Use of New Artificial Intelligence Technologies Policy Submission

My name is Madelin Burt-D'Agnillo. I am a first-year student at the University of Toronto's Faculty of Information, concentrating on Critical Information Policy Studies. My classmates and I have been studying, among other topics, issues of AI accountability, racial profiling in AI systems, and surveillance infrastructures. While the TPS's drafted AI Policy gestures toward considerations of racial profiling and other forms of harm exacted on marginalized communities because of AI, I think that it is grossly inadequate to address the potential of encoding inequities into machine learning algorithms.

My feedback reflects on Ruha Benjamin's work on race and AI, specifically two introductory chapters to books she authored and edited in 2019. I am choosing to focus on Benjamin's contribution to this field because she skillfully navigates the rhetorical dichotomy that can often polarize the conversation about AI: it is neither something that will save us, nor something that can slay us. By inviting a "race critical code studies analysis," Benjamin situates the current impacts of AI in a historical and social context (Benjamin, 2019a). My response will comment on three aspects of this proposed policy—the guiding principles, the risk categories, and the benefit analysis—and suggest, as Benjamin does, that "technology is not just a bystander that happens to be at the scene of the crime; it actually aids and abets the process by which carcerality penetrates social life" (Benjamin, 2019b, p. 2).

First, Benjamin traces the current racist outcomes of AI technologies to the history of slavery, Jim Crow-era politics, and racial discrimination in the United States. As an evolution of these histories, she defines the "New Jim Code" as the use of new technologies that reproduce existing social inequities; all of which are advertised as, and appear to be, neutral, objective, or even progressive (Benjamin, 2019a, p. 3). The New Jim Code is a set of practices, "some that

explicitly work to amplify hierarchies, many that ignore and thus replicate social divisions, and a number that aim to fix racial bias but end up doing the opposite" (Benjamin, 2019a, p. 4). This range of motivations and outcomes shows that the technology itself is a metaphorical blank canvas, onto which human actors display their values, ignorance, and biases. This nuanced perspective is paramount when evaluating the TPS's Policy, which is rich with passive, neutral, and evasive language.

Guiding Principle of the TPS Policy

In Benjamin's remarks at the Eighth International Conference on Learning Representations, she leads with a quote from Dr. Martin Luther King Jr., who said: "we have guided missiles and misguided men" (Benjamin, 2020). This quotation highlights the agency of humans to guide technologies, firmly placing the onus and the agency on humans to manage our own infrastructures. By contrast, the guiding principles of the TPS Policy utilizes passive language, stating: "there have been instances in which novel technologies were shown to incorporate and perpetuate preexisting and systemic biases, resulting in both individually and systemically discriminating decisions" (Toronto Police Services Board, 2021, p. 1). This passive construction decenters the role of the human in designing, encoding, and using these technologies; it presumes that they are "insulated from human influence" (Benjamin, 2019b, p. 3). Indeed, Benjamin calls this passive deference the "animating force" of the New Jim Code, whereby tech designers claim that the biases that they encoded into the technical systems are somehow entirely exterior to the encoding process (Benjamin, 2019b, p. 6). Thus, this single sentence in the TPS Policy projects the way that the Service might practice "digital denial" if their systems do produce individual and systemic discriminations in the future.

Risk Categories of the TPS Policy

The meat of the TPS Policy is a five-part risk categorization framework, which will determine how technologies will be decided on, deployed, and reported. While there is a great deal of scrutiny that these risk levels should be subjected to, my focus will be problematizing the notion of risk altogether. For instance, the "high risk category" of technologies includes applications which use data that may "be of poor quality, carry bias, or where the quality of such data is unknown," which can be "influenced or biased by malicious actors" (Toronto Police Services Board, 2021, p. 5). "Low risk" technologies, on the other hand, assist the TPS in "identifying, categorizing, prioritizing or otherwise making decisions pertaining to members of the public" (Toronto Police Services Board, 2021, p. 5). The subtle suggestion embedded in this logic is that there is less danger if the police are the ones who are making the decisions, as opposed to decisions made based on bad data and influenced by malicious actors. This minimizes and flattens the bias of individual officers, which has historically resulted in racial discrimination and violence against marginalized bodies, and the ways that AI technology could deepen inequities.

"Benefits" and the Human Calculus

Last, I think it is worth looking at some of the rationale that the TPS offers for this Policy. First, the financial angle: the TPS believes it is worth investing in new AI technology to provide "efficiencies and cost savings" (Toronto Police Services Board, 2021). Benjamin writes that the New Jim Code is also a symptom of capitalism and a push toward privatization, where efforts to cut costs and maximize profits often happen at the expense of other human needs (Benjamin, 2019a, p. 16). This transition to computational approaches is touted as not only good but necessary. But, as Benjamin warns, the "outsourcing of human decisions is, at once, the

insourcing of coded inequity" (Benjamin, 2019a, p.16). Second, the public safety angle: the undercurrent of this Policy is that in a risk/benefit analysis of AI in policing, the risk of AI policies is lower than the benefit to community safety. But this raises a necessary question: "safety for whom?" Or, as Benjamin postulates: "could it be that we don't need technocorrections to make us secure, that we need social insecurity to justify technocorrections?" (Benjamin, 2019b, p. 2). In other words, it is possible that the introduction of AI technologies into municipal policing services could create new harms that would justify new carceral systems.

Conclusion

In all, Benjamin's work provides an important entry point into the intersections of race and AI. As a final reflection, I will draw attention to my own positionality to this Policy and its implications. As a white, middle-class person with abolitionist values, it is important to me that I recognize my own stake in these carceral systems. While it is true that I could suffer the unintentional consequences of algorithmic policing, it is much more likely that the material, social, and lived consequences will impact marginalized communities and deepen existing inequities. Benjamin invites her white students to reflect on the way that their whiteness affords them invisibility and offers them immunity: "to be unmarked by race allows you to reap the benefits but escape responsibility for your role in an unjust system" (Benjamin, 2019a, p. 2). Plainly, I do not want to live in a community in which my safety comes at the expense of my neighbours'. In all, I ask that the Toronto Police Services Board reconsider the use of AI technologies altogether, and at a minimum put a stop to all procurement and deployment decisions until meaningful consultation with community members occurs.

References

Benjamin, R. (2019a). Introduction. In *Race after technology: Abolitionist tools for the new Jim code* (pp. 1–32). Polity Press. http://ebookcentral.proquest.com/lib/utoronto/detail.action?docID=5820427

Benjamin, R. (Ed.). (2019b). Introduction. In *Captivating technology: Race, carceral technoscience, and liberatory imagination in everyday life* (pp. 1–22). Duke University Press. https://doi.org/10.1215/9781478004493

Benjamin, R. (2020, April 28). *ICLR: 2020 vision: Reimagining the default settings of technology & society*. https://iclr.cc/virtual_2020/speaker_3.html

Toronto Police Services Board. (2021). *Artificial intelligence policy consultation*. https://tpsb.ca/consultations-and-publications/artificial-intelligence-policy-consultation

Toronto Police Services Board. (2021). *Use of artificial intelligence technology*. https://tpsb.ca/images/consultations/AI/Use_of_AI_Technology_Policy_-_DRAFT_-_20211108_Public_Consultation.pdf