

In Tech We Trust

– Maybe Too Much?

By Hannah Mayer, Jin Paik, and Jenny Hoffman
September 21, 2020

What happens when a belief in the power and opportunity of technology meets mindfulness about potentially resulting adverse impacts? An insightful conversation – about Google ads, Airbnb hosts, and Emmanuel Macron.

“Technology is not just a business issue, it is a policy issue. In fact, technology is a policy maker,” says Harvard Professor Latanya Sweeney in a conversation with Harvard Business School Professor Karim Lakhani, as part of the [AI in Enterprise series](#) hosted by the [Laboratory for Innovation Science at Harvard](#). Sweeney is a Professor of the Practice of Government and Technology at the Harvard Kennedy School and in the Harvard Faculty of Arts and Sciences, the director and founder of the Data Privacy Lab, and the former Chief Technology Officer at the U.S. Federal Trade Commission. She uses her background in computer science to create technology to assess and solve societal, political and governance problems, with a mission to teach others to do the same. In academic terms, this translates to her devoting time to the scientific study of technology’s impact on humankind (“technology science”). While doing so, she has founded and now serves as the Editor-in-Chief at the eponymous [Journal of Technology Science](#) dedicated to rapidly publishing academic papers exploring the benefits and adverse consequences of social, political, organizational and personal aspects of technology--of which there are many.

Take Google ads, for example. When typing a person’s name into the Google search bar, the results may include a personalized ad for public records about that person that can either be neutral (e.g., “Find [person X]’s contact details”) or may contain insinuations of police encounters (e.g., “[Person X], arrested?”). Sweeney finds that when [person X] has a racially associated name (such as Tamika Jones), [ad delivery is discriminatory](#). First names, previously identified by others as being assigned at birth to more Black or White babies, are found predictive of race. Those assigned primarily to Black babies, such as DeShawn, Darnell and Jermaine, generate ads suggestive of an arrest in more than 80% of name searches, while those assigned at birth primarily to White babies, such as Geoffrey, Jill and Emma, generate more neutral ad copy, with the word “arrest” appearing as rarely as in 20% of name searches.

These and other findings have led to the emergence of algorithmic fairness as a much needed area of study. “We tend to trust algorithms because we associate such things with it as mathematics and computers, thinking that such seemingly objective entities will not be biased. Similarly, we do not question algorithms because they are business property and somewhat

secretive. But the reality is that algorithmic discrimination is widespread, and it affects people's lives," Sweeney illustrates. She advocates for more accountability on the part of businesses because they indirectly create policies adversely affecting the livelihoods of many millions of individuals.

One key question remains: Where does the bias originate from? Sweeney likens technology (and algorithms) to a box dependent on inputs (such as training data) and external interactions (such as clicks on ads) to yield results. If an adverse impact emerges as a side effect of the output (say, the ad), what is the real cause of that adverse effect? Sweeney argues it can be rooted both in the inputs *and* the external reactions.

Training data is routinely biased. Take, for instance, facial recognition software. Facial recognition training data is almost exclusively composed of White male photos, making face detection for women and people of color difficult or even faulty. But even when facial recognition software does work, it can have enormous adverse impacts, such as in the case of recidivism algorithms. When a White male and Black female were arrested for independently committed petty thefts and their mug shots were factored into the algorithm predicting future crimes, the Black female was rated at a high risk of subsequent offenses, despite only having a record for some juvenile misdemeanors. Meanwhile the White male offender was rated as having a low risk of subsequent offences, despite prior offenses including multiple armed robberies. This goes to show there is much work to be done in correcting seemingly objective algorithms for biased input data. (In case you're interested: he later committed a grand theft, while she did not reoffend.)

However, external interactions (in other words, humans) can be just as much at fault for adverse impacts of technology. For instance, one paper in Sweeney's *Journal of Technology Science* found that [Asian American Airbnb hosts earn up to 20% less](#) than White hosts for comparable rentals. This is likely not because Asian Americans generally enjoy pricing their rentals cheaper, but because they are unable to achieve the same rental prices as their White peers in the market.

It is against this backdrop that Sweeney urges all stakeholders involved to support fair and equitable technology and algorithms – from the founders, software engineers and executives to the regulators, journalists and courts. French President Emmanuel Macron has undertaken one step in that direction, announcing that government data would be open in France and that all AI algorithms used for government decision-making would be public. Sweeney interprets this as an open call to both assess whether the data has any inherent biases and to understand and challenge the equitability of the algorithms themselves. It is this commitment to making algorithms a socially just instrument that Sweeney says we need now for technology to rightfully earn our trust (back).

Join us for the next *AI in Enterprise* event at 11am on October 14th with Roger Magoulas:
<https://bit.ly/ai-enterprise-magoulas>

Link to Podcast & Video: