

Engine Failure

Safiya Umoja Noble and Sarah T. Roberts on the Problems of Platform Capitalism



Safiya Umoja Noble and Sarah T. Roberts.

The internet is dominated by a handful of big companies. How do they organize information to maximize profit? What are the consequences for the public sphere? And what would a better model look like?

To answer these questions, we turned to two Los Angeles-based scholars who have thought deeply about the interaction between the internet and capitalism. Safiya Umoja Noble is an assistant professor in the Annenberg School of Communication at the University of Southern California, and Sarah T. Roberts is an assistant professor in the UCLA Graduate School of Education and Information Studies.

Tell us a bit about the project to build a new kind of search engine. What are the problems you're trying to solve for? What's wrong with the existing search architecture—and how do you propose to fix it?

Safiya:

I want to build a non-commercial search engine that makes its biases visible. If you think about Google, its architecture is based on using hyperlinks as an index of relevance. It's based in large part on popularity—the more people link to your site, the higher your rank.

When Sergey Brin and Larry Page first started figuring out how their search engine would work, they borrowed their ideas from library and information science practices, and in particular from

citation analysis. Citation analysis is a way of assessing the alleged importance of scholarship. The logic is that if you are cited by someone else, then your work is relevant.

But citation analysis fails in many ways because it doesn't tell us whether your work is being argued against rather than supported. For example, Sarah might cite my work and say that it's terrible. But she's still going to cite me—and the citation itself won't have a value. You won't be able to tell from the bibliography whether she's disagreeing or agreeing with me.

It also doesn't often pick up voices in the margins, where people are writing in small fields that are not represented by powerful journals or publishing houses. The metrics used don't capture all of the ways that knowledge is being created and disseminated.

This is similar to the logic of how search engines function. But here's the problem: one of the things that we know from information science is that the signal is not neutral. That's a fundamental flaw in search engine design, and some of its output. Simply pointing to something doesn't necessarily mean that it's legitimate or credible or valuable—as we are learning from watching the national political news landscape.

I would add that the system of “ranking” means something very specific in our cultural context in the United States. If a link ranks at number one, or appears on the first page of search results, it's generally considered to mean it's the best possible result. If a link ranks at the bottom, it's likely considered the worst.

But where content shows up in search engine results is also tied to the amount of money and optimization that is in play around that content. So we are operating on many flawed assumptions, given what we know about how those rankings are produced.

Why should we care about these flaws? What are the broader social stakes here?

Safiya:

As I mentioned, one of the most important factors that enables a link to rise to the top of the rankings is the amount of capital you have. This has been borne out by many different researchers.

For example, the porn industry has a tremendous amount of capital—and it uses that capital to influence search. One big porn company that owns thousands of websites will have those sites link to each other extensively, in order to bolster those sites' rankings and dominate a number of keywords. They can also buy interesting combinations of keywords that will guarantee that they will control the representation of women in search, especially women and girls of color. This is why for many years you could do a search in a commercial search engine like Google on the words “black girls,” “asian girls,” and “latina girls,” without even adding the words “sex” or “porn,” and get back pornography. Those identities have been synonymous with porn.

That's what happens when we take human decision-making out of knowledge management. We don't have cataloguers on the web the way we do in a library. Instead, we have people designing algorithms that exert tremendous power over our society but who, quite frankly, have very little understanding of our society.

I'm often asked, "What does Silicon Valley need to do to improve? Should they hire more black and Latino and women programmers?"

Yes, yes they should—but they also need to hire people with advanced degrees in ethnic studies and women's studies and sociology. There are some forms of knowledge that can't be algorithmically curated because they're forms of knowledge that have been contested for hundreds, if not thousands, of years. Sometimes you can't get a clear answer in .03 seconds.

When you're looking for banal information like the location of the closest Starbucks, Google is highly reliable. But not when you start asking more complex questions.

What would a search engine that is capable of acknowledging that complexity look like?

Safiya:

Well, let's look at an egregious case of search failing to acknowledge the complexity of something like the history of anti-black racism. I wrote about this in my new book, *Algorithms of Oppression: How Search Engines Reinforce Racism*. In it, I discuss Dylann Roof, the Charleston mass murderer, who said he Googled the phrase "black on white crime" after the Trayvon Martin shooting. He has talked about how important that experience was in forming his white supremacist views.

He noted in his online diary that when he Googled the phrase "black on white crime," the search engine gave him information that shocked him—and helped him come to a different understanding about the so-called truth about race and the value of a multiracial society. That's because his search only returned the white supremacist websites that use such a phrase—a phrase that is used by hate-based sites to radicalize white Americans against African Americans and other people of color, including Jewish people. Google didn't provide any context on the white supremacist movement. It didn't provide any counterpoints of view.

So the experiment I'm engaged in is to try to theorize, prototype, and imagine an alternative search experience—one that makes visible the engine's many biases, and points us toward information on the web that might be more contextualized. When you go to Google, it's just a simple box against a simple background. And that conveys, through its aesthetic, the idea that there's nothing going on. Its design logic is so simple: type in a word, and you'll get something back. It implies that you don't need to know what's going on "under the hood," so to speak.

I have spent many years thinking about search engines and other types of automated decision systems, and how I can change the public's consciousness about what these commercial information brokers are actually doing.

To what extent is the way that platforms like Google organize information due to the ignorance or negligence of the programmers who are designing the algorithms? And to what extent are these conscious decisions that are part of the business model?

Safiya:

Right after the election, when you did a search for election news on Google, it returned sites that discussed how Donald Trump won the popular vote—which we know is absolutely false. But because people are clicking on it, it's profitable.

It begs the question, how does misrepresentative information make it to the top of the search result pile—and what is missing in the current culture of software design and programming that got us here?

Facebook is another great example of how lucrative this model of organizing information without thinking about the consequences can be. Today, most people get their news through Facebook. Whenever they click some fake news story posted by Macedonian teenagers, Facebook makes money. It doesn't matter what the content is—so long as the content circulates. Virality and clicks generate revenue for the platform.

I've given interviews and talks about this in the context of the virality of videos of police murders of unarmed African Americans. No matter that these videos contribute to a culture of sustained trauma and violence against African Americans, they are heavily trafficked because they generate a lot of views and a lot of web traffic.

In the end, these companies are beholden to their shareholders and they're committed to maximizing profit before all else—and these videos contribute to a profitable bottom line. But we need more than just maximizing profit as the value system in our society.

So, engineers may not be malicious, of course. But I don't think they have the requisite education in the humanities and social sciences to incorporate other frameworks into their work. And we see the outcomes of that.

I think things will only spiral out of control, and we will increasingly see automated decision-making systems and other forms of artificial intelligence emerge as a civil and human rights issue that we cannot ignore.

Sarah:

For the past twenty years, the public has been told that the internet is a place of increased democracy and participation. And the public still largely labors under this image of the internet as a great democratizing force—an image actively encouraged by the tech industry.

In reality, the internet is a series of privatized islands with their own private governance. It's not a democratic place.

But the public doesn't fully realize that—and this confusion greatly impedes people's ability to have a critical lens on what's going on. The undemocratic quality of the internet—its domination by the profit motive—is obfuscated by the tech industry on purpose.

That said, I do think we've reached a point where more people are starting to question the mainstream view of the internet. They're saying, "Hmm, for years I've been bombarded by the son of some person I went to high school with on social media. I don't even know who that person is. That's mildly annoying. I've had to listen to my uncle rant at length about politics. That's a bit annoying too."

But they're also realizing that maybe there are much larger stakes at play. That all those things that are annoying and irritating might actually be truly dangerous. After the election, for instance,

people began to recognize that they're not all operating in the same informational landscape. That their online worlds are fracturing, divided up by the algorithms of these platforms.

As a result, there's a new appetite for information about how these platforms work. There's a new willingness to question the myths that we've been told about the internet. And whether you start to pull the thread at Google search or content moderation or something else, the entire ball of thread starts to unravel very quickly.

It sounds like what you're both saying is that profit-driven platforms produce algorithmic racism, algorithmic sexism, and misinformation. Broadly, they are producing a democratic deficit in the digital sphere. So what are some possible solutions? Should we think about trying to reform these companies from within? Should we think about regulation? Nationalization? Building alternatives?

Sarah:

The answer is yes.

Safiya:

All of the above.

Sarah:

I always go back to Jennifer Light's 1995 essay "The Digital Landscape: New Space For Women?" about online feminism. She was imagining the internet as having all this potential. It would provide new ways of being, interacting, communicating.

Since then, so much of that potential has been foreclosed by a model of rampant profiteering.

What's the first step for challenging that model? Where do we start?

Sarah:

I've sat in meetings with many high-level people from the tech industry, and these individuals were adamant that the firms they represented were tech firms—not media companies. This was said over and over again. Then, in the next breath, these individuals went on to explain all the ways in which their not-media company was manipulating the information on their platform.

We have to push back on these firms and get them to have a more honest depiction of what and who they actually are.

It's obviously very self-serving for these companies to say they're tech companies. They create and solicit and circulate media, day in and day out, but they don't want to take responsibility for it. One thing the public can do is say, "We know you're a media company and we're going to treat you like one." That's where regulation and law and other kinds of social accountability come in.

Safiya:

I don't think tech companies are equipped to self-regulate any more than the fossil fuel industry.

Certainly, our hyperinvestment in digital technologies has profound social, political, and environmental consequences. We're only beginning to scratch the surface in understanding these consequences, and what it means to be building these huge communications infrastructures.

Sarah:

Policymakers like to say, "The technology is too complicated, so I can't understand it. And if I can't understand it, I can't regulate it." The industry encourages that impression. But in fact, when states push hard enough on industry, industry finds a solution. It's not so complicated after all.

In 2000, the French government told Yahoo that it couldn't allow people to sell Nazi memorabilia on its auction site. Selling Nazi memorabilia is illegal in France. Yahoo refused. They argued that they couldn't possibly determine where their users were geographically located. And the French government said, "Guess what, you're going to do that—or you're not going to operate in France."

So Yahoo figured out that they could geolocate users pretty well by IP address. Which is why we now can't watch Netflix in some countries.

Government has the power to push firms. When you tell them they're going to lose access to an entire marketplace, they're going to make it happen.

State influence cuts both ways, of course. In other kinds of markets, companies cut different kinds of deals in order to uphold oppressive regimes. In Turkey, Facebook routinely takes down any material that relates to the Kurdistan Workers' Party (PKK), because that's a condition of them doing business in Turkey.

Technology can be manipulated to serve a variety of different visions. The political question is which vision will win.

This piece appears in Logic's third issue, "Justice." To order the issue, head on over to our store—or better yet, subscribe!