

Signal's Meredith Whittaker: 'These are the people who could actually pause AI if they wanted to'

Publication Date: 2023-06-11

Author: Ian Tucker

Section: Technology

Tags: Artificial intelligence (AI), The Observer, Social media, Digital media, Computing, interviews

Article URL: <https://www.theguardian.com/technology/2023/jun/11/signals-meredith-whittaker-these-are-the-people-who-could-actually-pause-ai-if-they-wanted-to>



Meredith Whittaker is the president of Signal – the not-for-profit secure messaging app. The service, along with WhatsApp and similar messaging platforms, is opposing the UK government's online safety bill which, among other things, seeks to scan users' messages for harmful content. Prior to Signal, Whittaker worked at Google, co-founded NYU's AI Now Institute and was an adviser to the Federal Trade Commission. After 10 years at Google you organised the walkout over the company's attitude to sexual harassment accusations, after which in 2019 you were forced out. How did you feel about that? Let me go back into some of the details, because there's a kind of broad story, and it matters for this moment. I was running a research group looking at the social implications of AI. I was pretty well known in the company and outside as somebody who discussed these issues in ways that were counter to Google's public messaging. I was an internal dissenter, an academic. Toward late 2017 I had a colleague flag me that there was a secret contract [known as Maven] between Google and the Department of Defense [DOD] to build AI systems for drone targeting. That for me was when my organising started, because I realised I had been presenting very clear arguments that people agreed with ... but it didn't really matter. This was not an issue of force of argument. This was an issue of the ultimate goals of this company are profit and growth and DOD contracts were always going to sort of trump these intellectual and moral considerations. So that was when you began organising a walkout? What do you do if you don't have the power to change it, even if you win the debate? That was where the organising started for me. I wrote the Maven letter [which gathered 3,000 employee signatures] and we got the Maven contract cancelled. The walkout was a big spectacle. There was a news story about Andy Rubin getting a \$90m (£72m) payout after an accusation of sexual misconduct [which he denies]. What came to a head was deep concerns about the moral and ethical direction of Google's business practices and an understanding that those moral and ethical lapses were also reflected in the workplace culture. How did I feel about that? I am very happy. I did not stand by and let my integrity get eaten away by making excuses for being complicit. Would you say Google is pretty typical of big tech companies, that they're only really bothered about moral and ethical concerns if they affect their bottom line? Ultimately every quarter the executives at Google meet the board and they need to be reporting up into the right projections on growth and revenue. Those are the key objectives of shareholder capitalism. If Sundar [Pichai, Google CEO] went to the board and said: "Morally we need to leave \$10bn on the table. Let Microsoft have this contract," he'd be fired in an instant. It is not a form or a set of goals that is going to be amenable with putting the social good first. Isn't there a business case for stamping out sexual harassment, and having a diverse workforce, and fixing biases in algorithms and so on? There is no Cartesian window

of neutrality that you can put an algorithm behind and be like, “This is outside our present and history.” These algorithms are trained on data that reflects not the world, but the internet – which is worse, arguably. That is going to encode the historical and present-day histories of marginalisation, inequality etc. There isn’t a way to get out of that and then be like, “This is a pristine, unbiased algorithm,” because data is authored by people. It’s always going to be a recycling of the past and then spitting that out, projecting that on to the present. We can say there’s a business case, but let’s be real about it. If you’re going to see a workforce transformed in the name of equality there are many powerful people who will have to lose their position, or their status, or their salary. We live in a world where people are misogynist and racist. Where there is behaviour from some that assumes women or non-white men shouldn’t be in the room. They’re making billions of dollars now. We can’t leave this up to a business case, I think, is the argument I’m making. So in 2020-21 when Timnit Gebru and Margaret Mitchell from Google’s AI ethics unit were ousted after warning about the inequalities perpetuated by AI, did you feel, “Oh, here we go again”? Timnit and her team were doing work that was showing the environmental and social harm potential of these large language models – which are the fuel of the AI hype at this moment. What you saw there was a very clear case of how much Google would tolerate in terms of people critiquing these systems. It didn’t matter that the issues that she and her co-authors pointed out were extraordinarily valid and real. It was that Google was like: “Hey, we don’t want to metabolise this right now.” Is it interesting to you how their warnings were received compared with the fears of existential risk expressed by ex-Google “godfather of AI” Geoffrey Hinton recently? If you were to heed Timnit’s warnings you would have to significantly change the business and the structure of these companies. If you heed Geoff’s warnings, you sit around a table at Davos and feel scared. Geoff’s warnings are much more convenient, because they project everything into the far future so they leave the status quo untouched. And if the status quo is untouched you’re going to see these companies and their systems further entrench their dominance such that it becomes impossible to regulate. This is not an inconvenient narrative at all. After Google you were appointed senior adviser on AI at the Federal Trade Commission (FTC). Did you find the process of trying to regulate disillusioning? I didn’t come in with rose-coloured glasses. I had been speaking with folks in Congress, folks in the FTC and various federal agencies for years and at one point I testified in front of Congress. I wasn’t new, I wasn’t coming to DC with a suitcase and a dream, but it did give me a sense of how fierce the opposition was. I’ll say the minute it was announced that I was getting the FTC role, far-right news articles dropped. The bat-signal went out. Of course, I’m not the only one. Lina’s office [Khan, FTC chair] was inundated with this. But within an agency where there’s a vision, where there is an analysis that is not grounded in sci-fi but in an understanding of these systems and this industry, there was a huge amount of opposition. That was coming from tech lobbyists, that was coming from the Chamber of Commerce, that was coming from some folks within the FTC who would like to go get a nice little general counsel job at a firm and don’t want their name on something that might be seen as anti-business. There was great vision there, but the machinery is difficult to operate under those conditions. Unlike many other tech entrepreneurs and academics you didn’t sign either of the two recent petitions, the Future of Life institute “pause AI” letter or last month’s Center for AI Safety “existential threat” letter. No. I don’t think they’re good faith. These are the people who could actually pause it if they wanted to. They could unplug the data centres. They could whistleblow. These are some of the most powerful people when it comes to having the levers to actually change this, so it’s a bit like the president issuing a statement saying somebody needs to issue an executive order. It’s disingenuous. What appealed to you about taking the position at Signal? I’ve been on their board for a number of years. Signal plays an important harm-reduction role in providing a mechanism for truly private communications in a world where in the past 30 years large corporations have made the decisions for us that all of our communications, our activity etc should be surveilled. You’ve threatened to pull Signal from the UK if the online safety bill currently making its way through parliament is passed. I’ll repeat what I said, which wasn’t quite the headline. We will never undermine our privacy promises. We will never adulterate our encryption, we will never participate in any regime that would enforce breaking those promises. The bill proposes that users’ messages are scanned “client side” – ie on the phone – and the messages would be encrypted when sent... The claims being made around client-side scanning, that it can protect privacy while also conducting mass surveillance, is magical thinking. It is cynically weaponising a meaningless semantic distinction between mass surveillance that happens before encryption takes effect and breaking encryption. You have politicians out there saying: “This doesn’t break encryption, don’t worry.” That is extraordinarily dishonest. And moderation in other spheres doesn’t work particularly well – software AI scanning messages and then difficult cases being referred to humans... This system could never work without massive amounts of human intervention. What you’re looking at is an unthinkable expensive system that would need constant human labour, constant updating. You’re looking at hundreds of thousands of people being locked out of their accounts with no recourse, grinding people’s daily lives, employment, and the economy. This is a fantasy made up by people who don’t understand how these systems actually work, how expensive they are, and how fallible they are. What is helping drive this, helping this seem plausible, is the AI hype. That is painting these systems as superhuman, as super capable, which leads to a sense in the public, and even among some politicians, that tech can do anything.