

AI for Climate Justice: Assessing Large Language Models from an Intersectional Lens

Victoria Nguyen, University of California-Irvine, victotn7@uci.edu

Ha Nguyen, Utah State University, ha.nguyen@usu.edu

Hunter King, University of California-Irvine, huntel1@uci.edu

Sara Ludovise, Orange County Department of Education, sludovise@ocde.us

Rossella Santagata, University of California-Irvine, r.santagata@uci.edu

Abstract: Generative Artificial Intelligence (AI) tools, including language models that can create novel content, have shown promise for science communication at scale. However, these tools may show inaccuracies and biases, particularly against marginalized populations. In this work, we examined the potential of GPT-4, a generative AI model, in creating content for climate science communication. We analyzed 100 messages generated by GPT-4 that included descriptors of intersecting identities, science communication mediums, locations, and climate justice issues. Our analyses revealed that community awareness and actions emerged as prominent themes in the generated messages, while systemic critiques of climate justice issues were not as present. We discuss how intersectional lenses can help researchers to examine the underlying assumptions of emerging technology before its integration into learning contexts.

Introduction

Climate change has far-reaching social, economic, political, and cultural implications. Integrating culturally sustaining pedagogy (Paris, 2021) and intersectional climate justice (Amorim-Maia et al., 2022) perspectives highlights how languages, practices, and identities interact within climate change contexts. In this work, we explore the potential of generative artificial intelligence (AI) to create science communication content that elevates intersectional identities. Our work is situated within a larger project, to develop an AI-guided, high school curriculum about climate change and science communication. We leverage OpenAI's GPT-4, a large language model capable of generating textual content with efficient human prompting. We prompt GPT-4 to generate 100 responses about locally situated climate justice issues. The responses integrate random combinations of science communication mediums and intersectional identities. Our research addresses the question: **To what extent are Large Language Models able to generate content pertinent to climate justice issues, solutions, and actions that represent intersectional and culturally sustaining perspectives?**

Intersectional climate justice & culturally sustaining pedagogies

We leverage the Intersectional Climate Justice framework (ICJ; Amorim-Maia et al., 2022) to consider intersectional perspectives within climate education. The framework includes five components: tackle underlying systemic reinforcers of racial and gender inequalities, redress drivers of differential vulnerabilities, take politics and ethics of care seriously, adopt place-based & place-making approaches, and promote cross-identity climate action and community resilience. These five components illuminate how climate change issues intersect with and impact various socioeconomic structures and identities. They account for the political and economic drivers that perpetuate climate vulnerabilities. Climate policies might exclude the voices of marginalized communities, while individuals with limited financial resources face increased challenges to cope with climate risks.

We also employ Culturally Sustaining Pedagogy (CSP; Paris, 2012, 2021), a framework that prioritizes understanding learning environments through the lens of strength-based pedagogy and intersectionality. Four features characterize CSP: firstly, a critical focus on the linguistic and cultural knowledge of communities; secondly, engagement of community actors like students, families, and Indigenous elders in providing input for learning contexts, termed "community accountability"; thirdly, the promotion of positive relationships with the land, Indigenous communities, and fellow learners; and finally, structured opportunities to challenge internalized assumptions that prevent individuals' full participation in learning settings. CSP further critiques systems and ideologies that might marginalize individuals due to their gender identities (Alim et al., 2020). Leveraging the framework enables us to formulate and assess climate education messages to reflect locally situated concerns.

Use of Large Language Models (LLMs) for science communication

Researchers have leveraged LLMs for science communication and examination of climate issues (Gero et al., 2022; Nguyen et al., 2023). However, there exist concerns that the models may harbor inaccuracies and social

biases based on gender, age, race/ethnicity, and ability status (Navigli et al., 2023). Further, LLMs' responses reflect the text corpora they are trained on. Within our study context, LLMs' responses may replicate gaps in the existing climate education literature and overlook human-nature relationships and human agency for climate action (Bang & Marin, 2015). Thus, it is crucial to examine the extent to which LLMs accurately present place-based contexts, and convey intersectional, multifaceted perspectives regarding climate issues.

Methodology

Our unique life experiences shape our approach to collaboration, analysis, and report of research findings. Our team comprises two university professors, an informal science educator, a PhD student, and an undergraduate student. The first author is an Education PhD student and brings her perspective as a Vietnamese American woman and a former informal environmental educator for K-12 students. The second author is a Learning Technologies assistant professor with Vietnamese descent. The third author is an undergraduate researcher and brings perspective as a Korean American with backgrounds in sustainability and law. The fourth author is a nonformal science educator of Jewish descent who coordinates a county-wide environmental education program. The last author is a professor of education of Italian descent whose research centers on improving STEM learning experiences for racially, culturally, and linguistically minoritized students.

Data

We leveraged an LLM (OpenAI's GPT-4) to create climate education content. The prompt included various science communication mediums (e.g., Twitter, forum discussion, personal narrative). We included descriptors for race, socioeconomic (lower-class, middle-class, upper-class), ability (e.g., able-bodied, sensory impairment, mental health), gender (female, male, transgender, non-binary, agender), and age range. Integrating these intersectional identities allows us to examine how LLMs represent diverse forms of privilege and discrimination within climate issues. Given the study context and the focus on place-based approaches, the prompts included the names of eight prominent counties across California. Finally, we included terms to cover different topics, such as "climate resilience", "sustainable ways of living", and "climate justice barriers". The prompt takes the form: Write a {medium} about {term} in {county} written by {socioeconomic} {race} {gender} {age} {ability}.

Analysis

We performed a content analysis to examine the LLM's responses, drawing from the ICJ (Amorim-Maia et al., 2022) and CSP (Paris, 2021) frameworks. Building on the CSP framework, we included LGBTQ+ community in our code definition to account for the underlying systemic reinforcers of racial and gender inequalities. We incorporated community accountability, as we examined how the LLM presented opinions from different identity groups. We refined the codes through three rounds of discussion between the first two authors based on 15 responses. We included codes for systemic and individual action, based on the focus on raising awareness and creating collective momentum in recent literature (Trott et al., 2023). Each response was coded for multiple instances of a code. The first and third author reached substantial inter-rater agreement on 15% of the code (81.25%) and split the remaining dataset to code. The researchers met weekly to resolve any disagreements.

Findings

Community awareness and actions are prominent

We found multiple instances of **cross-identity climate action and community resilience building** ($n = 262$, out of 667 coded instances). There were two prominent subcodes: **awareness justice** ($n = 129$) and **empower local communities** ($n = 107$). The **awareness justice** subcode highlighted the impact of climate change on marginalized communities, often linked to educational efforts such as community workshops and newsletters. Meanwhile, **empowering local communities** emphasized explicitly calling community members to action or providing examples of individuals conducting community-based climate justice work. Consider the following response, where GPT-4 was prompted to generate a podcast series about the success of "a White, female child with chronic illnesses" in promoting climate justice cultural knowledge and practices in Fresno County.

Listeners are invited into the life of our young protagonist, exploring the intersections of her cultural knowledge, advocacy for climate justice, and personal health battles. She shares her experiences and the wisdom her fight against chronic illnesses has bestowed upon her. [...] "Blossoming Voices" also engages with a diverse array of individuals from Fresno County and beyond ...

This response attends to both awareness justice and community empowerment. It provides a specific example of an individual preserving health (“fight against chronic illnesses”) and advocating for climate justice (“health and the environment are interconnected”). It also incorporates call for diverse individuals and youth voices from the community (“Fresno County and beyond”) in climate justice issues. While this example is promising, we noted responses that simply restated the prompts without much elaboration. The responses may give a generic overview such as “gender discrimination and injustices against Indigenous cultures” and use words like “diversity”, “resilience”, and “inclusivity”, without examples to provide meaning to the terms.

Systemic critiques of climate justice issues are not as present

There were also codes that were not as frequent, including **tackle underlying systemic reinforcers of racial and gender inequalities** ($n = 11$) and **call out colonial narratives** ($n = 1$). Recognizing the infrequent codes allowed us to further understand the underlying assumptions of LLMs. The first code (tackle systemic reinforcers) emphasized the adoption of climate principles and actions that incorporate *racial and gender equality* ($n = 5$) or critiqued systemic legacies that have perpetuated climate insecurity (*dismantle systems*; $n = 6$). For example, a generated response reads: “Transgender and disabled communities too often get left out of the conversation. Leilani is changing that narrative and reminding us all about the importance of inclusivity”. We coded these responses for *gender equality*, for their explicit acknowledgement of LGBTQ+ community.

Additionally, we only found one instance of the subcode **call out colonial narratives** under the code “adopting place-based and place making approaches”. The response reads: “I realize that it's similar to how indigenous people might look at their surroundings. It's so different from mainstream western thinking. Their eco-friendly ways of life, cultivating plants, and honoring seasons might be so beneficial if incorporated in our parks and around the city.” This response explicitly acknowledges “western thinking” as the dominant perspective and highlights that there are merits of incorporating Indigenous perspectives into urban planning. However, one might argue that the example might stereotype Indigenous knowledges. We examine this issue in the next section.

Place-based details, stereotypes, and deficit languages

We highlight occurrences of **place-based details** ($n = 32$) that identified accurate environmental issues that the respective county is facing such as “rampant wildfires”, “droughts”, and “precarious air quality”. In comparison, we found rare but notable instances of **inaccuracies** ($n = 8$) and explicit **stereotypes** ($n = 6$). The inaccuracies were found to generalize climate issues or include incorrect place-based details, for example, claiming “hurricanes displacing thousands” as a climate issue in California. We also spotted stereotypes, for example, when the LLM gendered Earth as a female, mother figure. Additionally, in simulating an online conversation about climate issues, GPT-4 used account handles like “FreeSpiritLove” when referring to an Indigenous individual. In another instance, the LLM assumed that all Middle Eastern communities were Muslim. Such simplification ignores the diversity of experiences, interests, and beliefs within these communities and can spread harmful associations.

Discussion

Findings reflect the assumptions and gaps for climate science communication in LLM-generated responses. Intersectional climate justice is predominantly portrayed through raising awareness and participating in environmentally friendly actions. These patterns are not surprising, given the emphasis on individual and systemic actions in climate justice literature (Martiskainen et al., 2020; Trott et al., 2023). However, the generated responses do not fully encompass the ICJ and CSP frameworks. Few responses acknowledged systemic, colonial perspectives or substantially represented intersecting gender identities in relation to climate justice. These findings reflect the gap in climate education, which rarely calls out colonial perspectives (Ajaps, 2023). Furthermore, they underrepresent women’s active participation in climate justice spaces (Godden et al., 2020) and undermine the disproportionate impact of climate change on LGBTQ+ communities and women (Goldsmith & Bell, 2022). LLMs reflect the text corpora they are trained on. The lack of coverage for calling out colonial narratives, LGBTQ+ voices, and women’s perspectives points to gaps in the training data.

Our findings have several practical implications. First, we illustrate the specification of perspectives, environmental issues, and locations, as one way to improve the response quality to include place-based details. Second, auditing the models’ responses with frameworks like ours can provide insights into their affordances and shortcomings. While we did not find frequent responses that were overtly inaccurate or biased, we include notes of caution for leveraging LLMs in science communication contexts. A major critique of models like GPT-4 is its knowledge cutoff, limiting non-finetuned models from providing up-to-date information. Science communicators can examine LLM-generated responses for accuracy and bias, before large-scale implementation. Relatedly, educators can adopt the framework into guiding prompts and learning activities, to help students develop critical understanding of missing or misrepresented perspectives, inaccuracies, and stereotypes in LLMs.

The limitations of our work can guide future research. First, we did not involve human evaluation outside of the research team. We call for future work to involve individuals with additional, diverse backgrounds to audit LLM-generated responses. Second, we find different levels of depth in the responses. Further research can explore different prompting strategies, to represent perspectives in climate science communication more accurately and comprehensively. Finally, we only experimented with one LLM (GPT-4) and did not include extensive fine-tuning. This decision was intentional, as we aimed to gain a baseline understanding of the model's assumptions. Researchers can experiment with other LLMs, to investigate the limitations and capacities of the different models.

Conclusion

Our research investigates the potential of using LLMs within climate education to construct effective science communication messages, based on Intersectional Climate Justice and Culturally Sustaining Pedagogy lenses. Findings indicate that while community awareness and actions related to climate justice are prevalent, the systemic critiques of climate justice issues are less prominent. Our work illuminates the underlying assumptions of LLMs regarding climate justices and intersecting identity dimensions. These findings highlight the need to develop more effective prompts for LLMs, enabling them to represent culturally sustaining knowledge, intersectional identities, and localized climate issues and solutions, prior to deployment in educational contexts.

References

- Ajaps, S. (2023). Deconstructing the constraints of justice-based environmental sustainability in higher education. *Teaching in Higher Education*, 28(5), 1024-1038.
- Alim, H. S., Paris, D., & Wong, C. P. (2020). Culturally sustaining pedagogy: A critical framework for centering communities. In *Handbook of the cultural foundations of learning* (pp. 261-276). Routledge.
- Amorim-Maia, A. T., Anguelovski, I., Chu, E., & Connolly, J. (2022). Intersectional climate justice: A conceptual pathway for bridging adaptation planning, transformative action, and social equity. *Urban Climate*, 41, 101053. <https://doi.org/10.1016/j.uclim.2021.101053>
- Bang, M., & Marin, A. (2015). Nature–culture constructs in science learning: Human/non-human agency and intentionality. *Journal of Research in Science Teaching*, 52(4), 530-544. <https://doi.org/10.1002/tea.21204>
- Gero, K. I., Liu, V., & Chilton, L. (2022). Sparks: Inspiration for Science Writing using Language Models. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference (DIS '22)* (pp. 1002-1019). Association for Computing Machinery, New York, NY. <https://doi.org/10.1145/3532106.3533533>
- Godden, N. J., Macnish, P., Chakma, T., & Naidu, K. (2020). Feminist Participatory Action Research as a tool for climate justice. *Gender & Development*, 28(3), 593-615. <https://doi.org/10.1080/13552074.2020.1842040>
- Goldsmith, L., & Bell, M. L. (2022). Queering Environmental Justice: Unequal Environmental Health Burden on the LGBTQ+ Community. *American Journal of Public Health*, 112, 79-87. <https://doi.org/10.2105/AJPH.2021.306406>
- Martiskainen, M., Axon, S., Sovacool, B. K., Sareen, S., Furszyfer Del Rio, D., & Axon, K. (2020). Contextualizing climate justice activism: Knowledge, emotions, motivations, and actions among climate strikers in six cities. *Global Environmental Change*, 65, 102180. <https://doi.org/10.1016/j.gloenvcha.2020.102180>
- Navigli, R., Conia, S., & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *Journal of Data and Information Quality*, 15(2), Article 10. <https://doi.org/10.1145/3597307>
- Nguyen, H. (2023, April). Exploring natural language models' responses to intersectional identities in climate change education. *AERA Annual Meeting 2023*. Chicago, IL.
- Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational researcher*, 41(3), 93-97.
- Paris, D. (2021). Culturally Sustaining Pedagogies and Our Futures. *The Educational Forum*, 85(4), 364-376. <https://doi.org/10.1080/00131725.2021.1957634>
- Trott, C. D., Gray, E.-S., Lam, S., Courtney, R. H., Roncker, J., & Even, T. L. (2023). People's action for climate justice: A systematic review. *Local Environment*, 28(9), 1131-1152. <https://doi.org/10.1080/13549839.2023.2187363>

Acknowledgements

This research is funded by the National Science Foundation (DRL-2241596).