic integrity. AI-driven platforms like QuillBot and ChatGPT further assist students in summarizing texts, generating ideas, and improving the clarity of their writing. These tools not only streamline the writing process but also offer personalized feedback that can help students refine their academic writing and research skills.Stefan A. D. Popenici1* and Sharon Kerr [4] in their recent study highlights the potential for AI to fundamentally change governance and the internal architecture of higher education institution

- For Educators, AI-powered software applications have been designed to enhance instructional practices and support classroom management. Tools like Kahoot! and Socrative use AI to create interactive quizzes and assessments, providing instant feedback and tracking student progress. Edpuzzle allows teachers to integrate questions and interactive elements into video lessons, fostering engagement and understanding. Classcraft uses AI to gamify classroom management, helping to motivate and monitor student behavior. Additionally, TeacherKit offers AI-driven analytics to manage student records, track performance, and organize lesson plans. These applications streamline administrative tasks, enhance teaching strategies, and provide valuable insights to support student learning

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24AUG1297

➢ *Key Challenges and Risks Related to the Application of Large Language Models in Education*

The integration of Large Language Models (LLMs) in education introduces numerous opportunities, but also significant challenges and risks that companies must address to ensure ethical implementation.

- From a company's perspective, **Data Privacy** is a paramount concern. To deliver personalized learning experiences, LLMs based Software require access to extensive datasets that often include sensitive student information, such as academic performance, behavioral data, and personal identifiers.Moreover, storing such large volumes of data makes companies a prime target for cyberattacks, heightening the risk of data breaches. Even with advanced encryption and security protocols, the sheer volume of sensitive information processed demands constant vigilance. The challenge lies in balancing the need for personalization with the obligation to minimize data collection and ensure its secure handling.

- Another critical issue is the risk of **Over-Reliance on AI**. As companies design increasingly sophisticated AI tools for education, there is a tendency for both students and educators to depend too heavily on AI-generated solutions. This over-reliance can diminish the role of critical thinking and human oversight in the educational process, which is problematic from an ethical standpoint. For instance, if teachers overly depend on AI for grading, feedback, or even instructional content, they may neglect their role as facilitators of deeper understanding and critical analysis. Similarly, students might turn to AI-generated answers instead of developing independent problem-solving skills. Companies face the challenge of designing software that enhances learning without replacing the human element, ensuring that AI serves as a supportive tool rather than a substitute for traditional teaching methods.

- **Bias and Fairness** : LLMs are trained on large datasets that may contain inherent biases related to race, gender, socioeconomic status, and cultural norms. These biases can manifest in the software's outputs, potentially leading to unfair treatment or misrepresentation of certain student groups. For instance, students from non-dominant linguistic or cultural backgrounds may receive responses that do not reflect their experiences or realities, contributing to a learning environment that is unintentionally exclusionary. The persistence of such biases in educational software poses serious ethical concerns, as it can perpetuate inequality and hinder the goal of providing an inclusive and equitable education for all students.

- **Misinformation and Content Accuracy:**The ability of large language models to generate human-like text can make it difficult for students to distinguish between real knowledge and unverified information. This can lead to students accepting false or misleading information as true, without questioning its validity[6 ,7].This risk is particularly dangerous in educational settings where students rely on accurate information to build their knowledge. A model might provide inaccurate explanations or present information in a way that is incomplete or biased, leading to confusion and misinformation.

- A further risk lies in **Equitable Access** to LLM-based educational tools. While these technologies hold great promise, not all students have the same level of access to the necessary resources, such as reliable internet connections or modern devices. This digital divide is particularly pronounced in underserved communities, where students may struggle to engage with AI-driven learning tools due to financial constraints or inadequate infrastructure. As LLM-based software becomes more integrated into education, there is a real concern that these disparities will widen the achievement gap, leaving behind students who lack access to the technology needed to fully participate in AI-enhanced learning environments.

III. PRACTICAL RECOMMENDATIONS

The following recommendations provide a comprehensive guide for companies seeking to navigate these complexities and implement LLM-based educational software responsibly.

- Development and implementation of robust data privacy and security policies that clearly outline the collection, storage, and use of student data in compliance with regulation (GDPR, HIPAA, FERPA) and ethical standards. Companies should also be upfront and transparent towards students and their families about the data collection, storage, and use practices, with obligatory consent before data collection and use. Educating the school staff and students about the data privacy and security policies, regulations, ethical concerns and best practices to handle and report related risks can also improve Ai-Software adoption.

- Incorporate safeguards to prevent over-reliance on AI by integrating human oversight features like manual overrides and teacher controls. Companies should offer training to educators on AI's limitations and encourage critical thinking, ensuring that AI supports rather than dominates the learning process.

- Develop solutions that cater to a wide range of technological environments, including offline capabilities, tiered pricing models, and multilingual support. Companies should aim to ensure that AI-driven tools are accessible to underserved communities and that the digital divide is narrowed by making education tools inclusive and cost-effective.For example, they could reduce the price of their offerings and provide their software in underrepresented languages.

We therefore conclude with UNESCO's call to ensure that AI does not widen the technological and educational divides within and between countries, and recommended important strategies for the use of AI in a responsible and fair

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24AUG1297

way to reduce this existing gap instead. According to the UNESCO education 2030 Agenda [8]: *"UNESCO's mandate calls inherently for a human-centered approach to AI. It aims to shift the conversation to include AI's role in addressing current inequalities regarding access to knowledge, research and the diversity of cultural expressions and to ensure AI does not widen the technological divides within and between countries.The promise of "AI for all" must be that everyone can take advantage of the technological revolution under way and access its fruits, notably in terms of innovation and knowledge."*

- Provide educators with foundational AI literacy training, focusing on understanding how LLMs work, their limitations, and their potential applications in education. This training should include practical sessions on integrating AI tools into lesson planning, assessment, and personalized instruction.Create ongoing professional development programs that encourage peer collaboration, where educators can share experiences, challenges, and best practices related to using LLMs Software in their teaching. These networks can foster continuous learning and adaptation of AI tools in diverse educational contexts.

- Companies' software should be used as supplementary tools to enhance personalized learning. They can provide tailored feedback, offer additional practice materials, and support differentiated instruction based on individual student needs. However, they should complement rather than replace traditional teaching methods and human interaction.for example administrative tasks such as grading and generating lesson plans, allowing educators to focus more on direct student engagement. Their integration in these areas should aim to reduce the administrative burden while maintaining the quality of education.It's important that students are guided to critically evaluate AI-generated content and not solely rely on it for their learning.

- Implement standardized assessments and benchmark tests to evaluate the impact of LLM software on student learning outcomes. Compare pre- and post-intervention performance to gauge improvements in areas such as literacy, critical thinking, and problem-solving skills. Analyze data to determine if there are significant gains in student achievement attributable to the use of Ai driven tools.Additionally, monitor changes in student engagement and participation using surveys, interviews, and classroom observations, utilize engagement analytics from the software, if available, to track student interactions and time spent on tasks.Collect qualitative feedback from teachers who integrate LLM software into their instruction through surveys, focus groups, and interviews. This feedback will help understand their experiences, challenges, and perceptions of how the software affects classroom dynamics and supports their teaching practices. Insights from this feedback will provide a comprehensive view of the software's practical effectiveness and highlight areas for potential improvement.

IV. CONCLUSION

The integration of Large Language Models (LLMs) into educational software holds transformative potential for enhancing teaching and learning experiences. However, this potential comes with significant ethical considerations that companies must navigate to ensure responsible and equitable implementation. This paper has explored key challenges including data privacy, misinformation, algorithmic bias, over-reliance on AI, and equitable access.

To address these challenges, it is essential for companies to implement robust data privacy policies, ensuring transparency and securing consent while safeguarding sensitive student information. Additionally, mitigating over-reliance on AI involves incorporating human oversight and promoting critical thinking, ensuring that AI tools complement rather than replace traditional educational methods. Professional development opportunities for educators should focus on AI literacy, ethical use, and collaborative learning networks to support the effective integration of LLMs into teaching practices.

Educational leaders must assess the effectiveness of LLMs through standardized assessments, engagement analytics, and qualitative feedback from teachers to understand the impact on student achievement and classroom dynamics. Finally, the extent of LLM integration in curricula should be carefully balanced to enhance learning without compromising the role of human interaction and traditional educational approaches.

By addressing these ethical concerns with thoughtful and responsible practices, companies can harness the benefits of LLMs while contributing to an equitable and effective educational environment. This balanced approach will ensure that advancements in AI serve to enhance learning opportunities for all students while upholding the principles of fairness and integrity in education.

REFERENCES

[1]. Language Models are Few-Shot Learners by Ilya Sutskever et al.
[2]. A Comprehensive Overview of Large Language Models. Humza Naveeda, Asad Ullah Khana,∗, Shi Qiub,∗ et al
[3]. Commercial Software Programs Approved for Teaching Reading and Writing in Primary Grades: Another Sobering Reality. Meridith Lovell and Linda Phillips
[4]. Comprehension Quiz Generation using Generative Pre-trained Transformers. Ramon Dijkstra, Zu ̈lku ̈f Genc ̧, Subhradeep Kayal, and Jaap Kamps. Reading .
[5]. Exploring the impact of artificial intelligence on teaching and learning in higher education. Stefan A. D. Popenici and Sharon Kerr

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24AUG1297

[6]. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. bioRxiv, 2022. Catherine A. Gao, Frederick M. Howard, Nikolay S. Markov, Emma C. Dyer, Siddhi Ramesh, Yuan Luo, and Alexander T. Pearson.
[7]. Chatting and Cheating. Ensuring academic integrity in the era of ChatGPT. EdArXiv, 2023. Debby R.E. Cotton, Peter A. Cotton, and J.Reuben Ship- way.
[8]. UNESCO. Education 2030 Agenda. https:// www.unesco.org/en/digital-education/ artificial-intelligence, 2023. Accessed: 22.01.2023.