

What is Algorithmic Bias?

Compelling Question: Why is algorithmic bias problematic and what should we do about it?

Grade Level		Time	Common Core Standards
K-2	3-5	45–50 Minutes	Reading: R1 Writing: W2 Speaking & Listening: SL1, SL2 Language: L4, L5
MS	HS		



Web Related Connections

Lessons

- [Addressing Hate Online: Countering Cyberhate with Counterspeech \(En español\)](#)
- [Everyday Bias](#)
- [Microaggressions in Our Lives](#)
- [Teens, Tech, Connect: How Technology Impacts Teenagers' Friendships](#)

Other Resources

- [Creating an Anti-Bias Learning Environment](#)
- [Navigating a Digital World: Tips for Youth \(En español\)](#)
- [Race Talk: Engaging Young People in Conversations about Race and Racism](#)

Key Words

(See ADL's [Education Glossary Terms](#).)

activism
algorithm
artificial intelligence
bias
discrimination
disputed
drone
inclusion
slurs
social justice
stereotypes
underrepresented

LESSON OVERVIEW

Algorithms—as basic as a recipe or as complex as the digital codes that create targeted ads in our social media feeds—are part of our daily lives. Anyone who uses social media such as Instagram or Facebook, search engines like Google, or almost any online platforms come into contact with algorithms regularly. These algorithms, which are often designed for good use, are not always neutral or objective in how they calculate, sort and present data. Algorithms can deepen the echo chamber by spreading fake information and can discriminate against people or perpetuate bias by constructing data sets based on perceived identities and stereotypes.

This lesson provides an opportunity for students to understand the role algorithms play in our everyday lives, explore how algorithmic bias functions in society and learn how to challenge it.

LEARNING OBJECTIVES

- Students will be able to define algorithms and algorithmic bias.
- Students will examine examples of algorithmic bias in their own lives, the lives of others and within societal institutions.
- Students will explore the difference between someone's actual and perceived identity through a social media analysis.

MATERIALS & PREPARATION

- *Your New Favorite Song has Been Chosen by an Algorithm* video (2017, 2½ mins., NBC News, www.youtube.com/watch?v=p7nn_jVQvWc)
- [Camps teach students how to fight racial bias in artificial intelligence](#), one copy for each student
- *How I'm fighting bias in algorithms* Ted Talks video (2016, 9 mins., Joy Buolamwini, www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms?utm_campaign=tedsread&utm_medium=referral&utm_source=tedcomshare)

PROCEDURES

Video Viewing: What is Algorithmic Bias?

1. Start the lesson by asking: *What is an algorithm?* Show the video [Your New Favorite Song Has Been Chosen by An Algorithm](#).

2. Engage students in a brief discussion by asking the following questions:
 - What did you learn from the video?
 - What are algorithms?
 - How are algorithms useful?
3. Elicit a definition of **algorithm** as a detailed set of instructions to reach a result based on given inputs, data or information. Explain that an algorithm can be digital or non-digital. In computing, programmers write algorithms that instruct the computer or digital platform how to perform a task.

Elicit from students that algorithms can be something as simple as following a recipe to prepare a specific meal. They can also be more complex and exist in digital spaces.
4. Ask students: *Can anyone share an example of an algorithm, either one that uses computer technology or not?* If students don't share any examples, provide these examples:
 - Algorithms define what ads appear on your Facebook page or pop-ups.
 - Algorithms suggest who to follow on Twitter and Instagram.
 - Algorithms select music just for you to listen to on Spotify.
 - Algorithms recommend books on Amazon and movies/shows on Netflix.
5. Explain that algorithms can be used to understand people's interests, personalities and preferences. However, they are not always neutral or "objective." In fact, algorithms can be used to define people in stereotypical ways based on aspects of identity (or perceived identity) and in this way, they can be biased and can spread and perpetuate bias.
6. Ask students: *What is bias?* Define **bias** as an inclination or preference either for or against an individual or group that interferes with impartial judgment.
7. Ask students: *What is algorithmic bias?* Elicit/define **algorithmic bias** as follows (this definition is adapted from [Wikipedia](#)):

Algorithmic bias describes systematic and repeated errors in a computer system that create discriminatory outcomes, such as privileging one group of users over another, based on aspects of perceived identity. Bias can emerge due to a variety of factors, including but not limited to: the design of the algorithm; unintended or unanticipated use or decisions relating to the way data is coded, collected selected or used; or lack of sensitivity to identity factors that contribute to bias in the design.
8. Share the following examples of algorithmic bias:
 - In the cash bail system within the criminal justice system, computer algorithms are used to determine who should be detained and who should be released. Under new policies in some states, "risk assessment" algorithms recommend to judges whether a person who's been arrested should be released. These advanced mathematical models appear to not factor in race. However, they are based on factors that are not only highly correlated with race and class, but are also significantly influenced by pervasive bias in the criminal justice system.
 - In terms of housing and mortgage lenders, recent studies have shown that lenders earn 11–17% higher profits by charging buyers who are African American and Latino higher rates. They pay 5.6 to 8.6% basis points higher interest on home purchase loans than their white and Asian counterparts with similar credit profiles. Researchers said the racial disparities could result from algorithms that charge higher rates to borrowers who may be less likely to shop around.
 - In the employment arena, a study revealed that Google's online advertising system showed an ad for high-income jobs to men much more often than it showed the ad to women.
9. Engage students in a discussion by asking the following questions:
 - What are your thoughts about this information?
 - Why and how do you think this happens?

- What impact does this have on the people involved? What impact does it have on society?
- How does this relate to our civil liberties and civil rights?
- What can be done about this type of systemic algorithmic bias?

Reading Activity

1. Distribute a copy of the article, "[Camps teach students how to fight racial bias in artificial intelligence](#)" and give students 10 minutes to read it silently. As they read, have students underline any words, phrases and quotes that stand out to them.

Optional: Assign the reading of this article the night before for homework.

2. After reading, engage students in a brief discussion by asking some or all of the following questions:
 - What did you learn that you didn't know before?
 - In what ways is artificial intelligence part of our everyday lives?
 - What are some examples of racial bias found in artificial intelligence?
 - How does the lack of representation among women and people of color factor into this situation?
 - Do you think creators of these programs have a responsibility to their community and to society at large? How so?
 - How has your thinking shifted by reading this article?

Analytical Activity: Perceived Identity vs. Actual Personality

Note: If you allow students to use smartphones/devices in the classroom for school-related activities, this activity can be done in class and if not, assign steps 4–6 as a homework assignment.

1. Explain to students that with some social media platforms, there is a function where the platform compiles information about you (e.g., likes, videos you've watched, pages you've visited) and makes a determination of what identity group(s) you belong to (e.g., race, gender, age, ethnicity, family status, sexual orientation, etc.). The platform then uses those aspects of your perceived identity to show you specific ads based on what they perceive people in those identity groups want to see. For example, if their algorithm decides you are a middle aged, single Asian-American woman, they may place more ads in your news feed for products they think are of interest to Asian-American, single, middle aged women. Ask students:
 - What do you think about this?
 - What are the different levels of bias and stereotyping in this process?
 - What impact does this have on the individual person?
 - What impact does it have on society?
2. Explain to students that they will do an activity to experience the difference between using people's information on social media to understand that person's personality traits, characteristics, interests and preferences as opposed to using a perceived identity as illustrated in the example above.
3. Have each student select a person whose social media accounts they will analyze. This should be someone they don't know personally such as a celebrity, athlete, politician or social media influencer. They should also decide which social media site(s) they will examine (Instagram, Facebook or Twitter).
4. With the whole class, brainstorm a list of items from social media they will use to construct a personality profile. The list might include:
 - Photos they posted
 - Hashtags they use
 - What's in their profile

- Quotes they've posted/shared
 - Articles they share
 - Who they follow/who follows them
 - Emojis, memes and gifs they post
 - What they talk about (products, ideas, people, points of view, interests, hobbies, etc.)
 - What they "liked"
 - Groups they follow
 - Surveys they organized or did
 - How often they post
5. Give students 10–15 minutes to look through the person's account and build their profile on the person. Remind them that they are going to come up with personality traits, characteristics, moods, attitudes, etc. and not focus on the person's social identity groups.
 6. After coming up with these traits, tell students they are going to do a "quick write," which means they will have five minutes (timed) in order to construct a short profile about the person that describes what they've learned about the person.
 7. If time permits, have some students share their "quick write" profile with the class.
 8. Engage students in a brief discussion by asking the following questions:
 - What was it like to do that?
 - In what ways was it limiting to only use their social media?
 - What impact did it have on you to take the time to learn about the person rather than jumping to conclusions about their identity?
 - Does this change the way you think about algorithms and their categorizing of information? How so?

Emphasize the point that the social identity groups to which we belong (race, gender/gender identity, religion, sexual orientation, etc.) are critically important and relevant to who we are. However, algorithms that use perceived aspects of our identity in stereotypical and inaccurate ways are problematic and perpetuate bias, discrimination and systemic inequities.

Video Viewing (Optional)

1. Show the video [How I'm fighting bias in algorithms](#).
2. Have students turn and talk with someone sitting next to them and take turns sharing their thoughts and feelings after watching the video.
3. Reconvene the class and engage them in a group discussion by asking the following questions:
 - What are your thoughts and feelings after watching the video?
 - How does Joy Buolamwini define algorithmic bias? How has algorithmic bias directed impacted her?
 - How does lack of racial and other identity group representation impact algorithmic bias?
 - How can this type of bias lead to discrimination?
 - What are some of Joy Buolamwini's solutions to addressing algorithmic bias?
 - How is Joy Buolamwini "fighting bias in algorithms?"

Closing

Have students identify one thing they can do to challenge or fight algorithmic bias and have them share their idea aloud.

ADDITIONAL READING AND RESOURCES

- [Algorithmic Bias](#) (Wikipedia)
- [“Algorithms learn from us, and we can be better teachers”](#) (NBC News, March 10, 2017)
- [“Artificial Intelligence Has a Problem With Gender and Racial Bias. Here’s How to Solve It”](#) (*Time*, February 7, 2019)
- [“How to Fight an Algorithm”](#) (ACLU, August 2, 2018)
- [“Mortgage algorithms perpetuate racial bias in lending, study finds”](#) (Berkeley News, November 13, 2018)
- [“New York City wants to make sure the AI and algorithms it uses aren’t biased. That’s harder than it sounds.”](#) (Vox, April 11, 2019)
- [“Some AI just shouldn’t exist”](#) (Vox, April 19, 2019)
- [“The Newest Jim Crow”](#) (*The New York Times*, November 8, 2018)
- [“Tips for Addressing Algorithmic Bias”](#) (Quirks Media, October 12, 2018)
- [“When Algorithms Discriminate”](#) (*The New York Times*, July 9, 2015)

Common Core Standards

CONTENT AREA/STANDARD
Reading
R1: Read closely to determine what the text says explicitly and to make logical inferences from it; cite specific textual evidence when writing or speaking to support conclusions drawn from the text.
Writing
W2: Write informative/explanatory texts to examine and convey complex ideas and information clearly and accurately through the effective selection, organization, and analysis of content.
Speaking and Listening
SL1: Prepare for and participate effectively in a range of conversations and collaborations with diverse partners, building on others’ ideas and expressing their own clearly and persuasively.
SL2: Integrate and evaluate information presented in diverse media and formats, including visually, quantitatively, and orally.
Language
L4: Determine or clarify the meaning of unknown and multiple-meaning words and phrases by using context clues, analyzing meaningful word parts, and consulting general and specialized reference materials, as appropriate.
L6: Acquire and use accurately a range of general academic and domain-specific words and phrases sufficient for reading, writing, speaking, and listening at the college and career readiness level; demonstrate independence in gathering vocabulary knowledge when encountering an unknown term important to comprehension or expression.



Camps teach students how to fight racial bias in artificial intelligence

By USA Today, adapted by Newsela staff on 09.06.18

Reprinted with permission from <https://newsela.com/read/artificial-intelligence-diversity/id/45426/>.

Word Count **989**

Level **1220L**



Stanford AI4ALL participants on their first day of class. Photo by: Stanford AI4ALL/Anna Wong

OAKLAND, California – High school students Aarzu Gupta and Lili Sun, through connections made at summer camp, used artificial intelligence to create a drone program to detect wildfires before they spread far. A drone is a small aircraft flown by remote control.

Artificial intelligence (AI) involves computer systems doing tasks that usually require human intelligence, such as making decisions.

Rebekah Agwunobi, a high school senior, learned enough to nab an internship at the Massachusetts Institute of Technology (MIT) Media Lab. She worked on using AI to evaluate the court system, including collecting data on how judges set bail.

Both projects stemmed from the Oakland-based nonprofit AI4All. It will expand its outreach to young underrepresented minorities and women with a \$1 million grant from Google.org, the technology giant's charity arm announced August 10.

AI is becoming increasingly commonplace in daily life. It's found in everything from Facebook's face detection for photos to Apple's iPhone X facial recognition.

The Risk In Using Unchecked AI

It's also one of the more disputed parts of technology. The late scientist Stephen Hawking and Tesla chief Elon Musk have warned human civilization is at risk from the unchecked development of artificial intelligence, which could lead to self-ruling weapons of terror.

The technology, still in its early stages, has also been criticized for built-in racial bias that can amplify existing stereotypes. That's worrisome as more companies use it for decisions such as hiring and police leverage AI-powered software to identify suspects.

MIT Media Lab researcher Joy Buolamwini, who is black, found facial recognition software could more easily identify her face when she wore a white mask. It was a result of computing rules that relied on data sets of mostly white faces.

Three years ago, Google apologized after its photo identification software mislabeled black people as gorillas. Even if the machine doesn't know what it's doing, this racist association has been used throughout history to say that black people are not human.

Microsoft also apologized after users quickly found a way to get an artificial intelligence-powered social chatbot to spew racial slurs. Chatbots are computer programs to simulate conversation with people.

Underrepresentation Of Women And Minorities

Tess Posner is the chief executive officer of AI4All. She said the problem is made worse by the fact that minority groups such as women and people of color historically have been left out of the tech industry, particularly in AI.

"We need to have people included that are going to be impacted by these technologies, and we also need inclusion to ensure that they're developed responsibly," Posner said. "Bias happens when we don't have people asking the right questions from the beginning."

Despite stated efforts to attract more women and more people of color, Google, Facebook and other tech giants have been slow to diversify their staff, and have not hired many women of color. African-American and Hispanic women make up no more than 1 percent of Silicon Valley's workforce.

Nonprofit Teaches AI To High School Students

Posner's organization believes the tech industry has to start including women and people of color at a much earlier stage. They're working to close that gap through summer camps for high school students.

AI4All, launched in 2017, is based on a two-week summer camp program out of Stanford University in Palo Alto, California.

Since then, AI4All's resources have expanded. In its first year, there were only two summer camps at Stanford University and the University of California at Berkeley. This year, it added four more at other colleges.

All of the camps are aimed at high school students who are women, people of color or low income.

Part of Google.org's grant will go toward opening more AI4All camps. The ultimate goal is to use the money to create a free, online AI course that will be accessible to anyone in the world. A course is already in the works. "We really need for AI to be made by diverse creators, and that starts with people having access to the learning opportunities to understand at its core what AI is and how it can be applied," said Hannah Peter. She is Google.org's AI4All partnership lead.

In addition to providing summer camps, AI4All offers three-month fellowships where students can develop their own ideas and pitch them to AI experts in the industry. There is also funding for students to launch independent projects.

Creating Meaningful Projects

One such project was former AI4All student Ananya Karthik's workshop, creAlte, which uses artificial intelligence to create artwork.

Other AI4All students, most still in high school, have turned their newly acquired technical skills toward current pressing issues, such as the wildfire project developed by Gupta and Sun. The two met during an AI4All fellowship this year. This idea came out of the fires that plagued northern California late last year.

Gupta and Sun appreciated the camp's talks featuring real-world examples of minority women who were able to succeed in the tech industry.

"I want to initiate change using artificial intelligence," Sun said. "I don't want to be just working on an iPhone or something like that."

Because of her experiences, Gupta said, she's looking forward to exploring a career in AI, particularly in its uses for health and medicine. She's already putting that interest to work with her internship this summer at the University of California at San Francisco, where the lab she's working at is doing research on Alzheimer's disease.

For Agwunobi, AI4All showed her how she could combine her passion for activism and social justice with her interest in technology.

At her MIT internship, Agwunobi took data gathered during the pretrial process to evaluate how key figures such as judges behave while setting bail. The goal is to arm activists with this data when pushing for prison and bail changes.

"You can work with tech and still be accountable to community solutions," Agwunobi said. AI4All "affirmed my desire to solve interesting problems that actually helped communities," she said.

Reprinted with permission from "Camps teach students how to fight racial bias in artificial intelligence," Newsela, September 6, 2018, <https://newsela.com/read/artificial-intelligence-diversity/id/45426/>

