International Society of the Learning Sciences

Toward Productive Multivocality in AI Development: Excavating Ethics Concerns among an Interdisciplinary Team

Grace (Yaxin) Xing, University at Buffalo, yaxinxin@buffalo.edu
Sanaz Ahmadzadeh Siyahrood, Georgia Institute of Technology, ssiyahrood3@gatech.edu
X. Christine Wang, University at Buffalo, wangxc@buffalo.edu
Jinjun Xiong, University at Buffalo, jinjun@buffalo.edu

Abstract: Adapting the *productive multivocality* framework (Suthers et al., 2013), we engage an interdisciplinary team, which designs AI tools for children with speech and language challenges, in developing guidelines for ethical AI development and deployment. In the initial phase of this 5-year study, we investigated ethical concerns about AI in general and the AI tools being designed in particular across the varied discipline teams. Employing thematic analysis, descriptive statistics, BERT topic modeling, and WordCloud tool, we analyzed survey data and focus group interviews of 13 researchers in Speech-Language Pathology/Learning Sciences, Human-Computer Interaction, Multimodality, and Core Technology. Our analysis uncovered prevalent unease about AI, along with nuanced and varying degrees of concerns regarding the AI tools under development across different disciplines. The findings inform our broader study and highlight the potential role of productive multivocality in fostering responsible and equitable AI development and its eventual implementation.

Introduction

As artificial intelligence (AI) increasingly permeates our daily lives, the imperative for ethical AI development and deployment becomes paramount, especially in educational contexts (The White House, 2023). While various ethical guidelines for AI exist, researchers challenge the efficacy of current AI ethics tools, pushing for methods rooted in best practices from different sectors to combat structural injustices (Ayling & Chapman, 2022; Heilinger, 2022). To address this gap, we are conducting an ethical study by following the IEEE Standard 7000-2020 (Olszewska, 2020) to establish comprehensive ethical guidelines for the tool development and deployment at the National Science Foundation (NSF)/Institute of Education Sciences (IES) jointly funded AI Institute for Exceptional Education (hereafter, *AI4ExceptionalEd*; https://www.buffalo.edu/ai4exceptionaled.html).

Our study adapts the *productive multivocality* framework (Suthers et al., 2013). Rooted in Bakhtin's (1981) seminal work on multivocality, the framework promotes knowledge expansion through interdisciplinary collaboration by transforming individual disciplinary contributions into a cohesive, interdisciplinary synthesis (Oshima et al., 2015). We apply its five core principles to help us move toward productive multivocality: valuing diverse perspectives, promoting open dialogue, engaging in reflective practice, co-constructing knowledge, and respecting complexity (Suthers et al., 2013). This paper concerns the first phase of our study, which focuses on excavating ethical concerns among the AI4ExceptionalEd's interdisciplinary team and investigates three overarching questions: (1) What are ethical concerns about AI in general across the team members' disciplines? (2) What are ethical concerns about the team's targeted AI tools across the team members' disciplines? and (3) How are these two types of ethical concerns related to each other?

Methods

Employing a case study approach, we investigated the complex ethical considerations confronting the interdisciplinary team at AI4ExceptionalEd. Supported by NSF and IES, AI4ExceptionalEd aims to mitigate the exacerbated shortage of Speech and Language Pathologists (SLPs) due to the COVID-19 pandemic by creating two advanced AI technology suites: the AI Screener and the AI Orchestrator. The Institute's unique focus on AI tools for young children, SLPs, and educators as end-users, combined with the diverse expertise of a large interdisciplinary team, makes it a "telling case" (Mitchell, 1984, p. 239) for examining responsible and equitable AI development and deployment. The ethical study's first phase began in September 2023, five months after the Institute's launch, capitalizing on a moment when team members' varying perspectives were most pronounced due to early cross-disciplinary interactions and collaborations.

Participants

We recruited participants from four key disciplinary fields (SLP/Learning Sciences, HCI, Multimodality, and Core Technology) at AI4ExceptionalEd. Fourteen agreed to participate in individual surveys, and 13 continued

participating in the focus group interviews. The largest group among the participants was the SLP/Learning Sciences (LS) with six members, followed by HCI and Core Technology fields, each with three participants, and the smallest group was Multimodal with just one researcher.

Data collection and analysis

Data was collected through individual surveys and focus group interviews organized within disciplines. The survey consisted of five open-ended questions about their experiences, attitudes, and concerns about AI in general. Structured interview questions focused on ethical concerns related to the AI Screener or/and the AI Orchestrator depending on which one(s) participants focused on. The interviews were conducted via Zoom and the length ranged from 20 to 90 minutes.

Survey data was organized in an Excel sheet, while interview recordings were transcribed and analyzed from CSV files after thorough data cleaning. This preparation involved standardizing text by removing irrelevant characters and converting it to lowercase for accurate word frequency analysis. We integrated thematic analysis, descriptive statistics, BERT topic modeling, and WordCloud tool in comprehensive qualitative and quantitative analysis of the survey and interview data. While the thematic analysis was conducted to identify recurring and divergent themes, descriptive statistics were used to show the general patterns of responses. BERT topic modeling was used to visualize key themes and topics within the data. Additionally, WordCloud was used to represent keyword frequency and conceptual relationships.

Findings and discussion

AI experiences, attitudes, and ethical concerns

There were varying experiences with AI across teams. The Core Technology team had the most extensive experiences and knowledge (e.g., “working on several topics in deep learning” “working on a fundamental theory for AI” and “decades of experience”). Although the Multimodality and the HCI teams had a similar amount of AI experiences as the Core Technology team, their work focused more on human-centered AI applications (e.g., “computer-child interaction” “social robotics” and “detecting cognitive states of learners”). The SLP/LS team had the least AI experiences, positioning themselves more as domain experts rather than technologists.

Individually, the participants' attitudes toward AI varied widely ranging from ambivalence to optimism. About 35% of participants had mixed feelings, recognizing both AI's strengths and weaknesses; 21% were neutral, viewing AI as a tool or a "superpower baby" shaped by humans; and 43% were positive, stressing the potential benefits of AI. Collectively, the Core Technology team was the most optimistic about AI's potential, followed by the HCI and Multimodality teams while the SLP/LS team was decidedly more mixed about AI’s potential.

Connecting participants’ AI experiences and their attitudes toward AI showed that individuals (mostly SLP/LS team members) new to AI adopted a balanced stance being “open” toward its “positive” effects while also concerned about its “negative” outcomes or potential “dangers.” But they often overlooked human agency and their abilities to address some of the problems. In contrast, those with rich AI experience (e.g., Core Technology, HCI, Multimodality teams) recognized AI as a tool that is controlled and reshaped by humans. They exhibited a pragmatic optimism, marveling at what AI could achieve, yet being aware of its limitations and social responsibilities. The enhanced familiarity with AI is associated with a more nuanced perspective that acknowledges the technology's promise alongside its ethical and societal implications, which is consistent with Ehsan et al.'s (2021) findings. This underscores the imperative of incorporating a broad spectrum of multidisciplinary expertise and multivocality in the development of AI to ensure a well-rounded and informed approach.

Regarding ethical concerns about AI in general, the most frequently cited concerns were “Disinformation and Misuse” (29%) and “Lack of Understanding of AI's Limitations” (29%), highlighting the risks of misinformation and overestimation of AI's capabilities. Other concerns included “Ethical and Legal Aspects” (21%), “Lack of Human Touch and Social Skills” (14%), “Over-Reliance of AI” (14%), and “Bias and Discrimination” (14%), pointing to the need for ethical and legal frameworks in AI regulation.

Ethical concerns about AI screener/AI orchestrator

Our analysis of participants’ potential ethical concerns about the AI Screener/AI Orchestrator revealed a collective agreement on fundamental issues as well as distinct insights reflecting their discipline-associated viewpoints.

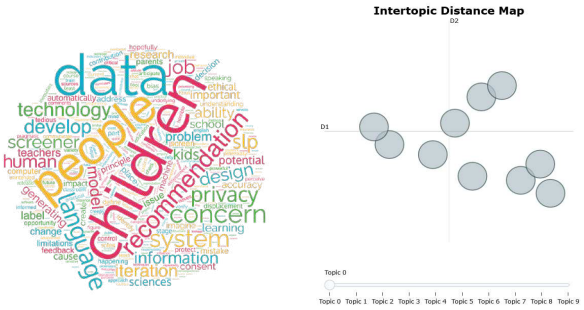
Common themes

The WordCloud in Figure 1 prominently features terms like “children,” “people,” “data,” “recommendation,” “language,” “concerns,” “privacy,” and “system.” The accompanying Intertopic Distance Map suggests a thematic

interconnectedness among these subjects. Also, each circle representing a topic has the same size. This suggests that all important topics were given equal importance or consideration in the analysis.

Figure 1

Word Cloud and BERT Topic Modeling Results



A deeper analysis of the interview transcripts yields four prevalent themes: (1) Data Privacy, Consent, and Ownership: All teams emphasized data privacy and the secure handling of sensitive information, e.g., young children's multimodal data (e.g., facial, behavioral, vocal, and semantic information). They raised questions about the extent and nature of informed consent in classrooms where not all may consent, as well as data access, ownership, and rights as children grow older. They were also concerned about the potential harm of AI rigidly labeling or categorizing children; (2) Human Oversight and Final Responsibility: There was a consensus across teams that AI should support, instead of replacing, decision-making by professionals, who should maintain final authority and accountability; (3) Transparency and Educating about AI's Limitations: Participants across teams advocated clear communication about AI's functionality and limitations, purshing for critical assessment of AI tools by end-users as well as educating public to prevent misunderstandings; and (4) Equity, and Fairness: The SLP/LS and Multimodality teams consistently expressed concerns about AI's potential to address or exacerbate inequalities. They stressed the need to ensure equitable access for marginalized groups and minimize bias, especially in language and speech assessments of bilingual and multilingual children.

Although different teams shared similar concerns or "voices," their contributions or "productivity" in addressing these concerns could vary significantly. For example, while the SLP/LS team and the Multimodality team both focused on equity and fairness, the former might provide insights into the pedagogical and practical implications, whereas the latter might offer more specific technological solutions. This variation is crucial as it allows for a more integrated approach to AI development and deployment.

Distinct themes

The Core Technology team emphasized AI's environmental impact and suggested smaller computing models to minimize carbon footprint. They also stressed the unrealistic *expectations* people often have of AI, emphasizing that individuals should understand that AI, much like *humans*, can make *mistakes*. The *HCI* team was optimistic about the AI Screener/AI Orchestrator's role in transforming SLP's jobs but discussed concerns about potential "job displacement" or SLP's "over-reliance" on these tools. The Multimodality team warned about privacy concerns and misuse of AI by educators, while the SLP/LS team stressed the need for safeguards against unintended AI outcomes in language education, advocating for features that give users more control and flexibility. They also highlighted the importance of equitable AI integration in schools.

The multivocality approach underscores that the distinct themes raised by each team are not merely separate strands but integral parts of a complex mosaic, reflecting the multifaceted nature of AI's implications. For instance, the Core Technology team's emphasis on the environmental impact of AI operations resonated with broader ethical considerations, intersecting with the HCI team's focus on job displacement and societal values, as both teams were essentially considering the long-term implications of AI on society and the planet. Similarly, the SLP/LS team's call for user-empowering features like a "redo" button dovetailed with the Multimodality team's apprehensions about data misuse, as both advocated for greater control and ethical oversight.

Connections between concerns about AI in general and the specific AI tools

The AI Screener and AI Orchestrator, designed to assist SLPs and young children with speech and language challenges, bring to the fore specific ethical concerns that resonate with the broader discourse on AI. Issues such as ethical issues, disinformation, and a lack of understanding about AI's limits are not unique to this tool but are part of a wider conversation about responsible AI deployment. For example, the risk of misinterpretation by

teachers and parents using the AI Screener's results could lead to unintended consequences, highlighting the need for clear communication and education about AI's capabilities and limitations.

However, these tools also offer chances to alleviate apprehensions about AI in general. The AI Screener and AI Orchestrator aim to augment, not supplant, SLP expertise, assuaging fears of job loss or diminished human skills. By automating mundane tasks, they could allow SLPs to focus on more complex cases, potentially improving intervention quality. Involvement in AI creation can dispel some of the mystique surrounding the technology, fostering a sense of agency and comprehension that might lessen ethical worries. As one participant shared, "My concerns are lessened by my involvement in the creation process; understanding its mechanics provides reassurance." Nonetheless, it's vital to avoid becoming complacent, assuming we need only worry about general AI ethical issues and not those specific to one's own team's creations.

Conclusions and implications

The discourse surrounding AI, especially when involving vulnerable groups like young children, is multi-layered. Ethical concerns articulated by interdisciplinary teams encompass a spectrum of issues, including disinformation, misuse, and lack of understanding of AI's limitations, underscoring the urgency for robust ethical and legal frameworks. Shared themes such as data privacy, informed consent, human oversight, and equity are juxtaposed with unique concerns like environmental sustainability, potential job displacement, and the amplification of biases within the AI Screener/Orchestrator framework. These general and specific ethical issues are interrelated, as the broader concerns about AI's societal impact are echoed in the nuanced apprehensions regarding the AI Screener/Orchestrator, highlighting the importance of integrating diverse expertise to navigate the ethical landscape of AI development and application.

Based on the findings, we strongly advocate for increased transparency in AI development, workforce training, proactive bias reduction, stringent data protection, efforts to reduce AI's ecological footprint, and increasing public's understanding of AI's capabilities and boundaries. Such a comprehensive strategy is critical for AI's ethical and responsible deployment in schools. The implications of this study highlight the essential role of multivocality in AI development, advocating for a collaborative, interdisciplinary approach that embraces diverse ethical perspectives to ensure responsible and equitable integration of AI technologies.

References

Ayling, J., & Chapman, A. (2022). Putting AI ethics to work: Are the tools fit for purpose? *AI Ethics, 2*, 405–429. https://doi.org/10.1007/s43681-021-00084-x

Bakhtin, M. (1981). Discourse in the novel. In M. Holquist (Ed.), *The Dialogic Imagination* (pp. 259-422). Austin: University of Texas.

Ehsan, U., Passi, S., Liao, Q. V., Chan, L., Lee, I., Muller, M., & Riedl, M. O. (2021). The who in explainable AI: How AI background shapes perceptions of AI explanations. ACM CHI. https://doi.org/10.48550/arXiv.2107.1350

Heilinger, J. C. (2022). The ethics of AI ethics. A constructive critique. *Philosophy & Technology, 35*(3), 61. https://doi.org/10.1007/s11023-020-09517-8

Mitchell, J. (1984). Typicality and the case study. In R. Ellen (Ed.), *Ethnographic Research: A guide to general conduct* (pp. 237 - 241). London: Academic Press.

Olszewska, J. I. (2020). IEEE recommended practice for assessing the impact of autonomous and intelligent systems on human wellbeing. *IEEE Standard*, 7010-2020. https:// doi.org/10.1109 /IEEESTD.2020.9084219

Oshima, J., Oshima, R., & Fujita, W. (2015). A multivocality approach to epistemic agency in collaborative learning. In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, S. Ludvigsen (Eds), *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning (CSCL '2015), Exploring the Material Conditions of Learning* (Vol. 1, pp. 62–69). Gothenburg, Sweden: International Society of the Learning Sciences. https://doi.dx.org/10.22318/cscl2015.146

Suthers, D. D., Lund, K., Rosé, C. P., & Teplovs, C., (2013). Achieving productive multivocality in the analysis of group interactions. In *Productive multivocality in the analysis of group interactions* (pp. 577-612). Boston, MA: Springer.

The White House. (2023, October 30). *Fact sheet: President Biden issues executive order on safe, secure, and trustworthy artificial intelligence*. The White House.

Acknowledgments

We thank all the participating members of the AI4ExceptionalEd team for their time and support for this project.