Research: AI is in danger of becoming too male

August 14 2019, by Juan Mateos-Garcia And Joysy John

Credit: Pavel Danilyuk from Pexels

Artificial Intelligence (AI) systems are becoming smarter every day, beating world champions in games like Go, identifying tumors in medical scans better than human radiologists, and increasing the

efficiency of electricity-hungry data centers. Some economists are comparing the transformative potential of AI with other "general purpose technologies" such as the steam engine, electricity or the transistor.

But current AI systems are far from perfect. They tend to reflect the biases of the data used to train them and to break down when they face unexpected situations. They can be gamed, as we have seen with the controversies surrounding misinformation on social media, violent content posted on YouTube, or the famous case of Tay, the Microsoft chatbot, which was manipulated into making racist and sexist statements within hours.

So do we really want to turn these bias-prone, brittle technologies into the foundation stones of tomorrow's economy?

Minimising risk

One way to minimize AI risks is to increase the diversity of the teams involved in their development. As research on collective decision-making and creativity suggests, groups that are more cognitively diverse tend to make better decisions. Unfortunately, this is a far cry from the situation in the community currently developing AI systems. And a lack of gender diversity is one important (although not the only) dimension of this.

A review published by the AI Now Institute earlier this year, showed that less than 20% of the researchers applying to prestigious AI conferences are women, and that only a quarter of undergraduates studying AI at Stanford and the University of California at Berkeley are female.

The authors argued that this lack of gender diversity results in AI failures that uniquely affect women, such as an Amazon recruitment

[system](#) that was shown to discriminate against job applicants with female names.

Our recent report, [Gender Diversity in AI research](#), involved a "big data" analysis of 1.5m papers in arXiv, a pre-prints website widely used by the AI community to disseminate its work.

We analyzed the text of abstracts to determine which apply AI techniques, inferred the gender of the authors from their names and studied the levels of gender diversity in AI and its evolution over time. We also compared the situation in different research fields and countries, and differences in language between papers with female co-authors and all-male papers.

Our analysis confirms the idea that there is a gender diversity crisis in AI research. Only 13.8% of AI authors in arXiv are women and, in relative terms, the proportion of AI papers co-authored by at least one woman has not improved since the 1990s.

There are significant differences between countries and research fields. We found a stronger representation of women in AI research in the Netherlands, Norway and Denmark, and a lower representation in Japan and Singapore. We also found that women working in physics, education, biology and social aspects of computing are more likely to publish work on AI compared with those working in computer science or mathematics.

In addition to measuring gender diversity in the AI research workforce, we also explored semantic differences between research papers with and without female participation. We tested the hypothesis that research teams with more gender diversity tend to increase the variety of issues and topics that are considered in AI research, potentially making their outputs more inclusive.

To do this, we measured the "semantic signature" of each paper using a machine learning technique called word embeddings, and compared these signatures between papers with at least one female author and papers without any women authors.

This analysis, which focuses on the Machine Learning and Social Aspects of Computing field in the UK, showed significant differences between the groups. In particular, we found that papers with at least one female co-author tend to be more applied and socially aware, with terms such as "fairness", "human mobility", "mental", "health", "gender" and "personality" playing a key role. The difference between the two groups is consistent with the idea that cognitive diversity has an impact on the research produced, and suggests that it leads to increased engagement with social issues.

How to fix it

So what explains this persistent gender gap in AI research, and what can we do about it?

Research shows that the lack of gender diversity in the science, technology, engineering and mathematics (STEM) workforce is not caused by a single factor: gender stereotypes and discrimination, a lack of role models and mentors, insufficient attention to work-life balance, and "toxic" work environments in the technology industry come together to create a perfect storm against gender inclusion.

There is no easy fix to close the gender gap in AI research. System-wide changes aimed at creating safe and inclusive spaces that support and promote researchers from underrepresented groups, a shift in attitudes and cultures in research and industry, and better communication of the transformative potential of AI in many areas could all play a part.

Policy interventions, such as the [£13.5m investment from government](#) to boost diversity in AI roles through new conversion degree courses, will go some way towards improving the situation, but broader scale interventions are needed to create better links between arts, humanities and AI, changing the image of who can work in AI.

While there is no single reason why girls disproportionately stop taking STEM subjects as they progress through education, there is evidence that factors including [pervasive stereotypes around gender](#) and a [teaching environment that impacts the confidence of girls more than boys](#) play a part in the problem. We must also showcase those role models who are using AI to make a positive difference.

One tangible intervention looking to tackle these issues is the [Longitude Explorer Prize](#), which encourages secondary school students to use AI to solve social challenges and work with role models in AI. We want young people, particularly girls, to realize AI's potential for good and their role in driving change.

By building skills and confidence in young women, we can change the ratio of people who study and work in AI—and help to address AI's potential biases.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation