

Dylan Lopez
WRT 205
Prof. Oakes
August 4, 2021

Annotation: The seductive diversion of “solving” bias in artificial intelligence (Powles & Nissenbaum, 2018).

Julia Powles and Hannah Nissenbaum don’t hold anything back when they dissect the bias problems of artificial intelligence. The authors are professors of technology, law and policy and information science, with research interests in online privacy and civic rights-based responses to emerging technologies. In particular, the speed at which AI is taking over the functions of human resources is of great interest to them. They understand that technology companies hold out a “golden promise” that AI is an unstoppable force that will solve our biggest problems. But they also point out that these companies are not so open about the bias problems that are being perpetuated by their technologies. These are the same systems that place men into higher-paying jobs, use facial recognition systems that fail people of color or of different nationalities, and use tests that filter out women and the disabled. The authors note that tech companies argue in defense that their technology will get better over time and that bias will eventually be engineered out of their systems altogether. This sounds like a hopeful future, but the authors point out that trying to solve bias using upstream engineering may have unintended social consequences, making bias worse downstream. For example, if AI systems are designed to better recognize gender and diversity, it may actually make these systems more powerful as a tool for discrimination. In other words, handling bias like a glitch to be fixed technically doesn’t get to the heart of the problem that bias is a social problem that needs a more holistic approach to fix. The bottom line is that despite their promise, systems driven by data from our world will replicate and amplify the racial, gender, and class inequality that exists in our world.

This source demonstrates the complexities of trying to use technology to solve a human problem like bias, and I will use it to highlight the challenges that our society will face as we wrestle with the wider use of AI for human resources. One of my ideas that this article will help me support is the false idea that people building technology are somehow smarter than the rest of us or that they know what is best for us. Just think of the plastic in the ocean and carbon in the atmosphere as a couple examples of the unintended byproducts of technological innovations. I will also use this article to highlight why we can’t look at the AI in HR problem in isolation. We live in a world where data about us is constantly being gathered by many different companies. What will happen when all this data converges so that our personality tests, credit scores, social media posts, social contacts, grades, driving record, browsing history, location history, etc. are available to our future employers? That future is what concerns me, and I think we are going to run out of time unless we establish some rules soon.

Powles, J., & Nissenbaum, H. (2018, December 7). *The seductive diversion of “solving” bias in artificial intelligence*. Medium. <https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>.