How AI discriminates and what that means for your Google habit

March 29 2024, by Julia Busiek

Safiya Umoja Noble is the David O. Sears Presidential Endowed Chair of Social Sciences and professor of Gender Studies and African American Studies at UCLA. She is the Director of the Center on Race & Digital Justice and Co-Director of the Minderoo Initiative on Tech & Power at the UCLA Center for Critical Internet Inquiry (C2i2) and the Interim Director of the UCLA DataX Initiative. Credit: UCLA

Safiya Umoja Noble swears she is not a Luddite. But she does think we

could all learn a thing or two from the machine-bashing textile craftsmen in 19th-century Britain whose name is now synonymous with technological skepticism.

"The Luddites knew that these new tools of industrialization were going change the way we created and the way we did work," said Noble, professor of gender studies and African American studies at UCLA.

Noble's scholarship examines new tools that are changing the way we work: Her 2018 book, "Algorithms of Oppression: How Search Engines Reinforce Racism," sparked a [national conversation](#) about algorithmic discrimination in [internet search engines](#) and earned her a MacArthur "genius" fellowship.

In February, Noble discussed her work during a keynote conversation with UCLA Executive Vice Chancellor Darnell Hunt at UC's recent Academic Congress on AI.

In response to the work of Noble and others, [tech companies](#) have fixed some of their most glaring search engine problems. But the algorithms that govern our Google results are just one of the multiplying ways that artificial intelligence is shaping our lives and livelihoods, and Noble says neither corporate self-policing nor government regulation has kept pace with the technology's growth and potential for harm.

We recently sat down with Noble to learn more about why the internet works the way it does, and what that means for an AI-powered future.

Why did you write 'Algorithms of Oppression' and what difference has your work made?

The book was an exploration of what happens when you ask search

engines about society, and especially about [vulnerable people](#). When I went to Google and searched Black girls, Asian girls, Latina girls, 80% of the first page of search results was pornography or hypersexualized content. And I didn't have to type the words "sex" or "porn." Black girls and women were already synonymous with pornography.

"Algorithms of Oppression" set out to describe what is happening here. And I think the breakthrough of that work, which is now kind of common sense, is that algorithms can discriminate. Everyday users understand that now. "Algorithm" is part of our lexicon.

On the academia side, now we have thousands of researchers all over the country and really tens of thousands probably around the world who are working on the harms of tech. And my work has really been taken up in industry. I thought Silicon Valley would be hostile to it, but I think it gave language to things that they also experienced.

What are some ways that search engine results affect people's lives?

When you have identities that have been historically oppressed, you experience systems in different ways. Some people get to still just share cats for the LOLs, while others are completely traumatized with racist content.

If all the Black women and girls in the country put all of our money together, we still would never have as much money as the porn industry. We will never be able to control the way in which our keywords are optimized. So how will we intervene upon these kinds of systems?

Google and other companies are the first to say that they know these problems exist and they're working on them. But I don't see them truly

being able to deal with the power and inequalities in their systems. If anything, they're just kind of tweaking these algorithms rather than remaking them in profound ways.

If search results aren't neutral, why do we get the results we do?

The prevailing logic ten years ago was that anything that showed up in a search engine was just reflecting us back on ourselves. There was absolutely no awareness of, at a fundamental level, the way these platforms are coded and how they prioritize certain types of values over others.

First of all, there is an algorithm that's sorting through millions of potential websites that could be surfaced to the first page. And then the results you see are influenced by industries, foreign operatives and [political campaigns](link) through the gray market of search engine optimization. That's a whole cottage industry that exists to figure out how to manipulate search engines.

The companies that own search engines are built to respond to the highest bidder. Yes, they are constantly trying to refine and detect where their systems are being gamed. But they've made products that can be gamed. So there's culpability on the part of companies who make products that are easily manipulable.

How has your research changed the way you use the internet?

I have so much skepticism around search. I don't use it like I'm going to the library. I use it for shopping, which is really what it's probably best designed for.

With something like television or movies, we understand that there is a subjective point of view. If there are racist or sexist misrepresentations in other media, we understand that there's a point of view from the director or the writers.

But we don't think of search results as subjective. We use search engines like fact checkers. And they're very reliable for certain kinds of facts: a mathematician can enter a formula and look to see where there's an error in the code. Or you can use it for banal things like, "Where's the closest coffee shop?" You'll get an answer and it's pretty reliable.

So if it's reliable for things that are fairly meaningless, then that reinforces your belief that it's meaningful for everything. But if you use that same tool for a question that is social or political, do you think you're going to get a reliable answer?

If you don't use Google and other search engines for social or political information, what do you do instead?

We have these things called libraries! Think about what happens before something even shows up in a library: There have been editors, reviewers, a publisher. And then there are librarians, who take very seriously their obligation to the public and to the preservation of knowledge. They're trained in categorizing knowledge in ways that make it understandable to the public what they are engaging with.

So if you're looking for a book about the Holocaust at the library, you are going to find that in context relative to many other works about the Holocaust. When you pull it off the shelf, you can discover other things around it: is it in history, or current affairs, or sociology?

If you just use a search engine, you might get a Holocaust denial website that looks like legitimate information, and it might be difficult for you to

discern what it is because it's completely decontextualized. It isn't like it's in a wrapper that says, "This is in the white supremacist section of the internet!"

Enter: ChatGPT. What are you seeing in terms of how generative AI is either reproducing or counteracting the kinds of discrimination you've identified in search, and why?

So ChatGPT is a type of AI that's built on what we call a [large language model](link). These models suck in basically everything that's available on the internet into their training data, which, for reasons we've talked about, isn't always a great idea. They take in copyrighted works and academic scholarship and also, like, random subreddits, as if these things are all equally reliable.

The companies that are producing generative AI have released products that are not ready for prime time. We're seeing stories in the news every week now around generative AI tools and their problems. A lot of what large language models produce isn't true, for one. And recently one of the generative AI art tools [refused to render an image of Black doctors treating white patients](link).

I mean, these large language models don't have agency, so they can't actually refuse. They're just statistical pattern matching tools. But the lack of possibilities for things like multiple types of gender expression, multiple types of racial and ethnic representation—where we are limited in society with those things, we're even more limited in the data that trains models. So the models are limited in being able to even produce certain types of results.

What are you doing to prepare your students to

navigate a world shaped by AI?

I've used Google search for more than a decade to teach media literacy to my students. I'll have students do searches on identities that are important to them and then bring back the results and we have a conversation about it. So I'd have students search for, say, "sorority girl," and you can imagine their reactions to the results, like, "That's so not right." The results often are so profoundly racist or sexist, or a misrepresentation in some ways of who students see themselves as.

In some ways, ChatGPT is just the next version of search, because it doesn't differentiate propaganda and evidence. Students come into my class and use propaganda sites as evidence, because they can't quite tell the difference. So being able to have these conversations in my classroom is important. But where I'd rather see more intervention is at the policy level. Right now, it's like, well AI is here, that's just how it is. You, the public, have to learn how to use it, you've got to outsmart it, you've got to figure it out. That seems woefully insufficient when we could regulate these tools differently and limit their adoption to where it makes sense.

Is there any cause for optimism that better algorithms could cut down societal biases faster than if we left all creation and decision-making up to humans, and our implicit biases, forever?

People who make the predictive AI models argue that they're reducing human bias. But there's no scenario where you do not have a prioritization, a decision tree, a system of valuing something over something else. So it's a misnomer to say that we endeavor to un-bias technology, any more than we want to un-bias people. What we want to do is be very specific: We want to reduce forms of discrimination or

unfairness, which is not the same thing as eliminating bias.

You could also characterize this as: We hold certain values. And we want to ensure that those values are met. So that means we would have to acknowledge that the technology does hold as a particular set of values or is programmed around a set of values. And we want to optimize to have more of the values that we want.

We try to do that with people. That's why we get an education. That's why we learn about other people. That's why we engage in art and things that sensitize us to the breadth of humanity.

To say you want a technology that's completely neutral and void of any markers that differentiate people, you're just defaulting to the priorities that are driven by your own biases. If we want to work toward pluralistic and pro-democratic priorities, we have to program toward the things we value.

What are some of the values that have governed tech as we know it?

Ten years ago, if search companies were prioritizing, let's say, the well-being of women and girls, they would not have let the porn industry be the highest bidder in their society. It just wouldn't have happened. They biased toward profit at the expense of women.

We're no longer in an era where porn dominates search because the companies have had to respond to their many critics. But there are still troubling ideological world views from leaders in Silicon Valley. They believe technocrats can design a better society and that democracy is too messy. Those politics are imbued in the products they make, who they're pointed toward, who's experimented upon and who's considered

disposable around the world.

I've had opportunities to think about doing this kind of work in other places. But my heart is at UC because the public institutions are the powerful counterweight to the private sector. It's very important that we do our research on behalf of the public, the people of the State of California and this country, because we hopefully will not be as beholden to the kinds of pressures that people who work in industry are.

What lessons about living with AI can we draw from past social transformations?

I tell my undergraduate students, "When my mom gave birth to me, she had a cigarette in her mouth. And the doctor did, too." And they're horrified! They're like, "There's no way that happened!" And I'm like, "That absolutely happened."

We used to look at magazines when I was little and see ads like, "Three out of four doctors prefer Camel cigarettes." On every movie, every TV show, to be glamorous was to be smoking. The tobacco industry shaped our sense of reality. That was so normalized that it still surprises me that I don't see anyone smoking anymore. And that's because we regulated smoking.

People growing up today just can't even imagine a time before. So I think we may look back on this era and say, why didn't I just ask my mom that question instead of Google? Why didn't I just read a book? How did we even get here?