Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #47381

Experiences with using an LLM-based Chatbot for a Multicultural Engineering Program Orientation (Experience)

Haya Alshayji, Pennsylvania State University

Haya Alshayji is advancing her academic journey as a Ph.D. candidate at Pennsylvania State University's Industrial and Manufacturing Engineering, focusing on integrating machine learning, data science, and statistical modeling in her studies. Recently, she expanded her research experience by joining THiCC Lab as a research assistant. Haya's involvement in these diverse research areas highlights her commitment to leveraging Al for social good and advancing the field of data science.

Deja Workman, Pennsylvania State University
Swapnika Dulam, Pennsylvania State University
Dr. Lauren A Griggs, The Pennsylvania State University

Dr. Lauren Griggs received her B.S. in Engineering Science from The University of Virginia. She received her Doctor of Philosophy in Biomedical Engineering from Virginia Commonwealth University (VCU), where she worked in the field of Cell and Matrix Biomechanics. She completed her Postdoctoral training in the Department of Biomedical Engineering at Vanderbilt University, where she sought to elucidate the mechanical linkage between breast cancer and diabetes. Dr. Griggs joined Penn State in the summer of 2019 as an Assistant Teaching Professor, Director of the Multicultural Engineering Program and Director of the Clark Scholars Program. In this role, Dr. Griggs drives initiatives to improve the recruitment and retention of historically underrepresented groups pursuing degrees in engineering and strives to foster a welcoming environment that celebrates culture and inclusion. She is the advisor for the National Society of Black Engineers and co-advisor for the Society of Hispanic Professional Engineers, and the Multicultural Engineering Graduate Association. Her passion lies in mentoring through meaningful career discussions and helping students gain confidence as well as succeed in their chosen degree fields.

Dixon Zor, Pennsylvania State University
Christopher L Dancy, The Pennsylvania State University, University Park

Dr. Chrisopher L. Dancy is an associate professor of Industrial and Manufacturing Engineering, Computer Science and Engineering, and African American Studies at Penn State where he also currently holds the Harold and Inge Marcus Industrial and Manufacturing Career Development professorship. He directs The Human in Computing and Cognition (THiCC) lab, is currently a faculty partner in the Center for Black Digital Research/#DigBlk, and is the incoming co-director of the Center for Black Digital Research/#DigBlk starting summer 2025. He is also an affiliate faculty in the Institute for Computational and Data Sciences, Center for Socially Responsible AI, and Center for Brain, Behavior, and Cognition. His expertise and interests lie at the intersection of human behavior, computational systems, and social structures with a goal of studying the Human in various computing systems and engineering processes. Dr. Dancy is an NSF CAREER award recipient, with a goal of moving AI and computing towards the development of socioculturally competent AI systems, and is co-lead of the Computational Representation of Human Decision Processes research thrust for the NSF funded AI Institute for Societal Decision-Making where he is using computational cognitive systems to study and intervene in the real impacts of racialization and anti-Blackness on decision making processes and outcomes.

Experiences with using an LLM-based Chatbot for a Multicultural Engineering Program Orientation (Experience)

Abstract

Given the pace with which AI systems are being developed and used, there is a growing need for more guidance around the ethical use of AI. Due to the prominence of artificially intelligent systems, future engineers need to be able to analyze the available AI models and make responsible choices critically. In the Fall of 2024, The Human in Computing and Cognition (THiCC) Lab collaborated with the Multicultural Engineering Program Orientation (MEPO) at Penn State to teach incoming engineering students about the responsible use of AI systems with the help of an interactive Large Language Model (LLM) based chatbot. The MEPO is a four-day program designed to welcome incoming first-year undergraduate engineering students, primarily from racially and ethnically minoritized groups, by fostering connections with upper-division student mentors, academic success resources, and professionals in the field while also exposing them to typical elements of the engineering curriculum such as teamwork and innovation. One exciting component of MEPO is an engineering design competition, where students are asked to design prototypes to solve their assigned problem, for which this year's theme was "decades." Students were assigned to one of 14 groups; the groups were then assigned to one of 4 decades (1910s, 1930s, 1970s, 1980s). Each decade had an accompanying disaster that the students would be responsible for helping to resolve–the students assigned to the 1910s, for example, were tasked with designing a context-appropriate technological solution to help mitigate the Spanish Flu. The final objective was to create a prototype and a presentation regarding their findings and solutions to their assigned problem. The chatbot was meant to aid specifically the students as they brainstormed different ideas and solutions, allowing them to think critically about these intelligent systems as they used the chatbot. Before the four-day MEPO event, our team at the THiCC lab spent some time building the chatbot for the students to use. For the chatbot, we chose LLaMA-2 because of its reliable text generation and open-source nature, which includes transparency about the data sources used to train the system. We focused on transparency, as we wanted to highlight the importance of data sources and the significance of community-engaged open-source development. Additionally, we integrated Retrieval-Augmented Generation (RAG) to allow the chatbot to pull specific information, like historical data and disaster scenarios, from a custom pamphlet prepared by the MEPO team. This ensured the chatbot gave factually correct answers tied directly to the decades they worked on, which was later hosted on Huggingface spaces. On the first day of the MEPO, the THiCC lab team directed a lesson to introduce the students to the chatbot and its utility. The first half of the lesson was spent educating the students on the dangers and potential considerations of using LLMs and AI. The second half was spent showing the students certain variations in the usage of chatbots and the differences in the answers they provide. The variations included using a chatbot with pre-trained data (vanilla version), using the RAG version to retrieve factually correct answers from

the pamphlet, and adding context to the RAG version to retrieve more nuanced answers. After the lesson, the students could use the various versions of the chatbot to help them in their design challenge and understand the difference in responses while using it for a given problem statement. During the four days, students presented a range of questions and feedback, from technical questions on how to access the chatbot to questions about motives and why they needed to use the chatbot. On the final day of the competition, students presented their designs and were able to thoughtfully consider the chatbot as an imperfect yet valuable tool in their competition. This report utilizes a structured survey and mixed-methods analysis to evaluate the educational impact of the chatbot and related activities on students' comprehension of AI ethics and their overall learning experience.

Keywords

Ethical AI Education, Multicultural Engineering, Large Language Models (LLMs), Critical Thinking in AI, Engineering Diversity Programs

1. Introduction

Artificial Intelligence (AI) and its role in our society are developing rapidly, making engineering education surrounding AI a crucial topic. Engineers who work with AI in the future will need the ability to think critically about AI-based systems and large language models. We encouraged these future engineers to question what artificial intelligence means, its limits, and what information is input into AI systems to produce outputs. Many large corporations intentionally keep their AI systems opaque for proprietary reasons, which leads to a lack of transparency in various AI technologies. People should consider whether they want to support that technology in these cases. In our experience with the Multicultural Engineering Program Orientation (MEPO) at Penn State, a program dedicated to helping incoming first-year undergraduate engineering students, we explored the usage of Large Language Models (LLMs) towards a critical understanding of AI systems. During the most recent iteration of the MEPO, we created an LLM-based chatbot to help engineering students with a design competition. Throughout this experience, the goal was to encourage the students to think critically about AI and help them try to put some of our teachings into practice.

1.1. Background on AI in Education

Artificial Intelligence has become a transformative force in education, especially in Science, Technology, Engineering, and Mathematics (STEM), offering tools like LLMs and chatbots that enhance teaching and learning. These technologies enable personalized learning experiences, real-time feedback, and interactive engagement, potentially fostering critical thinking and inquiry-driven approaches. For example, chatbots can guide students through structured learning paths, making STEM concepts more accessible and engaging [1,2]. Nonetheless, using AI in educational settings does not guarantee that students will be critically aware of the AI systems they interact with.

Critical AI education focuses on equipping students with the skills to evaluate these widely used AI systems critically, emphasizing digital literacy, ethical reasoning, and collaboration to use them better. As AI systems increasingly shape societal structures, integrating these educational objectives has become vital to preparing students for technology-driven futures. AI for Education (AI4EDU) is a multi-disciplinary field that uses state-of-the-art AI technologies, especially LLMs, to improve educational practices. These technologies facilitate data mining and the development of various applications that can support personalized learning experiences[1,3].

However, integrating AI into education has challenges, such as preprocessing bias, accessibility, and data privacy. Practical strategies to address these include providing professional development for educators, developing inclusive and adaptable curricula, and fostering a culture of continuous feedback [4,5]. By navigating these challenges, AI technologies can make education supportive, equitable, and tailored to diverse student needs.

1.2. MEPO and Competition Design Overview

The Multicultural Engineering Program Orientation at Penn State is an annual orientation hosted by the Center for Engineering Outreach and Inclusion. The program fosters a welcoming environment for all incoming first-year engineering student participants by offering various forms of support throughout the orientation, highlighting the unique perspectives of all student mentors and professors engaged in the program, and celebrating the value of diversity and inclusion in innovations within the engineering field. This orientation lasts for 4 days, during which students receive numerous resources and opportunities, including mentorship, networking prospects, professional development, and, as the focus of this paper, the chance to participate in an engineering design competition. In the engineering design competition, the students receive instructions to create a design that addresses a specific problem. For 2024, the theme was "decades." The instructor assigned students to one of fourteen groups, which focused on a different decade (1910s, 1930s, 1970s, 1980s) from which to derive a problem. For example, students assigned to the 1910s developed a technological plan to address issues stemming from the 1918 influenza pandemic. Students assigned to the 1930s actively helped resolve the Great Depression. Students in the 1970s worked to resolve the energy crisis, while those in the 1980s helped address the Exxon Valdez oil spill. While designing with these prompts, MEPO students were also explicitly instructed to consider social contexts (e.g., the ways access to resources such as transportation may not be equitable or trust amongst populations) of their assigned decade. During the design challenge, the students had access to mentors whom MEPO trained. The mentors assisted the students with their designs, answered questions related to the designs without necessarily providing solutions, and encouraged the design teams without necessarily giving them answers to their assigned problems. Students were also given some background on social contexts through MEPO. They were told to remember the decade and what people likely had access to or felt comfortable with; they were also told to consider how the chatbot's variations can improve the design process and collaboration for engineering students and professionals. At the end of the design competition, each group presented their prototype and thought processes for solving their assigned problem. Throughout the four days the students had to work on this prototype, they had access to the LLM-based chatbot designed by The Human in Computing and Cognition (THiCC) Lab. Students were encouraged to use the chatbot and its variations to gather the information required to design the prototype by incorporating social contexts and gaining an understanding of similar AI systems.

1.3. Objectives

This experience report aims to showcase how AI tools like chatbots can be integrated into a multicultural engineering program as pedagogical instruments. These tools foster critical engagement while enhancing students' understanding of AI ethics and its limitations. This report utilizes a structured survey and mixed-methods analysis to evaluate the educational impact of the chatbot and related activities on students' comprehension of AI ethics and their overall learning experience.

2. Related Work

2.1. Ethical AI Education

Garrett, Beard, and Fiesler [6] discuss two ways of incorporating ethics in AI education: hosting courses centered around AI ethics and AI ethics into pre-existing classes. We want to hone in on what they discussed regarding the implications of including ethics education in otherwise "technical" coursework. The authors specifically discuss educating computer science students, but many of these findings apply widely to other forms of engineering. The authors conducted systematic syllabi review of courses to determine how educators were teaching ethics in their AI courses; they found that the most common topics were avoiding bias, promoting fairness, and protecting privacy. However, they also found that these topics tended to come last in the syllabus, and there were times when they were marked only to be covered "if time allows." The authors argue that AI ethics should not be a backseat topic and should only be covered if convenient. Still, the primary inclusion of AI coursework must be taken seriously to convey its importance to students.

Borenstein and Howard additionally discuss the need for additional AI ethics education [7]. The authors argue that AI ethics has not truly been embedded in AI education yet, and some of the curriculum surrounding AI ethics needs to be rethought. They propose three elements to help students become aware of AI ethics in their education. First, they propose teaching students about participatory design and the "ethical design of AI algorithms." Second, they propose inviting lessons around ethical data acquisition. Third, they suggest offering these ethics-related lessons in different contexts across different courses.

2.2. LLMs in Educational Settings

Large Language Models are increasingly used in education, presenting benefits and challenges. These technologies can assist in automating tasks such as developing educational content, tailoring learning experiences to individual needs, and facilitating assessments [8,9]. They offer opportunities to enhance student engagement, support the creation of teaching materials, and deliver personalized educational approaches [10].

However, integrating LLMs into education presents several practical and ethical challenges, such as insufficient technological readiness, a lack of transparency, and privacy concerns [8]. To overcome these obstacles, researchers suggest updating existing innovations with advanced models, supporting open-sourcing initiatives, and taking a human-centered approach to development [4]. Furthermore, educators and students must cultivate new skills to understand and critically assess outputs generated by LLMs [10].

Despite these challenges, LLMs can transform educational practices and create more effective personalized learning environments [9]. However, their integration also raises concerns about equity and potential bias, particularly when considering the need for diversity and inclusion in engineering education.

2.3. Diversity in Engineering Education

Diversity and inclusion (D&I) in engineering education are essential for increasing the representation of underrepresented groups, especially women and minorities [11,12]. Despite ongoing initiatives, the demographics of the engineering field remain primarily unchanged and need to reflect societal diversity [12]. The rationale for promoting diversity varies, encompassing industry needs, social justice arguments, and the benefits of cognitive diversity [13].

D&I efforts include a variety of approaches, such as scholarship programs, extracurricular activities, and enhancements to the curriculum [13]. However, definitions and priorities regarding diversity differ across institutions, with some focusing specifically on women, ethnic minorities, or low socioeconomic status (SES) students [13]. Global collaboration on D&I initiatives could help share best practices and maximize learning opportunities [11]. Additionally, understanding the individual pathways to engineering and adopting a systemic perspective may offer insights for improving recruitment and retention efforts [12].

3. Methodology

Our methodology for integrating LLM-based chatbot into the Multicultural Engineering Program Orientation was organized into three key steps: Chatbot development, educational intervention, and assessment design and survey development. Each step was critical in ensuring the successful application of the chatbot as both a technical resource and an educational tool aimed at fostering critical thinking and ethical awareness in engineering students. Below the framework was outlined for the methodology and illustrated in Figure 1:

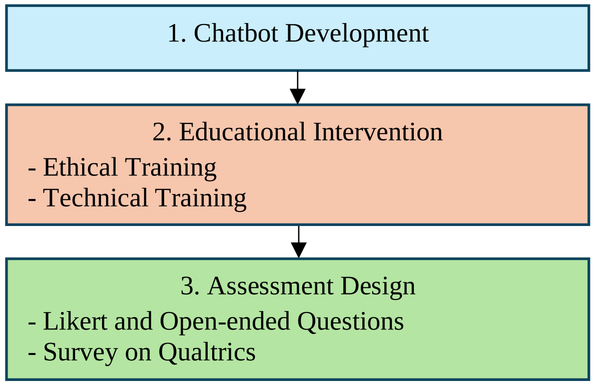

Figure 1. Framework Diagram for Study Methodology

3.1. Chatbot Development

Large Language Models are one kind of generative AI systems that typically use a series of neural networks (with a particular transformer architecture). These models are trained on extensive datasets with billions of parameters to learn the patterns and structures of human language and output a sequence of text or other media based on the input text (query/prompt). These models do not understand the input text or use human-like cognitive processing of information; instead, they complete (text) sequences with coherent and contextually relevant responses. This capability makes LLMs invaluable tools in fields ranging from customer support chatbots to creative writing.

We chose LLaMA2 [14] to build the chatbot because of the accessibility of the source code and the relative transparency of the model and its training data. We specifically chose the 7 billion parameters for the chat model from Hugging Face [15] (Llama-2-7b-chat-hf) for this. The participants were supposed to use this chatbot hosted on Hugging Face space to assist them in their design competition. Participants were expected to analyze AI-generated responses, identify biases, and explore practical applications of LLaMA2 in engineering design processes and user experiences. Ultimately, by asking the participants questions to reflect upon, we encouraged reflective learning to consider technical and ethical challenges.

3.2. Educational Intervention

3.2.1. Ethical Training

In the beginning stages of the design competition, a select few participants were chosen to be the "AI specialist" for their group. The AI specialists attended a lecture led by THiCC Lab, where they were specially equipped with additional training. This training was made up of ethical training and technical training. The ethical training covered various topics, beginning with AI, the biocentric Man, and anti-Blackness.

The biocentric Man is a term coined by Sylvia Wynter, and it is the idea of one Man, though not necessarily any one person, but rather one genre of person [16]. The term points to a pseudoperson that is White, cis-gendered, heterosexual, and male. We define this term because we and others argue that modern society is built around the biocentric Man, and this extends to technology and Artificial Intelligence. While the biocentric Man is placed on a pedestal, those without those characteristics are othered. Consequently, one may argue that AI is inherently sociocultural because it is created in and thus rooted in a socialized and racialized environment while persistently serving the biocentric Man [17,18].

The second topic we covered in our lesson was AI and the physical environment. Technology, including AI, affects the environment via energy usage, water consumption, and carbon dioxide emissions [19]. While these environmental challenges affect us globally, these universal effects are not felt evenly; these environmental struggles impact different groups of people unequally. Research has shown that racial minorities have higher health risks associated with air pollution [19]. Additionally, the data centers that consume water can compete with drinking water sources and electricity, causing the residents of places with data centers present or meant to be built to protest due to the draining of these resources. We taught the students about how pollution disproportionately affects marginalized communities.

Next, we covered research that describes large language models (LLMs) as stochastic parrots encoding the internet [20]. The LLMs behind various forms of AI use data from multiple sources, some of which are often scraped from the internet. However, the internet is disproportionately representative of White men as well, especially in certain corners of the internet such as Reddit. Not only is data sourced from the internet susceptible to being ill-representative of people who have been othered, but also this data gives AI the ability to mimic language in a deceptively passable way [20]. While the supposed measure of success and maturity of large language models is this ability to replicate human language, this goal is simultaneously dangerous because it creates what is known as stochastic parrots—models that push interpretable, passable responses without understanding the content of what they put out.

We then covered ethical AI practices and action items that the students should take when working with AI-based systems. This multi-faceted training area began with the understanding that data collection is typically extractive. Understanding that taking information from a population is an extractive process leads us to question how we can exchange rather than extract it and taking the appropriate actions to make this happen is an important part of building just AI. For data collection to be more of an exchange, we must consider who we are building for and who benefits from our actions. We taught the students to consider designing for those impacted by the technology concerning design justice perspectives [21]. In addition to understanding design justice, we taught about datasheets for datasets [22]. Datasheets accompany a dataset to keep the dataset acquainted with its roots and to potentially have the dataset developer document answers to critical social questions tied to the dataset. We taught that data will always be from somewhere and can be dangerous when emancipated from its original intentions or usages.

3.2.2. Technical Training

We guided the students through three different Python notebooks and explained how the difference in implementation would affect the chatbot's responses. We used the same three prompts based on the design competition theme to demonstrate the reaction difference.

The prompts were:

- How could a telemedicine system have been implemented using the technology available during the 1910s to combat the Spanish Flu pandemic?
- What affordable and sustainable housing solutions could have been developed during the Great Depression to address widespread homelessness?
- How might renewable energy technologies like solar or wind power have been developed in the 1970s to address the energy crisis?

For the first notebook, a vanilla version of the LLaMA2 model without any changes was used to give responses solely based on the data it was pre-trained on.

For the second notebook, we modified the notebook to allow system prompting, a technique for providing guided and contextualized responses. This is to enforce the theme of "decades" in the responses. For example, a context such as "You are living in the 1910s" could be added in the first prompt to get a more accurate response.

For the third notebook, we use a pamphlet relevant to the themes as the context and system prompts to get the responses. This notebook employed Retrieval-Augmented Generation (RAG), a model training technique that incorporates external data into the responses generated by the LLaMA2 model. By embedding the pamphlet into the pre-trained model, the chatbot could generate more accurate answers and contextually tied to the students' queries regarding their assigned decades.

This approach also served as a practical demonstration of how Large Language Models (LLMs) can be shaped to provide tailored and sometimes biased outputs, emphasizing the critical role of developers in training and contextualizing these systems. For instance, even with an open-source model like LLaMA2, the curated input data directly influenced the chatbot's responses. By exploring this aspect, the AI Specialists gained more profound insights into the implications of AI training processes.

The finalized RAG-enabled LLM-based chatbot was hosted on the Hugging Face platform so that students could interact with it for their design competition. This hands-on engagement encouraged reflective learning, enabling participants to critically evaluate the utility and limitations of AI-driven tools in addressing real-world problems.

After this exercise, the AI specialists were given the link to the hosted Hugging Face space based on the RAG-enabled notebook for their design competition. The AI specialists were expected to share this link with their teammates to aid them in their design competition and finally reflect on their experience with the AI system to understand its impact.

3.3. Assessment Design and Survey Development

Participants: To assess the impact of integrating an LLM-based chatbot in the Multicultural Engineering Program Orientation (MEPO), a Qualtrics survey was administered at the end of the academic semester. The survey targeted MEPO student participants and staff (mentors and facilitators). A total of 30 out of 142 eligible students completed the survey, yielding a response rate of 21%.

Mixed-methods approach: This study employed a mixed-methods research approach, combining quantitative (Likert-scale) and qualitative (open-ended) questions to evaluate the chatbot's effectiveness, aiming to assess students' understanding of AI ethics, perceived limitations of AI, and chatbot usability and impact on learning. The quantitative component provided measurable insights into student perceptions, while the qualitative responses captured deeper reflections on their challenges, concerns, and engagement levels. Mixed-methods research is particularly valuable in educational settings as it helps validate trends while incorporating nuanced perspectives that numerical data alone cannot provide [23,24].

Survey Structure and Data Collection: Data for this study was collected through an online survey conducted on the Qualtrics platform. The survey was administered after the MEPO program to gather insights into students' experiences and learning outcomes. It included Likert-scale questions to capture quantitative data on students' perceptions of the chatbot and open-ended questions to explore their reflections on its educational value.

The survey was structured as follows: The survey was structured into several sections:

1. Pre-Survey Eligibility Check – To identify participant roles (student vs. staff).
2. Understanding the Chatbot – Focused on usability, navigation, and effectiveness.
3. AI Limitations – Explored whether students recognized the constraints of AI tools.
4. Ethical Implications of AI – Assessed students' awareness of AI-related ethical concerns.
5. Additional Feedback – Open-ended reflections on improvements and overall experience.

Data collection adhered to Institutional Review Board (IRB) guidelines, ensuring ethical compliance and participant confidentiality. All responses were anonymous, participation was voluntary, and students received a $10 incentive upon survey completion.

Linking Data to Learning Outcomes: This study employed a mixed-methods research approach, combining quantitative (Likert-scale) and qualitative (open-ended) questions to evaluate the chatbot's effectiveness, focusing on students' understanding of AI ethics, perceived limitations of AI, and chatbot usability and impact on learning. The quantitative component offered measurable insights into student perceptions, while the qualitative responses provided more profound reflections on their challenges, concerns, and engagement levels. This mixed-methods research is particularly valuable in educational settings as it validates trends and incorporates nuanced perspectives that numerical data alone cannot capture. It has been widely endorsed in educational and behavioral sciences for addressing diverse research questions and generating actionable insights [25].

4. Results and Discussion

The Qualtrics survey included responses from 21% of MEPO participants. The results presented in this paper aim to provide a starting point for understanding the perceptions of MEPO participants and to reflect on their overall experience during the program. 71.5% of the survey participants were students, of which 26.7% engaged as AI specialists. Additionally, 28.5% were staff members, among which 33.3% served as design team members.

<u>Understanding the chatbot:</u>

Survey responses indicated a mixed experience with the chatbot. Figure 2 shows that while some participants appreciated its potential, nearly half (43.8%) strongly disagreed with the statement, "I was able to use the chatbot," highlighting significant usability challenges. Additionally, 37.5% found it complex or confusing, and only 6.3% felt it was particularly useful for their design tasks.

Key usability challenges included navigation difficulties, vague responses, and a disconnect between the chatbot's intended educational purpose and students' expectations. Students struggled with accessing the interface and often found the responses repetitive. Many questioned the necessity of using a chatbot when traditional search engines could serve similar functions.

To improve the experience, it is suggested that an interactive walkthrough be offered before use, that responses be more specific, and that better contextual framing be provided regarding the role of chatbots in AI ethics education. By implementing these changes, future iterations could better meet students' needs and clarify the benefits of engaging with the chatbot.

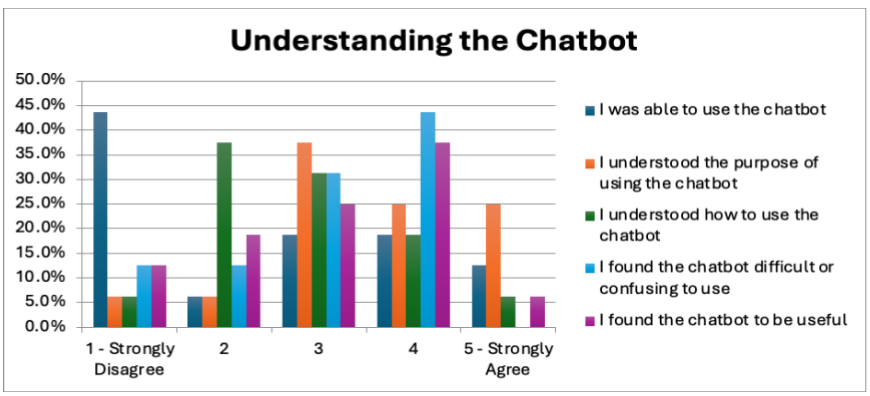

Figure 2. Participants' understanding of the chatbot

<u>Understanding AI limitations:</u>

In Figure 3, the survey revealed that 43.8% moderately agreed that the chatbot fell short and that this helped them see AI limitations. From the distribution of the responses, the chatbot successfully brought attention to the limits of AI, which aligns with its educational objectives regardless of room for improvement.

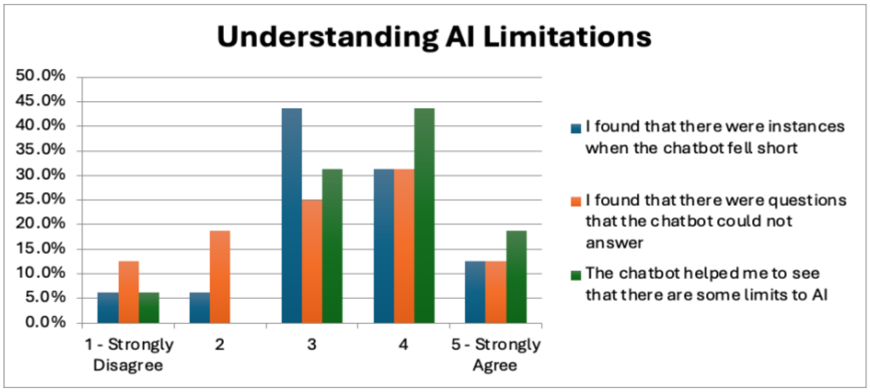

Figure 3. Participants' understanding of AI limitations

Understanding ethical implications:

Figure 4 shows that the survey indicated mixed perceptions about the chatbot's effectiveness in conveying AI's ethical implications and helping participants recognize its ethical drawbacks. These results suggest the chatbot successfully raised ethical awareness, yet targeted improvements in content and interactive scenarios could address the gaps for less impacted participants.

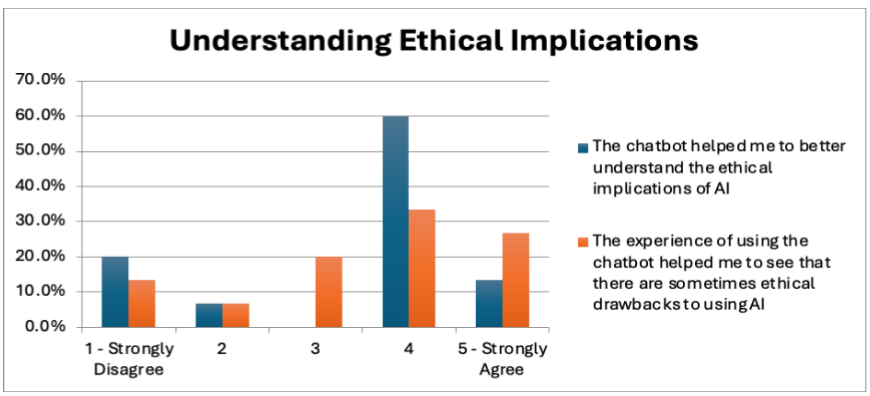

Figure 4. Participants' understanding of ethical implications

Positive and negative keywords were extracted from participant responses, and the percentage of participants cited each was calculated to identify crucial insights. 40% of participants—primarily mentors and students—used words like "useful" and "research aid" to describe the chatbot, while 30% mentioned words like "confusing" and "repetitive." 35% of participants had positive perceptions about "understanding AI" and "adjusted expectations" in conversations regarding AI limitations, while 50% cited drawbacks, such as "not accurate" and "unable to handle specifics." Concerning ethical issues, 30% voiced concerns about "misinformation" and "resource-heavy AI," while 50% emphasized benefits like "unbiased" and "transparency." These observations highlight the usefulness of the chatbot, its drawbacks, and the moral questions it raises.

Reflections and lessons learned:

These reflections highlight the benefits and drawbacks of an AI tool (e.g., chatbot) integrated into a multicultural engineering program as a pedagogical instrument. Participants valued the critical engagement fostered by discussions, such as introducing concepts like the biocentric human genre and how AI disproportionately affects marginalized communities. However, challenges arose due to knowledge gaps, particularly with tools like Jupyter notebooks and the chatbot's contextual understanding limitations. Some participants felt that integrating AI into historical contexts (e.g., the 1930s Great Depression) seemed incongruent, suggesting a need for more precise alignment between themes and AI applications.

Key lessons emphasize the importance of preparation and accessibility, ensuring contextual relevance, expanding ethical discussions, and implementing real-time feedback mechanisms for continuous improvement. Moving forward, program assessment enhancements can help address these challenges. Future iterations of this study could incorporate pre- and post-surveys to measure shifts in students' AI literacy levels, alongside focus groups or structured interviews to gain more profound insights into students' critical engagement with AI. Additionally, a longitudinal study could track how students' attitudes toward AI ethics evolve beyond MEPO, providing a more comprehensive understanding of the long-term impact of AI education in engineering programs. By refining assessment methods and integrating more structured evaluations, future implementations can ensure that students engage critically with AI and develop lasting AI literacy and ethical awareness.

5. Conclusion

Integrating an LLM-based chatbot in the Multicultural Engineering Program Orientation was a successful yet challenging effort. The chatbot's design focused on AI's ability—as a pedagogical and critical tool—to get students to consider its ethical implications and dependability. In contrast, this work also identified areas that demanded improvement, such as addressing participants' different technical readiness and improving alignment with historical themes. Considering these experiences, we intend to enhance subsequent applications and ensure AI-powered resources promote accessible, inclusive, and impactful engineering education.

References

[1] J. Tang, X. Zhou, X. Wan, M. Daley, and Z. Bai, "ML4STEM Professional Development Program: Enriching K-12 STEM Teaching with Machine Learning," *International Journal of Artificial Intelligence in Education*, vol. 33, pp. 185–224, 2022.

[2] Alsafari, B., et al., "Towards effective teaching assistants: From intent-based chatbots to LLM-powered teaching assistants," *Natural Language Processing Journal*, vol. 8, 2024

[3] W. Qingsong, J. Liang, C. Sierra, R. Luckin, R. Tong, Z. Liu, P. Cui, and J. Tang, "AI for Education (AI4EDU): Advancing Personalized Education with LLM and Adaptive Learning," in *Proceedings of the 2024 ACM International Conference on Learning Technologies*, pp. 6743–6744, 2024, Doi: 10.1145/3637528.3671498

[4] A. M. Al-Zahrani and T. M. Alasmari, "Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications," *Humanities and Social Sciences Communications*, vol. 11, 2024

[5] K. Stolpe and J. Hallström, "Artificial intelligence literacy for technology education," *Computers and Education Open*, vol. 6, Article no. 100159, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666557324000016. Doi: 10.1016/j.caeo.2024.100159

[6] N. Garrett, N. Beard, and C. Fiesler, "More Than 'If Time Allows': The Role of Ethics in AI Education," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES' 20)*, New York, NY, USA, 2020, pp. 272–278. Doi: 10.1145/3375627.3375868. [Online]. Available: https://doi.org/10.1145/3375627.3375868

[7] J. Borenstein and A. Howard, "Emerging challenges in AI and the need for AI ethics education," *AI and Ethics*, 2021

[8] J. Borenstein and A. Howard, "Emerging challenges in AI and the need for AI ethics education," *AI Ethics*, vol. 1, pp. 61–65, 2021. doi: 10.1007/s43681-020-00002-7. [Online]. Available: https://doi.org/10.1007/s43681-020-00002-7

[9] S. Wang, T. Xu, H. Li, C. Zhang, J. Liang, J. Tang, P. S. Yu, and Q. Wen, "Large Language Models for Education: A Survey and Outlook," *arXiv* preprint, vol. abs/2403.18105, 2024. [Online]. Available: https://arxiv.org/abs/2403.18105

[10] E. Kasneci et al., "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education." *Center for Open Science*, 2023

[11] D. Delaine, R. Tull, R. Sigamoney, and D. Williams, "Global diversity in engineering education: An exploratory analysis," in *Proceedings of the 2015 International Conference on Interactive Collaborative Learning (ICL)*, 2015, doi: 10.1109/ICL.2015.7318058

[12] A. McKenna, M. Dalal, I. Anderson, and T. N. Y. Ta, "Insights on diversity and inclusion from reflective experiences of distinct pathways to and through engineering education," in Proceedings of the *2018 Collaborative Network for Engineering and Computing Diversity Conference (CoNECD 2018)*, Crystal City, United States, 2018

[13] S. Appelhans, T. De Pree, J. Thompson, J. Aviles, A. Cheville, D. Riley, J. Karlin, S. Fatehiboroujeni, and A. Akera, "From 'Leaky Pipelines' to 'Diversity of Thought': What does 'Diversity' mean in engineering education?" in *Proceedings of the 2019 ASEE Annual Conference & Exposition*, 2019. doi: 10.18260/1-2--32861

[14] H. Touvron et al., "Llama 2: Open Foundation and Fine-Tuned Chat Models," *arXiv.org*, Jul. 19, 2023. Available: https://arxiv.org/abs/2307.09288

[15] “meta-llama/Llama-2-7b-chat-hf · Hugging Face,” huggingface.co. Available: https://huggingface.co/meta-llama/Llama-2-7b-chat-hf

[16] K. McKittrick, *Sylvia Wynter: On Being Humans as Praxis*. 2015

[17] C. Dancy and P. K. Saucier, “AI and Blackness: Towards moving beyond bias and representation,” *IEEE Transactions on Technology and Society*, 2021

[18] D. Workman and C. L. Dancy, “Identifying potential inlets of man in the Artificial Intelligence Development process: Man and antiblackness in AI development,” in *Computer Supported Cooperative Work and Social Computing*, 2023, pp. 348–353. doi:10.1145/3584931.3606981

[19] D. Berreby, “As use of A.I. soars, so does the energy and water it requires,” *Yale E360*, 2024. Available: https://e360.yale.edu/features/artificial-intelligence-climate-energy-emissions

[20] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, “On the dangers of stochastic parrots,” in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, pp. 610–623. doi:10.1145/3442188.3445922

[21] S. Costanza-Chock, *Design Justice: Community-led Practices to Build the Worlds We Need*. Cambridge, Massachusetts: MIT Press, 2020

[22] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, K. Wallach, H. D. III, and K. Crawford, “Datasheets for datasets,” *Communications of the ACM*, vol. 64, no. 12, pp. 86–92, 2021. doi:10.1145/3458723

[23] J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed. Thousand Oaks, CA: Sage Publications, 2017

[24] A. Tashakkori and C. Teddlie, *Handbook of Mixed Methods in Social & Behavioral Research*, 2nd ed. Thousand Oaks, CA: Sage Publications, 2010

[25] M. Fetters, L. Curry, and J. Creswell, “Achieving integration in mixed methods designs—principles and practices,” *Health Services Research*, vol. 48, no. 6, pp. 2134–2156, Dec. 2013