Algorithmic accountability

Algorithms are designed to make our lives easier. The problem is, they're designed by us.

Apr 30, 2017 | Megan Rose Dickey

When Netflix recommends you watch "Grace and Frankie" after you've finished "Love," an algorithm decided that would be the next logical thing for you to watch. And when Google shows you one search result ahead of another, an

biased algorithm that a real person developed made that call.

Algorithms are sets of rules that computers follow in order to solve problems and make decisions about a particular course of action. Whether it's the type of information we receive, the information people see about us, the jobs we get hired to do, the credit cards we get approved for, and, down the road, the driverless cars that either see us or don't see us, algorithms are increasingly becoming a big part of our lives.

But there is an inherent problem with algorithms that begins at the most base level and persists throughout its adaption: human bias that is baked into these machine-based decision-makers.

You may remember that time when Uber's self-driving car ran a red light in San Francisco, or when Google's photo app labeled images of black people as gorillas. The Massachusetts Registry of Motor Vehicles' facial-recognition algorithm mistakenly tagged someone as a criminal and revoked their driver's license. And Microsoft's bot Tay went rogue and decided to become a white supremacist. Those were algorithms at their worst. They have also recently been thrust into the spotlight with the troubles around fake news stories surfacing in Google search results and on Facebook.

But algorithms going rogue have much greater implications; they can result in life-altering consequences for unsuspecting people. Think about how scary it could be with algorithmically biased self-driving cars, drones and other sorts of automated vehicles. Consider robots that are algorithmically biased against black people or don't properly recognize people who are not cisgender white people, and then make a decision on the basis that the person is not human.

"driven" by algorithms mowing down black people because they don't recognize black people as human. Or algorithmic software that predicts future criminals, which just so happens to be biased against black people.

A variety of issues can arise as a result of bad or erroneous data, good but biased data because there's not enough of it, or an inflexible model that can't account for different scenarios.

The dilemma is figuring out what to do about these problematic algorithmic outcomes. Many researchers and academics are actively exploring how to increase algorithmic accountability. What would it mean if tech companies provided their code in order to make these algorithmic decisions more transparent? Furthermore, what would happen if some type of government board would be in charge of reviewing them?

Whatever approach is taken to ensure bias is removed from the development of algorithms, it can't dramatically impede progress, DJ Patil, former chief data scientist of the U.S., tells me. Solutions can only be implemented, and therefore effective, if tech companies fully acknowledge their roles in maintaining and perpetuating bias, discrimination and falsehoods, he adds.

developing new technologies, there needs to be more diversity on the team behind these algorithms. There's no denying that, Patil says, but the issue is scalability across the whole realm of diversity.

"There's diversity of race, there's diversity of religion, there's diversity with respect to disability. The number of times I've seen somebody design something where if they made a slight decision choice you could design for a much broader swath of society — just so easy, just to make a slight design change, but they just didn't know. And I think one of the challenges is that we don't have scalable templates to do that."

Google, for example, determines what many people see on the internet. As Frank Pasquale writes in his book, "The Black Box Society: The Secret Algorithms That Control Money and Information," Google, as well as other tech companies, set the standards by which all of us are judged, but there's no one really judging them (141, Pasquale).

When you conduct a Google Image search for "person," you don't see very many people of color, which perpetuates the normalization of whiteness and reinstills biases around race. Instead, you'll see many pictures of white men, says Sorelle Friedler, affiliate at the Data & Society Research Center and ex-Googler who worked on X and search infrastructure.

"That is perhaps representative of the way that 'person' is broadly used in our society, unfortunately," Friedler says. "So then the question is, is it appropriate for that sort of linguistic representation to make its way to image search? And then Google would need to decide, am I okay with black people only being represented only if you search specifically for black people? And I think that that's a philosophical decision about what we want our society to look like, and I think it's one that's worth reckoning with."

inputting" is acceptable in this scenario. But search queries related to the Holocaust or suicide have prompted Google to intervene.

Algorithms determine Google's search results and suggestions. Some of Google's algorithmic fails are more egregious than others, and sometimes Google steps in, but it often doesn't.

"If you search for ways to kill yourself, you're directed toward a suicide hotline," Robyn Caplan, a research analyst at Data & Society, tells TechCrunch. "There are things Google has deemed relevant to the public interest that they're willing to kind of intervene and guard against, but there really isn't a great understanding of how they're assessing that."

Earlier this year, if you searched for something like, "is the Holocaust real?," "did the holocaust happen" or "are black people smart?" one of the first search results for both queries was pretty problematic. It wasn't until people expressed outrage that Google decided to do something.
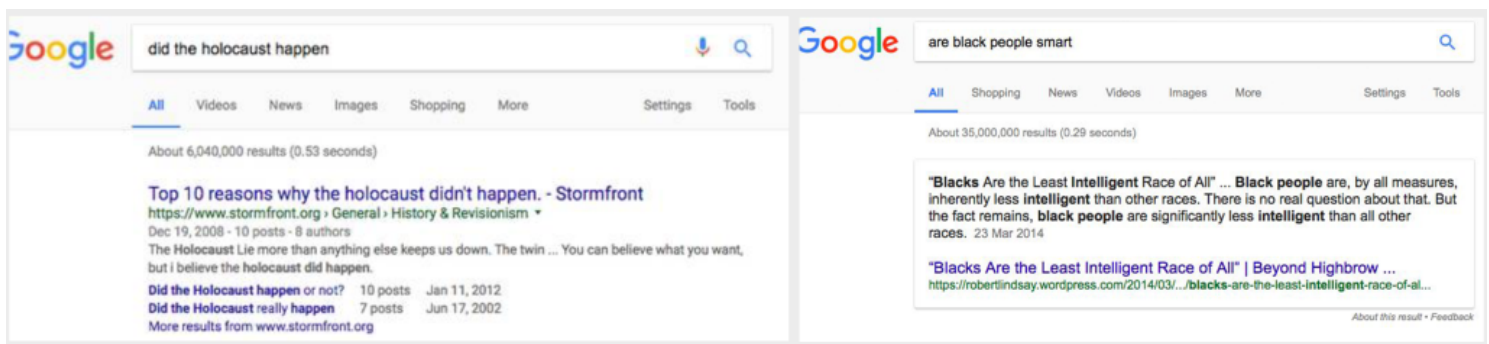
"When non-authoritative information ranks too high in our search results, we develop scalable, automated approaches to fix the problems, rather than manually removing these one-by-one," a Google spokesperson tells TechCrunch via email. "We are working on improvements to our algorithm that will help surface more

In addition to its search results and suggestions, Google's photo algorithms and the ads it serves have also been problematic. At one point, Google's photo algorithm mistakenly labeled black people as gorillas.

Google launched Photos in May 2015 to relatively good reception. But after developer Jacky Alciné pointed out the flaw, Bradley Horowitz, who led Google Photos at the time, said his inbox was on fire.

"That day was one of the worst days of my professional life, maybe my life," Horowitz said in December.

"People were typing in gorilla and African-American people were being returned in the search results," Horowitz said. How that happened, he said, was that there was garbage going in and garbage going out — a saying he said is common in computer science. "To the degree that the data is sexist or racist, you're going to have the algorithm imitating those behaviors."

Horowitz added that Google's employee base isn't representative of the users it serves. He admitted that if Google had a more diverse team, the company would have noticed the problems earlier in the development process.

Another time, Google featured mugshots at the top of search results for people with "black-sounding" names. Latanya Sweeney, a black professor in government and technology at Harvard University and founder of the Data Privacy Lab, brought this to the public's attention in 2013 when she published her study of Google AdWords. She found that when people search Google for names that traditionally belong to black people, the ads shown are of arrest records and mugshots.

What's driving mistakes like this is the idea that the natural world and natural processes are just like the social world and social processes of people, says

"And it's this assumption that if we can develop an algorithm that picks out all of the rocks as rocks correctly, we can have one that classifies people correctly or in a useful way or something like that," Pasquale says. "I think that's the fundamental problem. They are taking a lot of natural science methods and throwing them into social situations and they're not trying to tailor the intervention to reflect human values."

When an algorithm produces less-than-ideal results, it could be that the data set was bad to begin with, the algorithm wasn't flexible enough, the team behind the product didn't fully think through the use cases, humans interacted with the algorithm enough to manipulate it, or even all of the above. But no longer are the days where tech companies can just say, 'Oh, well it's just an app" or "Oh, we didn't have the data," Patil says. "We have a different level of responsibility when you're designing a product that really impacts people's lives in the way that it can."

While algorithms also have vast potential to change our world, Google's aforementioned fails are indicative of a larger issue: the algorithm's role in either sustaining or perpetuating historic models of discrimination and bias or spreading false information.

"There is both the harm data can do and the incredible opportunity it has to help," Patil says. "We often focus on the harm side and people talk about the way math is — we should be scared of it and why we should be so afraid of it.

"We have to remember these algorithms and these techniques are going to be the way we're going to solve cancer. This is how we're going to cure the next form of diseases. This is how we're going to battle crises like Ebola and Zika. Big data is the solution."

A barrier to tackling algorithmic issues that pertain to content on the internet is Section 230 of the Communications Decency Act, which states, "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."

It made it possible for tech companies to scale because it relieved platforms of any responsibilities for dealing with illegal or objectionable conduct of its users. The Electronic Frontier Foundation calls it "one of the most valuable tools for protecting freedom of expression and innovation on the internet."

If this law didn't exist, we could essentially deputize Google and other tech companies to operate as censors for what we consider to be objectionable speech. Something like that is happening in Europe with The Right to be Forgotten.

"American scholars and policy people are somewhat terrified because basically what has happened there is practically speaking, Google has become the arbiter of these claims," says Solon Barocas, a postdoc researcher at the NYC Lab of Microsoft Research and member of the Society, Ethics, and AI group. "It's not like a government agency is administering the decisions of what should be taken down. Instead, it said 'Google you have a responsibility to do this' and then Google does it themselves. That has frightened a lot of Americans."

But given the existence of Section 230 of the CDA and the fact that it provides many protections for platforms, it may be difficult to use legislative means in the

Outside the U.S., however, legislation could affect the way these tech companies operate inside America, Caplan says. In Germany, for example, the government has drafted a law that would fine social networks up to 50 million euro for failing to remove fake news or hate speech.

Meanwhile, the European Union's digital chief, Andrus Ansip, warned Facebook earlier this year that while he believes in self-regulatory measures, he's ready to take legislative action if it comes to that.

"What we've seen in the past is that these types of policies that take place outside of the U.S. do have a pretty big role in shaping how information is structured here," Caplan says. "So if you look at Google's autocomplete algorithm, you see a similar thing — that different auto-completions aren't allowed because of libel cases that happened abroad, even though Google is protected here. Those kinds of policies proposed by countries with a clear understanding of what they're willing to regulate media-wise may have an interesting impact here."

Even if Section 230 stays in place, and it most likely will, there are ways to reevaluate and reprogram algorithms to make better decisions and circumvent potential biases or discriminatory outcomes before they happen.

W different number of scenarios an algorithm may have to deal with, there's no straightforward, cut-and-dried solution to every company's algorithmic issues. But researchers have proposed several potential methods to address algorithmic accountability.

Two areas developing rapidly are related to the front- and backend process, respectively, Barocas tells me. The front-end method involves ensuring certain values are encoded and implemented in the algorithmic models that tech companies build. For example, tech companies could ensure that concerns of discrimination and fairness are part of the algorithmic process.

"Making sure there are certain ideas of fairness that constrain how the model behaves and that can be done upfront — meaning in the process of developing that procedure, you can make sure those things are satisfied."

On the backend, you could imagine that developers build the systems and deploy them without being totally sure how they will behave, and unable to anticipate the potential adverse outcomes they might generate. What you would do, Barocas says, is build the system, feed it a bunch of examples, and see how it behaves.

Let's say the system is a self-driving car and you feed it examples of pedestrians (such as a white person versus a black person versus a disabled person). By analyzing how the system operates based on a variety of inputs/examples, one could see if the process is discriminatory. If the car only stops for white people but decides to hit black and disabled people, there's clearly a problem with the algorithm.

"If you do this enough, you can kind of tease out if there's any type of systematic bias or systematic disparity in the outcome, and that's also an area where people are doing a lot of work," Barocas says. "That's known as algorithmic auditing."

"In terms of algorithmic accountability, a first step is transparency with respect to data and algorithms," Pasquale says. "With respect to data, we can do far more to ensure transparency, in terms of saying what's going into the information that's guiding my Facebook feed or Google search results."

So, for example, enabling people to better understand what's feeding their Facebook news feeds, their Google search results and suggestions, as well as their Twitter feeds.

"A very first step would be allowing them to understand exactly the full range of data they have about them," Pasquale says.

The next step is something Pasquale calls qualified transparency, where people from the outside inspect and see if there's something untoward going on. The last part, and perhaps most difficult part, is getting tech companies to "accept some kind of ethical and social responsibility for the discriminatory impacts of what they're doing," Pasquale says.

The fundamental barrier to algorithmic accountability, Pasquale says, is that until we "get the companies to invest serious money in assuring some sort of both legal compliance and broader ethical compliance with personnel that have the power to do this, we're not really going to get anywhere."

Pasquale says he is a proponent of government regulation and oversight and envisions something like a federal search commission to oversee search engines and analyze how they rank and rate people and companies.

Friedler, however, sees a situation in which an outside organization would develop metrics that measure what they consider to be the problem. Then that organization could publicize those metrics and its methodology.

accountable," Friedler tells me. "I also would like to believe that many of these tech companies want to do the right thing. But to be fair, determining what the right thing is is very tricky. And measuring it is even trickier."

Algorithms aren't going to go away, and I think we can all agree that they're only going to become more prevalent and powerful. But unless academics, technologists and other stakeholders determine a concrete process to hold algorithms and the tech companies behind them accountable, we're all at risk.

19 Comments

Sort by Oldest

Add a comment...

Donald Winiecki
Collaboration between stakeholders is the surest way through this newly created mess. Sociologists, Public Policy, ...as well as Computer Science and corporate interests, should be in on the action.

Like · Reply ·　3 · Apr 30, 2017 8:27pm

Andrew Powys · Teacher at GroupeSet
Great writing, thanks!

Like · Reply · May 1, 2017 5:11am

Luke Hsieh · Toronto, Ontario
After reading the first few paragraphs, I can not continue. Please do more research on a subject which is not your speciality before you start talking about it in public.

Like · Reply ·　4 · May 1, 2017 9:09am · Edited

> **Tom Parrett** · Freelance writer at Self-Employed
> If only specialists were allowed to write, we'd never read about ideas. Oh, and ". . . subject 'that' is not your specialty . . . "
>
> Like · Reply ·　5 · May 1, 2017 10:44am

Gaston Ferrari
I'm sorry, but if someone has to ask Google if the holocaust did happen then there's something really wrong with them. Google is a search engine, it's supposed to deliver information that relates to the words you entered, not judge if the content of those articles is correct or not. If it worked that way we'd probably only get a few hits per search.