Trans Resistance to AI Technologies

SAMUEL MAYWORM, University of Michigan, USA

MICHAELANNE THOMAS, University of Michigan, USA

OLIVER L. HAIMSON, University of Michigan, USA

In this workshop paper, we argue that most AI technologies enforce gender binarism and gender surveillance in an inherently transphobic way, preventing these technologies from providing "solutions" that support the goals and well-being of trans people. We argue that transness itself is a form of resistance against AI, and encourage trans people and allies to resist the increased proliferation of transphobic AI technologies in mainstream society. We can, instead, actively subvert AI technologies.

1 Deleting three years of my online trans life as an act of resistance against OpenAI

On May 16 2024, Reddit announced that it would allow OpenAI to "bring Reddit content to ChatGPT and new products" and "access Reddit's Data API, which provides real-time, structured, and unique content from Reddit," allowing OpenAI tools to "better understand and showcase Reddit content, especially on recent topics" [13]. Some Reddit users (including myself[1]) expressed frustration and concern with Reddit and OpenAI's partnership. I primarily used Reddit to post on trans-specific subreddits about issues related to gender transition and navigating the world as a trans person. I felt uncomfortable with the idea that my trans-specific account activity could fuel algorithmic tools that systematically exclude and demonstrate bias toward trans people while policing our genders and gender presentation. Did I really want years of my vulnerable online questions about gender and transition to be fed to a large language model?

In response, I followed the footsteps of other disgruntled Reddit users by mass-overwriting my Reddit post and comment history before deleting all my comments (and later my account) entirely. Simply deleting a Reddit account does not delete the account's comments and posts, which remain publicly visible to other users [12]; additionally, deleted Reddit comments are still stored on Reddit's servers, allowing online content retrieval tools to display the text

---

[1]The Reddit account deletion vignette described throughout this workshop paper refers to the first author's experiences; the use of "I" throughout these sequences refers to the first author.

Authors' Contact Information: Samuel Mayworm, mayworms@umich.edu, University of Michigan, Ann Arbor, Michigan, USA; Michaelanne Thomas, mmtd@umich.edu, University of Michigan, Ann Arbor, Michigan, USA; Oliver L. Haimson, haimson@umich.edu, University of Michigan, Ann Arbor, Michigan, USA.

of deleted Reddit posts. The nature of Reddit content and account deletions fueled speculation among users that even if they deleted their comments, posts, and accounts, their content may still be shared with OpenAI against their wishes.

Particularly disgruntled Reddit users like myself responded to this speculation by mass-editing our post and comment history into strings of random words before deleting our Reddit content and accounts, attempting to sabotage our online comments and posts before they were shared with external parties. I personally used an online tool called PowerDeleteSuite [17] to overwrite my Reddit account's entire comment history with random strings of words before deleting all comments and posts from the account. Once all my scrambled comments and posts were deleted, I glanced through my now-empty account one last time before deleting it from the platform. Three years of my online life – my early years learning about transness, iterating on pronouns and names, meeting other trans people, asking countless first-timer questions about identity, transition, belonging, struggle, and finding hope for a happier future – were overwritten and deleted in under three minutes.

I am aware that my act of "resistance" may have been entirely ineffective, and that my cumulative account history – every gender joke, vent post, and question I posted on multiple trans subreddits for years – may have been added to some generative AI database long before I overwrote it all. No social media user can peek into the opaque relationship between platforms, online generative AI tools, and user data shared from one to the other, let alone know whether their acts of account sabotage actually prevented their data from being used by these services. Still, this act was the closest thing I could envision to "resisting" AI in some way, even at the cost of my own personal history; my little Library of Alexandria burning to the ground, taking years of my own trans life down with it. I miss some of the people I met there. I wonder how they're doing?

2   The incompatibility of transness and the algorithmic operationalization of binary gender

Past work has found that AI systems often act on biased datasets that misrepresent, or fail to represent, marginalized populations [3–5, 11]. Many AI technologies act on gender classification systems that operationalize binary gender dichotomies and norms for gender presentation, putting these technologies at direct odds with transness itself [14, 15]. Not only do these technologies generally perform worse for trans people than cisgender people, but they often perform exceptionally badly for nonbinary people – their databases generally do not include images of nonbinary people, nor can these systems detect whether a person is nonbinary [14, 15].

Trans people have written extensively about the real-world harm they experience when interacting with AI systems that surveil and police gender, such as interactions with TSA scanners that assess "gender" based on physical shape and flag people with "anomalies" as potential security threats [6], or algorithmic content moderation systems that misgender trans people while incorrectly flagging their images as rule-breaking "nudity" [8, 10]. Exclusionary algorithmic systems are a near-unavoidable part of regular life for trans people who use internet technologies, exacerbating the marginalization we already face on a daily basis.

AI technologies' adherence to gender binarism and gender surveillance are inherently at odds with both transness itself and with the goals and well-being of trans people, making the authors of this work skeptical that AI technologies can provide "solutions" for trans people's sociopolitical struggles. We argue that the inherent tension between AI technologies and transness, and the routine ways in which trans people and bodies confound AI technologies, reveals how transness itself can be a form of resistance against AI technologies and their normative values. The problems caused by AI systems' goals and operationalization of gender cannot be solved by the increased "representation" of trans people's data in AI systems' datasets. Instead, we encourage both trans people and trans allies to resist the increased proliferation of transphobic AI technologies in mainstream society, and to subvert such technologies when possible.

We say this while fully aware that no one can truly know the most effective ways to "subvert" opaque algorithmic systems, creating the potential for acts of "resistance" that may not achieve the intended result. We also say this while acknowledging that certain acts of subversion require personal sacrifice, such as Reddit users sacrificing their online histories to prevent their data from being shared with OpenAI.

3 Trans people building AI technologies to resist AI

Some trans people resist transphobic AI technologies by building their own AI technologies. Allucquére Rosanne "Sandy" Stone argues that "in cyberspace, the transgendered [sic] body is the natural body" [16], describing how trans people use technologies to survive, experiment with identity, and challenge mainstream cultural norms surrounding bodies, materiality, and gender expression. Similarly, some trans people have built and used trans-specific AI technologies that subvert and critique mainstream AI technologies and normative attitudes about gender embedded in mainstream AI technologies. Trans artists like Ada Ada Ada [2] use custom-built machine learning tools to create digital art that interrogates gender and online gender surveillance; Ada Ada Ada describes the precautions she takes to limit the environmental impact and privacy implications of her generative AI use, while sharing her commitment to "never use AI without accompanying it [with] a criticism of the technology itself" [1]. Similarly, trans artist Lucas LaRochelle (creator of Queering the Map, a "community generated counter-mapping platform for digitally archiving LGBTQ2IA+ experience in relation to physical space") developed QT.bot, an AI technology that generates hypothetical queer stories based on real Queering the Map entries; LaRochelle's goal for QT.bot is to "revel in the potential of failure, chaos and incommensurability in the queer use of machine learning technologies" by producing "visions of the many futures contained within [Queering the Map's] data... fabulating on the absences of the archive, orienting us away from what is, and towards what could be" [9]. When trans people develop AI-based "trans technologies" [7] for their personal ends, they often create technologies that explicitly subvert and resist mainstream AI technologies' goals, while openly challenging the harm AI technologies inflict upon trans people.

4 Conclusion

I deleted three years of my formative trans life in an act of resistance against OpenAI; I cannot take back this act, nor can I know whether it was an effective act of "resistance." Yet I would make the same decision today even without knowing whether it would meaningfully subvert OpenAI's products, simply because I would rather resist the increasing proliferation of transphobic AI technologies than to passively allow my own data to fuel it, even if that resistance is symbolic. Reflecting on this decision and the circumstances that informed it, the authors of this paper propose the following questions for our workshop:

(1) What does it mean for an act of resistance against AI technologies to be "effective?" And to what extent does the "effectiveness" of an act of AI resistance matter?

(2) How should one envision an act of *trans* resistance against AI technologies? Could they be targeted, explicit acts of sabotage, or symbolic acts of resistance with less guaranteed impact on AI technologies – or both?

References

[1] Ada Ada Ada. 2025. 6 steps for using AI ethically as a conceptual artist. https://www.patreon.com/posts/6-steps-for-ai-122338507

[2] Ada Ada Ada. 2025. Ada Ada Ada. https://ada-ada-ada.art

[3] Nazanin Andalibi, Cassidy Pyle, Kristen Barta, Lu Xian, Abigail Z. Jacobs, and Mark S. Ackerman. 2023. Conceptualizing Algorithmic Stigmatization. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–18. doi:10.1145/3544548.3580970

[4] Cynthia L. Bennett, Cole Gleason, Morgan Klaus Scheuerman, Jeffrey P. Bigham, Anhong Guo, and Alexandra To. 2021. "It's Complicated": Negotiating Accessibility and (Mis)Representation in Image Descriptions of Race, Gender, and Disability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–19. doi:10.1145/3411764.3445498

[5] Clementine Collett and Sarah Dillon. 2019. *AI and Gender: Four Proposals for Future Research.* Technical Report. Apollo - University of Cambridge Repository. doi:10.17863/CAM.41459

[6] Sasha Costanza-Chock. 2018. Design Justice, A.I., and Escape from the Matrix of Domination. *Journal of Design and Science* (July 2018). doi:10.21428/96c8d426

[7] Oliver L. Haimson. 2025. *Trans technologies.* The MIT Press, Cambridge, Massachusetts.

[8] Oliver L. Haimson, Daniel Delmonaco, Peipei Nie, and Andrea Wegner. 2021. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–35. doi:10.1145/3479610

[9] Lucas LaRochelle. 2021. Queering the Map: On designing digital queer space. In *Queer sites in global contexts: technologies, spaces, and otherness.* Routledge, Abingdon New York, NY, 133–147.

[10] Samuel Mayworm, Kendra Albert, and Oliver L. Haimson. 2024. Misgendered During Moderation: How Transgender Bodies Make Visible Cisnormative Content Moderation Policies and Enforcement in a Meta Oversight Board Case. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency.* ACM, Rio de Janeiro Brazil, 301–312. doi:10.1145/3630106.3658907

[11] Jennifer Mickel. 2024. Racial/Ethnic Categories in AI and Algorithmic Fairness: Why They Matter and What They Represent. doi:10.48550/ARXIV.2404.06717 Version Number: 1.

[12] Reddit. 2024. How do I delete my account? https://support.reddithelp.com/hc/en-us/articles/204579509-How-do-I-delete-my-account

[13] Reddit. 2024. Reddit and OpenAI Build Partnership. https://redditinc.com/blog/reddit-and-oai-partner

[14] Morgan Klaus Scheuerman, Jacob M. Paul, and Jed R. Brubaker. 2019. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–33. doi:10.1145/3359246

[15] Morgan Klaus Scheuerman, Kandrea Wade, Caitlin Lustig, and Jed R. Brubaker. 2020. How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 1–35. doi:10.1145/3392866

[16] Allucquère Rosanne Stone. 2001. *The war of desire and technology at the close of the mechanical age* (4. print ed.). MIT Press, Cambridge, Mass.

[17] Ronald Troyer. 2018. PowerDeleteSuite. https://github.com/j0be/PowerDeleteSuite