The dark side of AI: Book warns that algorithms may be generating hate and discrimination

February 15 2023

Credit: Rowman & Littlefield

In an increasingly digital world, technology will promote historic social inequalities unless the system is challenged and changed, warns a new publication from Professor Yasmin Ibrahim from Queen Mary's School of Business and Management.

Prof Ibrahim's latest book draws on research from computer science to sociology and critical race studies, in a ground-breaking demonstration of how digital platforms and algorithms can shape social attitudes and behavior.

"Digital Racial: Algorithmic Violence and Digital Platforms" explores how algorithms can target and profile people based on race, as well as how digital technologies enable online hate speech and bigotry. The book also details how algorithms are no longer exclusively used for digital platforms, such as social media and online shopping; they play a hidden but growing role in vital public services like health and social care, welfare, education and banking.

There are countless examples of the dangers that digital technologies can pose—from infamous scandals like Cambridge Analytica data misuse and racial bias in the U.S. courts' risk assessment algorithm, to emerging issues like self-driving cars being more likely to hit darker skinned pedestrians and virtual assistants not understanding diverse accents.

Prof Ibrahim highlights real-world examples of how digital platforms can reveal and reinforce deep-seated inequalities—such as Facebook's algorithm contributing to the Rohingya genocide, in which an estimated 25,000 people have been killed and 700,000 more displaced. Amnesty International found that "Facebook's algorithmic systems were supercharging the spread of harmful anti-Rohingya content," and the platform failed to remove dangerous posts because it was profiting from the increased engagement.

More recently and closer to home, "The A—Level Fiasco and U-turn by the UK Government" (2020) saw an algorithm created by exam regulator Ofqual downgrade students at state schools and upgrade those at privately funded independent schools, disadvantaging young people from lower socio-economic backgrounds.

Similarly, Dutch Prime Minister Mark Rutte and his entire cabinet resigned in 2021, when investigations revealed that 26,000 innocent families were wrongly accused of social benefits fraud partially due to a discriminatory algorithm. This led to tens of thousands of parents and caregivers being falsely accused of childcare benefit fraud by the Dutch tax authorities—often those with lower incomes or belonging to ethnic minorities.

Touching on her recent work in "Technologies of Trauma," Prof Ibrahim's new book raises the issue of how best to moderate digital platforms. Since algorithms lack the humanity needed to judge what may be harmful to people, this task often falls to low-paid workers on unstable contracts, forced to look at vast amounts of traumatic content. Prof Ibrahim argues that content moderation should be treated as a hazardous business, with regulation for employers and support for employees.

Commenting on the publication of "Digital Racial," Prof Ibrahim said, "Digital technologies have the potential to bring about positive social change, but they also carry with them the risk of spreading and intensifying existing inequalities. I'm thrilled to finally be able to share this book with the world, which I hope will start a critical conversation about the role of digital platforms and the implications it can have on equality.

"With the rise of technology and its increasing role in our lives, it's more important than ever to ensure that digital spaces are not replicating racial

3/4

inequalities in our society. We must challenge algorithmic inequality to stem discrimination, hate, and violence and push for more inclusion and representation in our digital platforms."

Provided by Queen Mary, University of London