Computer Science Faculty Research and Publications

Computer Science, Department of

10-27-2020

Autobiography of an Audit: Tracing the Roots and Repercussions of the HRT-Transgender Database

Os Keyes

Jeanie Austin

Michael Zimmer

AUTOBIOGRAPHY OF AN AUDIT: TRACING THE ROOTS AND REPERCUSSIONS OF THE HRT-TRANSGENDER DATABASE

Os Keyes
University of Washington

Jeanie Austin
San Francisco Public Library

Michael Zimmer
Marquette University

Introduction and Overview

Much concern has been expressed regarding the difficulty in attending to harms in data collection, distribution, and use in sociotechnical systems (Seaver, 2017; Kitchin, 2016). Arguments around both internal and external audit mechanisms frequently focus on the work of private-sector organisations, which are often seen as responsible for the increasing role of data collection and analytic systems in society (Kerssens, 2019), as well as the site of particular practical barriers to data auditing due to concepts such as trade secrets.

This approach tends to reify the idea that research conducted by universities and other public-sector parties is both more ethical and more easily lends itself to auditing. In reality, the possibility for scrutiny of research utilizing online content to create databases and to train algorithmic prediction is heavily reduced by limited IRB oversight and researcher's guarded approach to their own data collection and sharing practices. Further, approaches to audits rarely examine the practices of identifying and addressing *discursive* harms (Hoffmann, 2019; Dencik, Jansen & Metcalfe, 2018) and the affective work and cost of undertaking audits of violent systems.

Documenting our attempts to audit the HRT-Transgender Database - a database collected by a public university in the United States - we engage in a critical examination

of not only data collection practices, but the practical and phenomenological experience of auditing violent research. Doing so raises vital questions that those seeking to design audit mechanisms should attend to.

Background

Transgender people often face difficulties in accessing reliable information about navigating the medicolegal systems that gatekeep access to hormones, surgery and other biomedical technologies used in the process of physical transition (Pohjanen and Kortelainen, 2016). Social media communities, particularly those that utilize video recording, have become one of the primary means by which transgender people share vital information that can ease the strain of accessing needed medications and therapies while building a sense of community and social support (Haimson, Dame-Griff, Capello, & Richter, 2019; Horak, 2014).

Repurposing and appropriating these community resources and personal narratives, researchers at the University of North Carolina, Wilmington developed the "HRT-Transgender Database", a collection of transition progress videos scraped from YouTube and acquired for use in the design of facial recognition systems (Mahalingam & Ricanek, 2013). This database was then made available to other researchers through the project website. Several years after its creation, the database and its authors attracted (largely-negative) media coverage for their work, and eventually shut the project down (Vincent, 2017)

On the surface, this controversy centers on issues that are well-covered in critical data studies and internet research ethics; the problem of treating data as "already public" and the insufficiency of existing accountability mechanisms around algorithmic systems (Zimmer, 2010; Neyland, 2016). Our interest in it, however, stems from our desire to expand on research addressing two other aspects of data's negative effects and efforts to ameliorate them; the *discursive harms* of data and the *fluid nature* of database distribution. By the former, we mean the way that researchers frequently not only repurpose data, but resituate it in cultural contexts where it can be violently deployed against the creator - something that, as Anna Lauren Hoffmann notes, is undertheorised in work around algorithmic ethics (Hoffmann, 2019). By the fluid nature of distribution, we are referencing the way that database repurposing is commonly done across organisational and disciplinary boundaries, undermining accountability mechanisms that assume a singular entity to audit (be that "internally" or "externally") (Raji *et al.,* 2020). In order to build on this work, we document and narrate our efforts to audit the HRT-Transgender Database and its secondary reusage. Drawing on feminist holistic reflexivity (Cooky, Linabery, Corple, 2018), we not only engage in a critical examination of the practical difficulties in auditing a (theoretically) publicly-accessible data collection process, but grapple with our own emotive experiences of auditing a violent, trans-focused database as a (predominantly trans) research team, and surface the discursive violence central to how the database is framed, narrated and reused. Our work brings into frame vital and critical issues that researchers seeking to design oversight mechanisms should address, and begins a conversation about the visceral and often painful work of *providing* that oversight.

Conclusion

Our experience of auditing the creation, distribution and reuse of the HRT-Transgender Database provides a vital case study in expanding on existing work on algorithmic accountability and data ethics. Complicating the traditional image of private-sector organisations causing direct material harms due to (for example) issues of database bias, we demonstrate the discursive harms and obfuscatory practices involved in even a supposedly "public" research project. Further, we attend to the cost borne by researchers undertaking audits of violent systems - audits that frequently require close encounters with that violence. We argue from this that conversations around algorithmic accountability must take a more systemic and reflective view, recognising not only idealised accountability processes but the complex and very human failures contained within them in practice, and the affective cost that (frequently, marginalised) auditors take on by seeking to surface harms in this space.

References

Cooky, C., Linabery, J. R., & Corple, D. J. (2018). Navigating Big Data dilemmas: Feminist holistic reflexivity in social media research. *Big Data & Society*.

Dencik, L., Jansen, F., & Metcalfe, P. (2018, August 30). A conceptual framework for approaching social justice in an age of datafication. DATAJUSTICE project. Retrieved from https://datajusticeproject.net/2018/08/30/a-conceptual-framework-for-approaching-social-justice-in-an-age-of-datafication

Haimson, O.L., Dame-Griff, A., Capello, E. & Richter, Z. (2019). Tumblr was a trans technology: the meaning, importance, history, and future of trans technologies, *Feminist Media Studies*. DOI: 10.1080/14680777.2019.1678505

Hoffmann, Anna Lauren. (2019). "Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse." *Information, Communication & Society* 22.7: 900-915.

Horak, L. (2014). Trans on YouTube: Intimacy, visibility, temporality. *TSQ: Transgender Studies Quarterly*, 1(4), 572 - 585.

Kerssens, Niels. (2019). "De-Agentializing Data Practices: The Shifting Power of Metaphor in 1990s Discourses on Data Mining." *Journal of Cultural Analytics*: 1-26.

Kitchin, Rob. (2017). "Thinking critically about and researching algorithms." *Information, Communication & Society* 20.1: 14-29.

Mahalingam, Gayathri, and Karl Ricanek. (2013). "Is the eye region more reliable than the face? A preliminary study of face-based recognition on a transgender dataset." *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*..

Neyland, Daniel. (2016). "Bearing account-able witness to the ethical algorithmic system." *Science, Technology, & Human Values* 41.1: 50-76.

Seaver, Nick. (2017) "Algorithms as culture: Some tactics for the ethnography of algorithmic systems." *Big Data & Society* 4.2.

Pohjanen, A. M., and Kortelainen, T. A. M. (2016). Transgender information behavior. *Journal of Documentation*, 72 (1), 172-190.

Raji, Inioluwa Deborah, et al. (2020). "Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing". In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 33–44.

Vincent, James (2017). "Transgender YouTubers had their videos grabbed to train facial recognition software"*. The Verge.* Retrieved from https://www.theverge.com/2017/8/22/16180080/transgender-youtubers-ai-facial-recognition-dataset

Zimmer, Michael. (2010). ""But the data is already public": on the ethics of research in Facebook." *Ethics and information technology* 12.4: 313-325.