The Challenge of Incorporating End-User Values into Design: A Methodological Perspective of Using Provotypes

Marie Utterberg Modén*[1], Tiina Leino Lindell[1], Johan Lundin[1], and Marisa Ponti[1]

[1] *University of Gothenburg, Department of Applied IT, Sweden*

Abstract

This study addresses the critical need for ethics and fairness in AI and machine learning, focusing on the often-overlooked inequalities behind biases in these technologies. Adopting a value-sensitive design approach, it investigates a design method that use conflicts through the introduction of 'provotypes', intended to enhance end-user agency in the AI tool design process. Specifically concentrating on the educational sector, centers on teachers, this paper offers an in-depth perspective at both the application of these methods and the outcomes they produce, covering both methodological insights and findings related to values on AI in education.

Keywords

end user design, value sensitive design, artificial intelligence, education

1. Introduction

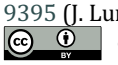
There is an increasing recognition of the importance of ethics and fairness in machine learning and AI systems research. Despite this, much of the effort has been on examining and addressing biases by implementing 'fairness-aware' algorithms, rather than on understanding and handling the deeper, systemic inequalities that these biases may reflect or reinforce [1]. There is an identified gap and need for more proactive, inclusive approaches, such as participatory methodologies, that involve stakeholders in conceptualizing and designing AI tools [2, 3]. The active involvement of practitioners in the design phase is critical for enhancing the quality of AI-tools and ensuring sustainable implementation [4]. Moreover, engaging end-users and stakeholders from the beginning is vital for safeguarding the freedom, values, and rights of those the AI-tool is designed for [5, 6, 7]. Fairness is a value often emphasized in the design and use of AI-tools [2], highlighting the "ethical need to understand the historical and social contexts into which these systems are being deployed" [8, p. 2]. To explore fairness of AI-tools, they should be examined within both a broad societal perspective and the specific context of its application. The notion of

Proceedings of the 8th International Workshop on Cultures of Participation in the Digital Age (CoPDA 2024): Differentiating and Deepening the Concept of "End User" in the Digital Age, June 2024, Arenzano, Italy

* Corresponding author.

marie.utterberg@ait.gu.se (M. Utterberg Modén); tiina.leino.lindell@ait.gu.se (T. Leino Lindell); johan.lundin@ait.gu.se (J. Lundin); marisa.ponti@ait.gu.se (M. Ponti)

0000-0002-3820-4063 (M. Utterberg Modén); 0000-0001-9444-7513 (T. Leino Lindell); 0000-0001-5547-9395 (J. Lundin); 0000-0003-4708-4048 (M. Ponti)



fairness is not just about abstract principles but involves critically analyzing who stands to determine what fairness means, who benefits from these definitions, and the power dynamics influencing these determinations [1]. In this way, "many of the sources of unfairness are not straightforward to identify but instead require thorough domain knowledge" [9, p. 382]. However, articulating and meaningfully translating the values and needs of stakeholders into the design process can be challenging [10]. Moreover, exploring the values and needs of various stakeholders may uncover conflicts [2]. This study adopts a value-sensitive design [6] perspective and explores a design method aimed at taking advantage of conflicts and using them to provoke critical discussions and the exploration of novel ideas. By doing so, it introduces 'provocative prototypes', or provotypes [11, 12], as a tool that advocates diverse perspectives and tensions of values as a catalyst for creativity and innovation. The current study centers on teachers, who play a critical role in implementing fairness within their daily activities in schools. It details an ongoing project intending to empower secondary school teachers in the design of prototypes of AI-based educational tools that uphold fairness. This paper offers an in-depth perspective at both the application of these methods and the outcomes they produce, covering both methodological insights and findings related to end-users' perspectives on AI in education.

2. The design process

2.1. Participants

Forty participants took part in this study, which included teachers, who were considered direct users of AI tools, and principals, students, and pedagogical developers, who were identified as indirect users, with 10 individuals from each category. They were purposefully recruited based on their interest in exploring AI in education, representing secondary schools across four municipalities in Sweden. We chose to engage with these specific participants to gain insights from a group that generally does not have a say in design of AI tools [13] but where teachers, principals and pedagogical developers are responsible for distributing fairness in their everyday work [14].

2.2. Workshops design

A series of two workshops and focus group discussions were conducted for each category of participants for a total of 40. Both the first and second workshops lasted approximately three hours. The first workshop was audio-recorded, while the second was video-recorded. The focus group discussions, which lasted about one hour, involved smaller groups from each category and were audio-recorded. All sessions took place at the university.

Workshop 1 - End users' perception of fairness. A week before the workshop, we distributed a video to teachers, pedagogical developers and principals, providing them with an introduction to the basics of AI. It has been highlighted that providing participants with knowledge of advanced and complex technologies is necessary to bridge the gap between nonprofessional and professional designers/researchers [15]. Thus, the first workshop began with a lecture about AI, with a particular emphasis on understanding biases, both broadly and within the specific area of education.

The workshop consisted of two sessions. The first session aimed to highlight the participants' efforts in promoting fairness within their schools. It emphasized that teachers and principals naturally promote equity through their daily practices. Meanwhile, pedagogical developers are tasked with supporting schools, which includes advocating for equitable education for all students. Participants were asked to reflect on the following question: What specific actions do you take in your work to create a fair and equitable school? Initially, they individually documented on post-it notes strategies they had personally implemented. Subsequently, they formed small groups to collaboratively discuss those strategies.

The second session was designed to focus on the role of AI in education. Participants were encouraged to pair up and create a storyboard illustrating a future scenario, a 'sketch of use' [16], scenarios that highlights values emphasizing the social and ethical considerations of new technologies [17] focusing on how AI could be used to accomplish work tasks and other activities. The workshop concluded with a collective discussion where participants shared and reflected on their work in relation to fairness.

Workshop 2 - Provotypes illustrating conflicting values. In preparation for the second workshop, our analysis of transcribed recordings from workshop 1 revealed tensions in participants perspectives. We identified three primary areas of tension that reflect both the participants' values regarding fairness and their conceptualizations of using AI tools in education. The first area of tension addresses the challenge of balancing personalized education for each student against the need to frame the classroom as a space for collaborative work and discussion. The second area points the tension between the efficiency of monitoring students through AI for time-saving analysis and the depth of understanding that teachers achieve through direct interaction with their students. Lastly, the third area of tension contrasts the benefits of obtaining data-driven insights on student with the ethical imperative to respect students' privacy and protect their personal information.

Although these tensions in participants values are widely recognized in the field of AI and fairness research [18, 19], the persistent challenge is how to thoughtfully mitigate these tensions and coherently integrate them into AI tool design [3] In design processes, tensions stemming from conflicting values among stakeholder perspectives have traditionally been addressed by devising strategies to foster consensus within or among stakeholder groups [20, 21].

Instead of circumventing these tensions, we have used them as a resource, creating provocative prototypes (provotypes) that embody these very tensions [12]. From a theoretical perspective, Activity Theory [22] can serve as a foundation for a provotyping approach [11]. Activity theory is based on the concept that activities are inherently subject to systemic contradictions, which act as catalysts for change processes, ultimately leading to transformation of the activity [22]. This understanding has guided our development of three provotypes designed to actively engage with such contradictions.

The participants were divided in small groups and each group was given paper-provotypes. The participants were told that the provotypes illustrated ideas capturing their varied imaginaries of AI-tools in education from previous discussions. Provotyping, as a method, focuses on identifying and highlighting contradictions within a specific practice

[11]. Thus, by interacting with the provotypes, participants from different educational backgrounds and schools were able to critically confront and reflect upon the varied and sometimes conflicting ideas of AI in education. The provotypes aimed to act as catalysts, stimulating creativity and encouraging new ideas by questioning norms and values while designing for future practices [12] in education. In this way, provotyping was viewed as an intermediary, linking the exploration of current concrete practices with the imagination of future opportunities by uncovering values and intended to facilitate the transition from analysis to design [12]. By exposing contradictions, provotyping aimed to address the identified contradictions and inform design [11]. To do so, participants redesigned the provotypes and created their own interface prototypes. They were equipped with plain paper prototypes along with a selection of pens in various colors, sticky notes, scissors, rulers, glue, and pre-made stencils of elements like buttons, icons, and form fields.

Focus group discussions - Prototype-stimulated discussions. In preparation for the focus group discussions, students were invited to engage with, and respond to, the three provotypes and create prototypes. They were divided into small groups and encouraged to reflect on the perceived strengths and weaknesses of the provotypes, their personal values related to its use and functionality, and any ethical or practical concerns they identified.

Together, teachers, principals, pedagogical developers, and students generated a wide variety of detailed prototype designs in response to the provotypes. A selection of these prototypes and video-recorded reflective discussions, which took place during their creation, collectively served as 'stakeholder prompts' [17] in subsequent focus group discussions. These prompts were intended to elicit the underlying values, guiding the conversation and analysis in these groups, which were homogeneous, each consisting of either teachers, principals, or pedagogical developers.

3. Results and discussion

Participants design ideas and values were reflected through a diversity of methods and representations, including scenarios, prototypes, and stakeholder prompts, alongside verbal presentations of the designs in relation to fairness. During the design process, we recognized the necessity for participants to switch back and forth between concrete and abstract thinking to uncover their values.

Teachers value direct, personal interaction with students, viewing it as essential for comprehensively understanding each student's emotional, social, and practical needs. Thus, the teacher's role is seen as irreplaceable, with human insight and empathy being crucial for supporting students. While integrating AI, many participants also highlight 'conservation values' [23], emphasizing the importance of preserving the stability and maintaining traditional educational practices, such as age-based classes and student collaborative group work.

At the same time, the participants were open to assigning a wide range of tasks to AI tools, motivated by the desire to enhance student outcomes. They were willing to delegate tasks such as automated individualization and progress monitoring (bordering on surveillance) to prevent students from falling behind, along with other supportive functions, to AI tools to augment teacher capabilities. This willingness could be interpreted

as placing the value of high-quality education for all students above concerns for students' data privacy.

Throughout the design process, a significant challenge has been inspiring participants to broaden their understanding of AI tools. Encouraging them to envision how AI can innovate and transform education necessitates a shift in perspective and moving beyond traditional approaches to explore the novel possibilities and risks that AI introduces to the field. An additional challenge has involved raising awareness about the potential consequences of AI usage on the teaching profession, including the risk of deskilling and maintaining elements of teaching that contribute to job satisfaction.

Acknowledgements

This work was supported by the Marianne and Marcus Wallenberg Foundation under Grant MMW 2021.0030.

References

[1] A. D. Selbst, D. Boyd, S. A. Friedler, S. Venkatasubramanian, and J. Vertesi, "Fairness and Abstraction in Sociotechnical Systems," in Proceedings of the Conference on Fairness, Accountability, and Transparency, Atlanta GA USA: ACM, Jan. 2019, pp. 59–68. doi: 10.1145/3287560.3287598.

[2] I. Gan and S. Moussawi, "A Value Sensitive Design Perspective on AI Biases," presented at the 55th Hawaii International Conference on System Sciences (HICSS), Hyatt Regency Maui, Hawaii, USA: University of Hawai'i at Manoa, Hamilton Library, Jan. 2022, pp. 5548–5557.

[3] J. Lundin, M. Utterberg Modén, T. Leino Lindell, and G. Fischer, "A Remedy to the Unfair Use of AI in Educational Settings," IxD&A, no. 59, pp. 62–78, Dec. 2023, doi: 10.55612/s-5002-059-002.

[4] E. B.-N. Sanders and P. J. Stappers, "Co-creation and the new landscapes of design," CoDesign, vol. 4, no. 1, pp. 5–18, Mar. 2008, doi: 10.1080/15710880701875068.

[5] S. Bødker and M. Kyng, "Participatory Design that Matters—Facing the Big Issues," ACM Trans. Comput.-Hum. Interact., vol. 25, no. 1, pp. 1–31, Feb. 2018, doi: 10.1145/3152421.

[6] B. Friedman, P. Kahn, A. Borning, and A. Huldtgren, "Value sensitive design and information systems.," in Early engagement and new technologies: open up the laboratory, N. Doorn, D. Schuurbiers, I. van de Poel, and M. Gorman, Eds., Springer International Publishing, 2013, pp. 55–95.

[7] G. Fischer, D. Fogli, and A. Piccinno, "Revisiting and Broadening the Meta-Design Framework for End-User Development," in New Perspectives in End-User Development, F. Paternò and V. Wulf, Eds., Cham: Springer International Publishing, 2017, pp. 61–97. doi: 10.1007/978-3-319-60291-2_4.

[8] J. Silberg and J. Manyika, "Notes from the AI frontier: Tackling bias in AI (and in humans)," McKinsey Global Institute, vol. 6, no. 1, pp. 1–31, [Online]. Available: https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Tackling%20bias%20in%20artificial%20intelligence%20and%20in%20humans/MGI-Tackling-bias-in-AI-June-2019.pdf

[9] S. Feuerriegel, M. Dolata, and G. Schwabe, “Fair AI: Challenges and Opportunities,” Bus Inf Syst Eng, vol. 62, no. 4, pp. 379–384, Aug. 2020, doi: 10.1007/s12599-020-00650-3.
[10] B. Friedman, “Value-sensitive design,” Interactions, vol. 6, no. 3, pp. 16–23, 1996, [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/242485.242493
[11] L. Boer and J. Donovan, “Provotypes for participatory innovation,” in Proceedings of the Designing Interactive Systems Conference, Newcastle Upon Tyne United Kingdom: ACM, Jun. 2012, pp. 388–397. doi: 10.1145/2317956.2318014.
[12] P. Mogensen, “Challenging practice: An approach to cooperative analysis,” Doctoral Dissertation, Aarhus university, Denmark, 1994.
[13] M. Utterberg Modén, M. Ponti, J. Lundin and, M. Tallvid, “When fairness is an abstraction: Equity and AI in Swedish compulsory education,” Nordic journal of educational research, In preparation 2024.
[14] “Education Act 2010:800,” Department of education, Stockholm, Sweden, SFS 2010:800, 2010.
[15] T. Bratteteig and G. Verne, “Does AI make PD obsolete?: exploring challenges from artificial intelligence to participatory design,” in Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial - Volume 2, Hasselt and Genk Belgium: ACM, Aug. 2018, pp. 1–5. doi: 10.1145/3210604.3210646.
[16] M. B. Rosson and J. M. Carroll, Usability engineering: scenario-based development of human-computer interaction, 1st ed. in Morgan Kaufmann series in interactive technologies. San Francisco: Academic Press, 2002.
[17] D. Yoo, A. Huldtgren, J. P. Woelfer, D. G. Hendry, and B. Friedman, “A value sensitive action-reflection model: evolving a co-design space with stakeholder and designer prompts,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris France: ACM, Apr. 2013, pp. 419–428. doi: 10.1145/2470654.2470715.
[18] M. Utterberg Modén, M. Tallvid, J. Lundin, and B. Lindström, “Intelligent Tutoring Systems: Why Teachers Abandoned a Technology Aimed at Automating Teaching Processes,” presented at the 54th Hawaii International Conference on System Sciences, 2021, pp. 1538–1547.
[19] N. Selwyn, “The future of AI and education: Some cautionary notes,” Euro J of Education, vol. 57, no. 4, pp. 620–631, Dec. 2022, doi: 10.1111/ejed.12532.
[20] E. Grönvall, L. Malmborg, and J. Messeter, “Negotiation of values as driver in community-based PD,” in Proceedings of the 14th Participatory Design Conference: Full papers - Volume 1, Aarhus Denmark: ACM, Aug. 2016, pp. 41–50. doi: 10.1145/2940299.2940308.
[21] R. M. Jonas and B. V. Hanrahan, “Designing for Shared Values: Exploring Ethical Dilemmas of Conducting Values Inclusive Design Research,” Proc. ACM Hum.-Comput. Interact., vol. 6, no. CSCW2, pp. 1–20, Nov. 2022, doi: 10.1145/3555182.
[22] Y. Engeström, Learning by expanding: an activity-theoretical approach to developmental research, Second edition. New York, NY: Cambridge University Press, 1987.
[23] C. Russo, F. Danioni, I. Zagrean, D. and Barni, D.. Changing personal values through value-manipulation tasks: a systematic literature review based on Schwartz’s theory of basic human values. European Journal of Investigation in Health, Psychology and Education, 12(7), 2022, 692-715.