Can a new independent institute improve the ethics of AI?

Former Google AI researcher Timnit Gebru is launching an independent AI institute to explore how the technology can exacerbate inequality. The Distributed Artificial Intelligence Research (DAIR) Institute will document the harmful effects AI can have on marginalised communities, while looking for ways to improve access to AI development for those groups. Gebru is a longstanding critic of facial recognition software's bias against people of colour.

Temps de lecture : minute

---

13 December 2021

<u>*Article originally published on Curation*</u>

She said publishing papers that were picked up by academia, journalists and regulators was more successful at changing Google's policies than raising issues internally in her former role as co-lead of the ethical AI group. (<u>*The Washington Post*</u>)

A quick recap

This news comes just a year after Gebru left Google following a dispute with the company over her published research. In a <u>*paper*</u> on AI language models, Gebru and colleagues asked questions and made recommendations to the industry about how large firms mitigate risks of bias, particularly bias against racial diversity and minority ethnic groups. In response, Google <u>*pushed Gebru to withdraw*</u> the paper, after which she was let go.

In protest of what many saw as Google cracking down on ethical research, which might pose barriers to their work, and the loss of the important voice of a woman of colour advocating for more responsible standards of development, 2,000 Google employees signed a petition demanding justice.

After widespread backlash against the decision, Google CEO Sundar Pichai issued an _apology_, but then the firm made headlines again in February 2021 for _firing_ Margaret Mitchell, Gebru's co-lead on the company's AI ethics team.

Why does this matter?

In their paper Gebru and her co-authors _posed questions_ about the responsible production of AI technology. With only a handful of firms with the resources and talent to pursue large-scale products, and outdated regulatory frameworks designed either by the companies themselves or by regulators without the technical expertise to understand the complex issues such technologies face, the paper asked "how big is too big?".

A key issue the DAIR Institute looks to tackle is _changing the culture_ around tech development to make it more inclusive, both internally and in use. Hiring research staff with diverse ethnic and social backgrounds is a priority, says the institute, as is creating a space where success for researchers is measured by the quality of work rather than the volume produced. DAIR's work will also focus on social impact technology, which isn't typically prioritised by corporates, such as one project it's undertaking to create a public data set of aerial images to evaluate the traces of apartheid in South Africa.

How's Google doing after the break-up?

To some extent it seems the outrage following Gebru's departure has

motivated a new drive within the company to champion ethical research and technology with a social purpose – it launched an inclusive _writing assistant_, has joined a _number_ of coalitions pushing for greater diversity in tech, launched a _number_ of technologies focused on improving _accessibility_, and has even grappled with larger social issues such as police brutality in the wake of the Black Lives Matter protests with a _VR offering_ to train law enforcement in de-escalation. In its 2021 AI Principles Progress Update, Google also reports to have published 500 papers around ethical research since 2018.
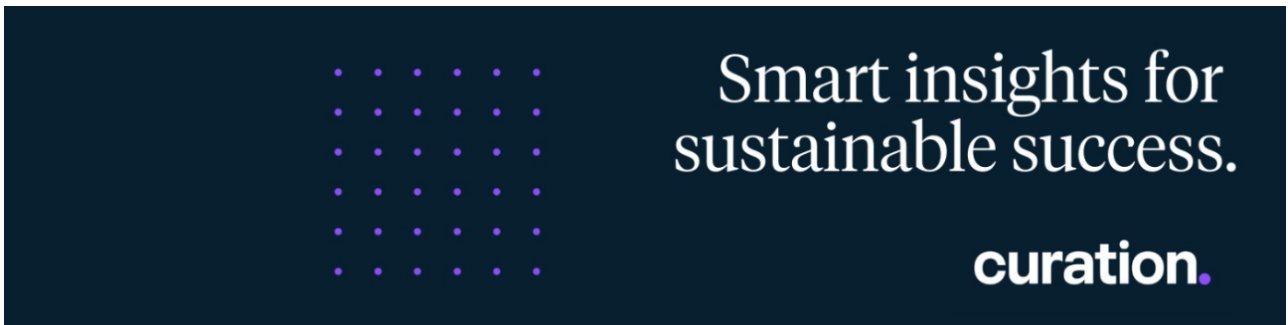
However, some of its news has not been so positive. In November, the company _confirmed_ it was pursuing a defence contract it had previously bid for in 2018. At the time, Google was forced to withdraw its bid after 4,000 employees launched a protest against the use of their technology in weapons development.

A key takeaway

Now that technology (particularly Big Tech) has become essential to our lives, more and more people are questioning it. Many voices, like those of Timnit Gebru, Margaret Mitchell, _Frances Haugen_, Aerica Shimizu Banks, Laurence Berland… (_the list_ goes on) are coming from the inside, but others – like US Federal Trade Commission Chair Lina Khan, are trying to dismantle the monopoly these firms have on tech development and expertise from the outside.

Something they all agree on? The work being done at such companies is being done without proper regulatory oversight, which allows for ethics to be compromised in pursuit of rapid technological development. The work being done by organisations like DAIR, _NYU's AI Now Institute_, the _Algorithmic Justice League_, and _Data for Black Lives_ is a key step towards bridging the private-public knowledge gap and creating frameworks for more sustainable, responsible and impactful technology.

Sara Trett is Sustainability Editor at _Curation_ where this article was originally published

Article by Sara Trett