

# Data ethics is more than just what we do with data, it's also about who's doing it

June 22 2018, by James Arvanitakis, Andrew Francis And Oliver Obst

---



Credit: AI-generated image ([disclaimer](#))

If the recent Cambridge Analytica data scandal has taught us anything, it's that the ethical cultures of our largest tech firms need tougher scrutiny.

But moral questions about what data should be collected and how it

should be used are only the beginning. They raise broader questions about who gets to make those decisions in the first place.

We currently have a system in which power over the judicious and ethical use of data is overwhelmingly concentrated among white men. Research shows that the unconscious biases that emerge from a person's upbringing and experiences can be baked into technology, resulting in negative consequences for minority groups.

These biases are difficult to shed, which makes workplace diversity a powerful and necessary tool for catching unsuspected bias before it has a chance to cause damage. As the impact of data-driven algorithms and decisions grows more profound, we need to ask: how is this going to change in the future?

Unfortunately, the indicators suggest the answer is: not much.

## **What consequences are we talking about?**

[Algorithmic bias](#) is now a [widely studied problem](#) that refers to how human biases creep into the decisions made by computers.

The problem has [led to](#) gendered language translations, biased criminal sentencing recommendations, and racially skewed facial recognition systems.

For example, when an automated translation tool such as Google Translate is required to translate a gender-neutral language (such as Turkish) into a gender-specific one (such as English) it makes a guess as to which gender to assign to the translated text.

[People noticed](#) that Google Translate showed a tendency to assign feminine gender pronouns to certain jobs and masculine pronouns to

others – "she is a babysitter" or "he is a doctor" – in a manner that reeked of sexism. Google Translate bases its decision about which gender to assign to a particular job on the training data it learns from. In this case, it's picking up the gender bias that already exists in the world and feeding it back to us.

If we want to ensure that algorithms don't perpetuate and reinforce existing biases, we need to be careful about the data we use to train algorithms. But if we hold the view that women are more likely to be babysitters and men are more likely to be doctors, then we might not even notice – and correct for – biased data in the tools we build.

So it matters who is writing the code because the code defines the algorithm, which makes the judgement on the basis of the data.

## **Who holds the power?**

Only ten years ago the first smartphones were making their mark. Today some of the most powerful people on the planet are those who control data gathered through mobile technologies.

Data is central to the functioning of the modern world. And power over business, democracy and education will likely continue to lie with data and data-dependent tools, such as machine learning and artificial intelligence.

Currently, the people who have the power to make ethical decisions about the use of data are typically white males from [high-earning, well-educated families](#).

One research company, Open MIC, which describes itself as "investing in racial diversity in the tech world", [reviewed data](#) from some of the biggest tech firms and found a consistent pattern: disproportionate

percentages of white employees compared with the wider working population.

Adobe's workforce is 69% white, Apple's is 56% white, Google is 59% white and Microsoft is 58% white. The [list](#) goes on: "Black people, Latinos, and Native Americans are underrepresented in tech by 16 to 18 percentage points compared with their presence in the US labour force overall."

This is made far worse by a crippling lack of gender diversity.

In a 2017 Microsoft [report](#), a survey of UK IT and tech leaders found that on average, the gender mix among their teams was 80% male and 20% female. A staggering 35% of respondents had no plans in place to change this imbalance.

The numbers are similar in Australia, according to a [study](#) of Australian professional profiles on the social network LinkedIn.

It revealed that just 14% of executive roles in the local tech industry were held by women. Of the 435,000 people in IT listed on LinkedIn in Australia, only 31% were women. Even these numbers may be optimistic, according to Australia's Chief Scientist, Alan Finkel, [who noted](#) that women make up less than one-fifth of Australians qualified in science, technology, engineering and maths.

## **Will this change?**

Those likely to be in charge of developing the algorithms of the future are those who are studying computer science and mathematical sciences right now. Sadly, the groups dominating those subjects at schools and universities largely reflect the current workforce.

Australian domestic students enrolled in tertiary level information technology [dropped](#) from a peak of 46,945 in 2002 to 27,547 in 2013. While the numbers have improved slightly according to AEN University Rankings, females in engineering and IT still [represent less than one in five students](#).

Meanwhile, the number of girls at the senior high school level taking the advanced computing and mathematics subjects needed to enter these roles remains [resolutely low](#).

This ship is taking a long time to turn around.

## **What can we do about it?**

If the coders of the future are today's middle-class boys, how are we preparing them to make unbiased ethical choices when they become the Zuckerbergs of tomorrow? And how can we steer the ship so that the wealth and power that will continue to flow from mastery of such technical skills is not denied to those who are not white and male?

Our education system is unwittingly allowing boys to train as technical people without the skills to put their work in a social context, and allowing girls to do the reverse.

Indeed, while many of the smartest young women are choosing to go into medicine or law, these professions are vulnerable to the advance of [artificial intelligence](#) – paralegals, radiologists, and those making [preliminary diagnoses](#).

We are in a structure in which the same old imbalances are strengthening and look to persist. But this is not the way it should be. Unless we confront the culture through big shifts in educational trends, nothing will change.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: Data ethics is more than just what we do with data, it's also about who's doing it (2018, June 22) retrieved 18 April 2024 from <https://phys.org/news/2018-06-ethics.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.