



How Does Machine & Algorithm Bias Affect All of Us?

Fast Facts

Access to good information is necessary for us to make informed decisions, so search engines and software use patterns and mathematics to guess what users want. However, the algorithms that machines use can also reflect gender, class, and racial bias. Search engines may favor results with a financial incentive. Unintentional (and perhaps intentional) machine bias affects our everyday lives. Below are fast facts on how this happens.

In what ways could a computer be biased? Can't algorithms avoid bias by not collecting irrelevant information?

- [Friedman and Nissenbaum's 1996 study](#) was one of the first to show that software can systematically and unfairly discriminate against certain individuals and groups. **Pre-existing bias** in society can affect system designs, **technical bias** can occur due to technical limitations, and **emergent bias** can arise after software is completed and released.
- **Artificial Intelligence (AI)** bots created to connect jobs and job applicants didn't collect information about the sex of applicants but did [exclude most female applicants](#). A bot can be programmed to make connections between past successful male employees and the language male applicants typically use on resumes. If programmed to collect applicants who are more likely to apply, bots often rejected profiles of female job seekers who statistically apply for a limited pool of jobs they're qualified for and instead prioritized men who applied widely for all kinds of jobs.
- **Risk assessments** calculated by an **algorithm**, or set of programmed computer calculations, are given to judges in court during sentencing to predict whether criminals will reoffend. Whether the algorithms work has only been examined in a few studies, primarily by those selling the software. When ProPublica obtained [the risk scores of more than 7,000 people](#) arrested in the same Florida county, only 20 percent of the people predicted to commit violent crimes had actually done so, and only 61 percent had committed any crime including misdemeanors. Black defendants were consistently mislabeled as more likely to reoffend and white defendants mislabeled as less likely, though they were asked about home life, employment, and education rather than race.
- The algorithm that helps human Transportation Security Administration (TSA) agents assess scanned bodies for suspicious activity has two gender options, meaning that [transgender travelers](#) are far more likely to be flagged as suspicious in an airport because the millimeter wave scanner's data on a typical body type doesn't match.

What about just retrieving the information I want to know? Can a computer's algorithms have bias then?

- Search engines are defined as biased if they favor results of their partners and bury competitors. [Google has had several antitrust](#) (laws that determine legal business practices) battles over bias.
 - Google entered a [two-year trial that ended in 2013](#) to determine whether their **Universal Search function** (including maps, news, and other options in results rather than just 10 blue links) was harming competitors by defaulting to Google products, such as Google Places rather than TripAdvisor. Google argued that other search engines such as Yahoo! and Bing have copied the Universal Search interface and favor their own results just as much. This practice continues primarily because [the suggested solutions](#) proved unconstitutional or impractical:
 - To have the government somehow monitor Google algorithms for **neutrality**
 - To constrain Google's ability to partner with websites and apps
 - To require Google to release more information about how their algorithms work, which would allow competitors to copy and spammers to manipulate them

Need help finding a resource?

Ask library staff for help if you are unable to find or access a resource. We're glad to help you find the information you are looking for.





Fast Facts

- To invert [Fair Use Doctrine](#) (legally using quotes from content for reporting purposes) to restrict Google from [crawling](#) webpages (searching the Internet for content) to use in search results, but letting other search engines crawl Google's content
- To restrict Google from Universal Search and return to the ten blue links model
- Similarly, in April 2021 a United States Senate Subcommittee [held a hearing to discuss Google's and Apple's](#) roles in the mobile app ecosystem, and whether they are making competing apps such as Spotify, Match Group, and Tile difficult to access or unaffordable in their app stores.
- Users often assume that search engines are impartial because they are automated, but search engine algorithms are created by humans and most companies [include a human element](#) to search results as well. Google and others have admitted that employees manually **blacklist** or **whitelist** certain websites, especially if the algorithm didn't rank the company as expected. Human operators also evaluate and can remove information at the request of governments.
- There are over [200 different factors used by Google](#) to calculate a page's rank in search results. Not all of the factors are public knowledge, and neither is how they are weighted (how factors are valued versus others). Google's worldwide data centers are also not in sync, so data center algorithms may differ depending on location.
- [Algorithms that make predictions](#), such as Bing's incorrect predictions about the vote to leave the European Union or Google's Flu Trend predictions that were off by 140 percent, sometimes fail because they can only collect data that people put on the internet, and enough people act offline and not online to skew results.
- Social media algorithms often [weigh novelty over popularity](#). A new term is more likely to become a "Twitter Trend" than a widespread movement that has existed for longer, even when more people are currently tweeting about the movement. This pushes the focus to viral memes and makes it more difficult to track slowly building news. It's also why you're more likely to order from businesses that post while you're online. Search engines work the opposite way—a **domain** that has been established for longer will be ranked higher in results.
- **Personalization algorithms** collect information on what [an individual user does online](#) in order to make predictions about what information will be most useful to them. When you search for "Spokane businesses," the search engine will look at your location and give you results for Spokane, Washington, rather than Spokane, South Dakota. Google uses signals like your location, previous search keywords, and social media contacts; whereas Facebook and YouTube use your interaction with others (**social gestures**) and what you've watched, liked, and shared.
 - Personalization algorithms hide information that they [predict will be unwanted](#), which can mean that you stop seeing social media posts from friends you don't contact very often or that certain websites will never show up in your search—all without you knowing you're missing them.
 - A **recommendation mechanism** [collects information](#) on what you've purchased or read, looking for similar items and curating online content such as ads specifically targeting you.
 - Rather than returning results that might go against a user's current beliefs, even if accurate, and prompt them to decide on a new course of action or thought, personalized algorithms return results that are comfortable for the user. When it comes to news and facts, users might end up in "**filter bubbles**" that give them [information they want to hear over information that is correct](#).

Need help finding a resource?

Ask library staff for help if you are unable to find or access a resource. We're glad to help you find the information you are looking for.

