Challenges in Designing Racially Inclusive Language Technologies

Kimi V. Wenzel
Carnegie Mellon University
5000 Forbes Ave,
Pittsburgh PA 15213
kwenzel@cs.cmu.edu

Geoff Kaufman
Carnegie Mellon University
5000 Forbes Ave,
Pittsburgh PA 15213
gfk@cs.cmu.edu

Abstract

We apply a critical lens toward the pursuit of racially inclusive language technologies and identify several areas necessitating future work. Our discussion focuses on identifying the potential harms of conversational technologies, outlining three challenges that arise in inclusive design, and lastly, arguing that conversational user interface designers and researchers should go beyond racially inclusive design and strive for anti-racist and race positive designs.

Author Keywords

Inclusion; Decolonization; Anti-racism; Race positive design;

CCS Concepts

•**Social and professional topics** → **Race and ethnicity;**
•**Human-centered computing** → *HCI theory, concepts and models;*

Introduction

Attention to "ethical" and "inclusive" developments in language technology has increased dramatically in recent years. A great deal of work has been devoted to understanding how racial bias can emerge in language model production, performance, and accuracy [4]; however, work aimed at understanding how conversational user interfaces (CUIs) may themselves perpetuate racism is limited. In this

paper, we introduce the potential harms caused by racism in CUIs, discuss three broad challenges of inclusive design, and conclude by urging CUI researchers to invest in anti-racist and race positive design research.

Harms of Non-Inclusive CUIs

Our recent work has revealed that miscommunication errors with conversational technologies can mirror that same effects as racist microaggressions in interpersonal interactions: A high speech recognition error rate can significantly increase Black users' levels of self-consciousness and significantly decrease their self-esteem and emotional affect. This effect was not found among white users [15]. Thus, CUIs that are not inclusive of marginalized dialects and vernaculars may pose not only a usability threat to these users but a psychological threat as well. These findings only increase the urgency for language technologists to create racially just CUIs.

Three Challenges of Inclusive Design

We outline three primary ethical challenges in designing racially inclusive conversational user interfaces: (1) Collecting representative voice and language data in a non-extractive, decolonial manner; (2) Understanding the CUI utility-surveillance trade off for vulnerable populations; and (3) Contributing to unmediated anti-racism.

Collecting representative voice and language data
The high error rate in conversational technologies for marginalized users has largely been attributed to lack of representative voice and language data [9]. Thus, the direct solution to minimizing this error rate is to gather voice data from those underrepresented groups. The methods for achieving this solution, however, are conventionally colonial and harmful [10, 2]. Many researchers have contributed frameworks and guidelines for non-extractive data collection [7];

however, further investigation is needed to understand how these practices may uphold specifically in data collection for language models and conversational user interfaces. How might we create a representative language data set for conversational technologies while being respectful, non-extractive, and non-exploitative of the marginalized groups from whom we seek data from? Perhaps more importantly, in line with community-based participatory research practices [6], we should not ask how we may seek data *from* these groups, but rather, ask how we may seek data *with* these groups. This further necessitates an understanding of whether or not these technologies are of significant value to marginalized groups at all. Even if such technologies appear to be of significance on a superficial level, utility-surveillance tradeoffs must also be considered and analyzed.

Understanding surveillance risks
Current research follows a rhetoric of techno-solutionism [11], assuming that a positive outcome is inherent in racially inclusive technology. But how much would conversational technologies, even in the absence of their racial disparities, cause harm? There are various concerns regarding how emerging technologies can serve as tools to target and surveil marginalized groups [5]. While there are many privacy researchers focusing specifically on conversational devices [13] and CUI researchers working on inclusive design [14], these two lines of work are often divorced. When designing CUIs for marginalized groups, how might we consider their unique privacy needs? In response, we urge the CUI community to consider increased collaborations with privacy and security researchers.

Contributing to unmediated anti-racism
Framing the issue of inclusion as merely a dataset problem also ignores the structural and historical roots of dis-

crimination, and fails to acknowledge the larger web of discrimination in which technology-mediated racism exists. Acknowledging this larger web of discrimination necessitates our understanding that CUIs are just one medium for present-day discrimination, and that mediums of discrimination have, and will continue to, change over time. While these mediums of discrimination may evolve, however, the targets of discrimination seldom do. Thus, we should not limit our focus solely to technology-mediated interactions. Prioritizing research on technology-mediated forms of discrimination often de-prioritizes unmediated forms of discrimination [12], ultimately implicitly suggesting that the *form* of discrimination matters more than the marginalized groups and individuals themselves. To this end, we suggest CUI scholars and practitioners take care to reflect on how their work may contribute to anti-discrimination in unmediated contexts. This suggestion is in line with broader efforts to go beyond inclusivity, toward anti-racism.

Beyond Inclusion: Toward Anti-Racism and Race Positive Design

"Inclusion" first and foremost centers the status-quo of privileged groups. "Inclusion" suggests that a privileged status-quo exists, and that previously excluded groups should be accepted to the said status-quo. Instead of simply aiming to offer marginalized groups the same opportunities as privileged groups through "inclusion," we should (1) actively aim to dismantle the structures that caused this power imbalance through anti-racist practices and (2) center and celebrate silenced and oppressed identities through race positive design.

Anti-racist design
As Ibram X. Kendi writes, "the only way to undo what is racist is to consistently identify and describe it – and then dismantle it" [8]. Dismantling this structure requires compre-

hensive anti-racist education, and a reorienting of research and design work. How might we shift from critiquing, writing about, and discussing racism toward actively producing change in both our scholarly and local communities? (See [1] and [8] for further articulation and case studies.)

Race positive design
Race positive design "enables thinking about race as a positive presence—as cultural capital, histories of resistance, bindings between lands and peoples—and the means by which it can be a generative force in technologies for just and sustainable futures" [3]. Recent work suggests that race positive identity affirmations may be a potential method for addressing the harms that discriminatory conversational technologies inflict on people of color; however, more work needs to be done to validate these design interventions [15]. Future work may consider: How might CUIs serve as a medium for race and culture to be celebrated? How might race and culture serve as a generative feature of language technologies, uplifting marginalized groups?

Conclusion

We outline three challenges in designing for inclusive conversational user interfaces, and provide suggestions for how researchers may begin to address these challenges: (1) Researchers should practice non-extractive decolonial data collection, and consider what legitimate value their data collection efforts may bring to the communities they are working with. (2) CUI designers and researchers should engage further in privacy and security work, especially as surveillance through CUIs may be especially harmful for marginalized populations. (3) Researchers should reflect on how their research on inclusivity in technology-mediated tools may contribute to broader unmediated anti-racism efforts. In addition to appreciating these challenges, we encourage researchers to consider race-positive CUI designs

in their future work.

REFERENCES

[1] Veronica Abebe, Gagik Amaryan, Marina Beshai, Ali Ekin Gurgen, Wendy Ho, Naaji R Hylton, Daniel Kim, Christy Lee, Carina Lewandowski, Katherine T Miller, and others. 2022. Anti-Racist HCI: notes on an emerging critical technical practice. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–12.

[2] Bagele Chilisa. 2019. *Indigenous research methodologies*. Sage publications.

[3] Ron Eglash, Audrey Bennett, Michael Lachney, and William Babbitt. 2020. Race-positive design: A generative approach to decolonizing computing. In *Human factors in computing systems*.

[4] Anjalie Field, Su Lin Blodgett, Zeerak Waseem, and Yulia Tsvetkov. 2021. A survey of race, racism, and anti-racism in NLP. *arXiv preprint arXiv:2106.11410* (2021).

[5] Seeta Peña Gangadharan. 2017. The downside of digital inclusion: Expectations and experiences of privacy and surveillance among marginal Internet users. *New Media & Society* 19, 4 (2017), 597–615.

[6] Karen Hacker. 2013. *Community-based participatory research*. Sage publications.

[7] Paul Agu Igwe, Nnamdi O Madichie, and David Gamariel Rugara. 2022. Decolonising research approaches towards non-extractive research. *Qualitative Market Research: An International Journal* ahead-of-print (2022).

[8] Ibram X Kendi. 2023. *How to be an antiracist*. One world.

[9] Allison Koenecke, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R Rickford, Dan Jurafsky, and Sharad Goel. 2020. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences* 117, 14 (2020), 7684–7689.

[10] Sandra Kouritzin and Satoru Nakagawa. 2018. Toward a non-extractive research ethics for transcultural, translingual research: perspectives from the coloniser and the colonised. *Journal of Multilingual and Multicultural Development* 39, 8 (2018), 675–687.

[11] Evgeny Morozov. 2013. *To save everything, click here: The folly of technological solutionism*. Public Affairs.

[12] Seeta Peña Gangadharan and Jędrzej Niklas. 2019. Decentering technology in discourse on discrimination. *Information, Communication & Society* 22, 7 (2019), 882–899.

[13] Anne Pfeifle. 2018. Alexa, what should we do about privacy: Protecting privacy for users of voice-activated devices. *Wash. L. Rev.* 93 (2018), 421.

[14] Jaisie Sin, Cosmin Munteanu, Jenny Waycott, Robin N. Brewer, Sergio Sayago, Amanda Lazar, and Astrid Weber. 2022. Alexa, Tell Me a Joke!:" Voice Interfaces are Truly Inclusive". In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–3.

[15] Kimi Wenzel, Nitya Devireddy, Cam Davidson, and Geoff Kaufman. 2023. Can Voice Assistants Be Microaggressors? Cross-Race Psychological Responses to Failures of Automatic Speech Recognition. *arXiv preprint arXiv:2302.12326* (2023).