

Systemic racism in data practices

Trevor Watkins¹

Jonathan O. Cain²

Positionality statement

As we begin to discuss this issue, its origins, and its importance in contemporary society, I wanted to acknowledge my positionality and the role that it may play in the formation of this issue. Jonathan O. Cain is an African-American male working in the LIS field. Before moving into administration, I taught data and digital literacy and worked on developing programs that focused on improving access to these critical skills at zero cost to learners.

It is important to acknowledge my positionality and the lens through which I see the data science field. Trevor Watkins is an African American male working in the LIS field at an academic institution in an academic library. I teach critical data literacy workshops and engage in diversity and BIPOC-related digital projects with faculty, students, and the broader academic community across the country. I am also a researcher and practitioner in artificial intelligence (AI) and data science.

The global pandemic, its impacts, and why it matters

We first met in August 2020 to discuss the possibilities of this special issue about five months into the pandemic. We spent a good chunk of that meeting getting to know each other and, most importantly, discussed the toll the pandemic placed on our communities and us. It is probably safe to say that many of you, at some point, were uncertain of the future. Like most people worldwide, we lost family and friends or knew of people who succumbed to Covid-19 and other illnesses that weren't treated because the focus shifted to Covid-19. We get it. At one point, Covid-19 killed over three thousand people per day (Centers for Disease Control and Prevention (CDC), 2022). According to data from the CDC, 90% of the 385,676 people who died between March and December 2020 had Covid-19 listed as the underlying cause of death on their death certificate. The murders of Ahmaud Arbery in February, Breonna Taylor in March, and George Floyd in May 2020 sparked civic unrest across the United States (US) and protests across the globe in solidarity against racial injustice. When we announced this special issue and initiated a call for papers, we didn't get much of a response initially. We expected and acknowledged that it would probably take some time before we received inquiries or proposals about the issue, the intent to submit, or any submissions.

Like many of you, we are still picking up the pieces from 2020 and dealing with the aftermath of Covid-19. The pandemic may be over now, depending on whom you ask, but the emotional scars are still there and may remain so for quite some time. Patience was the one quality we all had throughout this process, which is why we can present this publication today.

Data and liberatory technology

Liberatory technology. This is a concept that invited contemplation as we sat down to record our reflections on this special issue. In drawing together scholars, educators, and practitioners to address the issue of data and its relationship to race, ethnicity, and representation, we, as coeditors, were making a statement about the importance of data, the material impact that this seemingly abstract and ethereal object can and does have on individual and community lives. And thinking about that

impact brought liberatory technology to the front of our minds. The definition of liberator technology offered by the IDA B. Wells Just Data Lab intrigues us and invites us to grapple with that topic. They defined liberatory as something that "supports the increased freedom and wellbeing of marginalized people, especially black people outside of capitalism and settler colonial power structures" and technology as "a tool used to accomplish a task." And as we contemplate this set of definitions, we are left to question whether data can be a liberatory technology or not. (LIBERATORY TECHNOLOGY AND DIGITAL MARRONAGE, n.d.)

In *Liberation Technology: Black Protest in the Age of Franklin*, Richard S. Newman draws parallels with the asserting ownership and mastery of new communication technologies and black liberation activities. Reflecting on the transformative nature of print technology, he writes, "If the Marquis de Condorcet was right in 1793 that print had unshackled Europe from medieval modes of thought and action, then it is also true that print was perhaps the first technology to liberate blacks from the servile images that had long haunted their existence in Western culture." And draws a 19th-century example of how it expressly connects to black lives post-emancipation noting "W. E. B. Du Bois certainly thought that black history and print history worked in tandem. Wherever one found newspapers in the post-Civil War South, he observed, one found some form of black freedom" (Richard S. Newman, 2009, p. 175). He even notes how scholars note that black activists embraced other communication technologies like photography "to reshape the image of African Americans in nineteenth-century culture." (Richard S. Newman, 2009, p. 175)

We have no shortage of examples of how data and data-driven technologies fail to support the "increased freedom and wellbeing of marginalized people outside of capitalism and settler colonial power structures." In 2016, ProPublica published *Machine Bias*, a report that looks at Risk assessment technologies used in arraignment and sentencing. They report that "The formula was particularly likely to falsely flag black defendants as future, wrongly labeling them this way at almost twice the rate as white defendants" and "white defendants were mislabeled as low risk more often than black defendants" (Julia Angwin, 2016). A 2021 article, *Fairness in Criminal Justice Risk Assessments: The State of the Art*, in their analysis, noted, "The false negative rate is much higher for whites so that violent white offenders are more likely than violent black offenders to be incorrectly classified as nonviolent. The false positive rate is much higher for blacks so that nonviolent black offenders are more likely than nonviolent white offenders to be incorrectly classified as violent. Both error rates mistakenly inflate the relative representation of blacks predicted to be violent. Such differences can support claims of racial injustice. In this application, the trade-off between two different kinds of fairness has real bite." (Berk et al., 2021, p. 33)

These are just a few examples of how these technological developments, on their own merits, fail to meet the definition offered by the authors of the "Liberatory Technology and Digital Marronage" Zine from the Ida B. Wells Just Data Labs. Reflecting on the technological path illustrated by Newman, the work of ownership and mastery of the tool provides the potential for it to be liberatory. Through this lens, the work of the Just Data Lab is exemplary for this meditation; it draws a direct line from technology, education, mastery, and liberatory technology.

Data in higher education

Data literacy education is an area that has been a focus of our careers in librarianship. It's a space where we saw the libraries' ability to make a meaningful impact. Data has had a tremendous impact on college campuses, from how research is conducted to the pressures colleges feel from stakeholder groups: students, governments, funders, donors, and employers to prepare students with the data and technology skills to gain employment in the knowledge economy.

As colleges and universities have turned (with varying degrees of success) to meet the needs of these communities, a myriad of explorations on the importance of the representation of these marginalized communities in these systems—to combat and dismantle the harmful practices that we see embedded in the systems that drive society and the potentially debilitating consequences they produce. That is partly why the works in this special issue are so important at this moment in time. These scholars and scholar-practitioners are engaging with these issues that drive the opaque structures surrounding us. And hopefully, their work can give us another perspective on how to engage with these structures and transform them to support liberatory practices.

The entries in this issue

We have some fantastic articles for you to read in this issue. We open with an article by Kevin Manuel, Rosa Orlandini, and Alexandra Cooper, who discuss how the collection process of racial, ethnic, and indigenous data has evolved in the Canadian Census since 1871, the erasure of minorities and indigenous citizens from those censuses, and the work to restore and accurately identify and categorize racialized groups.

In the next article, Leigh Phan, Stephanie Labou, Erin Foster, and Ibraheem Ali present a model for data ethics instruction for non-experts by designing and implementing two data ethics workshops. They make important points about the failure of academia to incorporate the ethical use of data in course curriculums and digital literacy training and demonstrate how academic libraries have become an essential resource for the academic community. Their workshop structure can be modeled for any academic library that endeavors to provide a similar service to its community.

In the third article, Natasha Johnson, Megan Sapp Nelson, and Katherine Yngve, interrogate the collective and local purposes of institutional data collection and its impact on student belongingness and propose a framework based on data feminism that centers the student as a person rather than a commodity.

Finally, our closing article from Thema Monroe-White focuses on marginalized and underrepresented people in the data science field. The author proposes that racially relevant and responsive teaching is necessary to recruit more people from these groups and diversify the field. She discusses how the Ladson-Billings model of cultural relevant pedagogy has been applied and is beneficial to STEM curriculums, and how a liberatory data science curriculum could promote a student's voice and sense of belonging.

Conclusion

We want to thank all those involved in producing this special issue. We want to thank the authors first. Their patience, dedication, and perseverance throughout this process were much appreciated. The

reviewers provided timely, very detailed, and thorough feedback. We would be remised if we didn't acknowledge their hard work and labor. We would like to thank the IQ Editorial Team, Michele Hayslett and Karsten Boye Rasmussen, for working with us over the last two years, and Ofira Schwartz-Soicher, for helping us get to the finish line.

Trevor Watkins
Jonathan O. Cain

References

- Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2021). Fairness in Criminal Justice Risk Assessments: The State of the Art. *Sociological Methods & Research*, 50(1), 3–44.
<https://doi.org/10.1177/0049124118782533>
- Flipsnack. (n.d.). *Liberatory Technology Zine*. Flipsnack. Retrieved December 17, 2022, from <https://www.flipsnack.com/EBC8CD77C6F/liberatory-technology-zine.html>
- LIBERATORY TECHNOLOGY AND DIGITAL MARRONAGE. (n.d.). IDA B. WELLS JUST DATA LAB. Retrieved December 17, 2022, from <https://www.thejustdatalab.com/tools-1/liberatory-technology-and-digital-marronage>
- Mattu, J. A., Jeff Larson, Lauren Kirchner, Surya. (n.d.). *Machine Bias*. ProPublica. Retrieved December 17, 2022, from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Richard S. Newman. (2009). Liberation Technology: Black Printed Protest in the Age of Franklin. *Early American Studies: An Interdisciplinary Journal*, 8(1), 173–198.
<https://doi.org/10.1353/eam.0.0033>

Endnotes

¹ Trevor Watkins is the Teaching and Outreach Librarian at George Mason University Libraries. He can be reached at twatkin8@gmu.edu.

² Jonathan O. Cain is the Associate University Librarian for Research and Learning at Columbia University Libraries and can be reached at joc2122@columbia.edu.