Computer Science Ethics

15-110 – Monday 12/07

Learning Goals

- Understand the current extent of **data collection** on the internet and how data is used

- Recognize the uses and drawbacks of **facial recognition** algorithms in different contexts

- Identify the societal impact when **AI decision making** replaces human decision making due to the explainability problem

Recitation Final Exam Review Poll

The last recitation will review topics for the final exam. Fill out this poll to request specific topics for your recitation to review!

https://forms.gle/NHebYsn2LHVE7SBr8

Ethics in Computer Science

When we move from theoretical concepts of computer science to applying those theories in real life, the decisions we make have consequences.

The professional field of computer science has only recently adopted a code of ethics, and the code is not yet uniformly taught to new computer scientists or programmers. There is still much to debate over what the responsibilities of computer scientists are.

We'll discuss three areas where people debate how computing should be used in the current time: data collection, facial recognition, and AI decision-making.

Data Collection

User Data

Most applications collect data about users from various sources. We'll discuss three main categories: data provided by the **user**, data provided by the **browser/system**, and data provided by **other sources**.

As a user of the internet and various applications, you already voluntarily share a lot of data with the world!

- Internet – profile information, tweets, searches
- Applications – preferences, locations, images
- Real life – purchase history, contact info, location

Browser/System Data

Behind the scenes, your browser or phone/computer is sending additional information to the services you use.

This is not done maliciously – services can put this information to good use. However, you may be surprised by some of the data being shared.

Check out the data your browser shares here:
https://webkay.robinlinus.com/

There are plugins you can install that limit the information your browser sends, but this may also limit functionality of websites.

Other Data Sources

Cookies are used by websites to store temporary information about people using their services (like which items you've put in a shopping cart). A cookie is a small packet of data that is sent back and forth between the website and your browser.

Cookies that are shared between two or more websites are called **tracking cookies**, or just trackers. These cookies attempt to collect a portfolio of information on you, the user, by gathering information on the websites you visit. This is commonly done through ads that are placed on websites.

With enough data collected from tracking cookies and the browser, a website may be able to create a **fingerprint** that identifies you as a user. Read more here.

You can check what kinds of trackers your browser stops and what your fingerprint looks like here: https://coveryourtracks.eff.org/

Data Economy – Data Collection

Why are so many companies interested in data collection? **Data has become the economy of the internet**. Most websites are supported by advertising, and advertisers pay more for targeted ads.

Websites have a strong incentive to get the best data possible on their users, so they get paid more for advertisements. This has led to **hyper-targeting** in ads, with ads attempting to reach more niche populations.

If you have a Facebook account, try going to Settings > Ads > Ad Settings > Categories Used to Reach You. This will show you the niche groups Facebook thinks you might be a part of.

Data Economy – Selling Data

Even companies that don't rely on advertising have a use for user data – they can **sell it to other companies**. This data is aggregated by companies that can then sell portfolios of individuals to advertisers or insurance companies.

Even when companies promise not to sell individual data, it still isn't entirely private. For example, consider online DNA services like 23andMe. This site (and many others) sell **aggregated data**; though this data does not have a user's name or address attached, the genetic information is still shared.

Data and the Government

There are also concerns around how companies share data with police forces and the government. For example, the smart doorbell company Ring formed a partnership with police forces across the US to share video data, with homeowner permission.

Some governments have gotten more directly involved in large-scale data collection; in particular, China has instituted the Social Credit System, where data collected on individuals has direct impacts on their ability to interact in everyday life.

Most governments have not yet determined how to handle questions surrounding privacy and the sale and collection of data. However, there are some legal restrictions on the treatment of specific types of data.

Restrictions on Data Sharing

In the EU, the **GDPR** (General Data Protection Regulation) gives all users certain rights over their data; they must be told when data is being collected, data must be stored securely, and users have the right to obtain their data and/or ask for it to be deleted. The EU also has the **Right to Be Forgotten**, which lets users request that certain pages be removed from search results after time has passed.

In the US, data collected about minors is protected though **COPPA** (Children's Online Privacy Protection Act) and **FERPA** (Federal Educational Rights and Privacy Act). There is no general law about data privacy yet, but California passed the **CCPA** (California Consumer Privacy Act), which institutes some regulations for that state.

Protecting Your Data

If you want to protect your data online, you have a lot of options! Most browsers let you block cookies and can request that websites do not track you. You can also restrict permissions given to websites and applications on your devices.

For advanced protection, you can also use a VPN (Virtual Private Network) to connect to the internet. CMU has a VPN (though then CMU will know which websites you're accessing):

https://www.cmu.edu/computing/services/endpoint/network-access/vpn/how-to/

Facial Recognition

Image Recognition on Faces

In the Machine Learning lecture, we briefly discussed how machine learning algorithms can be applied to images, to recognize objects that occur in the images.

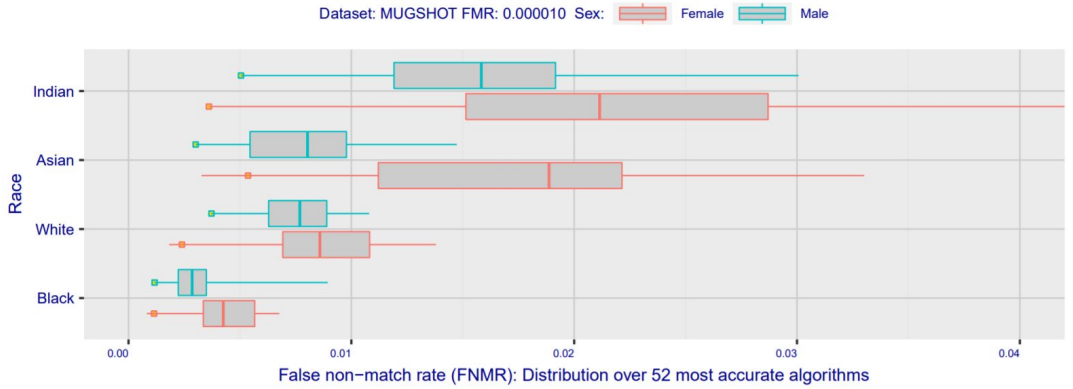
This can also be applied to people. **Facial Recognition** is used to automatically match the face of a person in a photo to their identity. (Further algorithms can even identify expressions!)

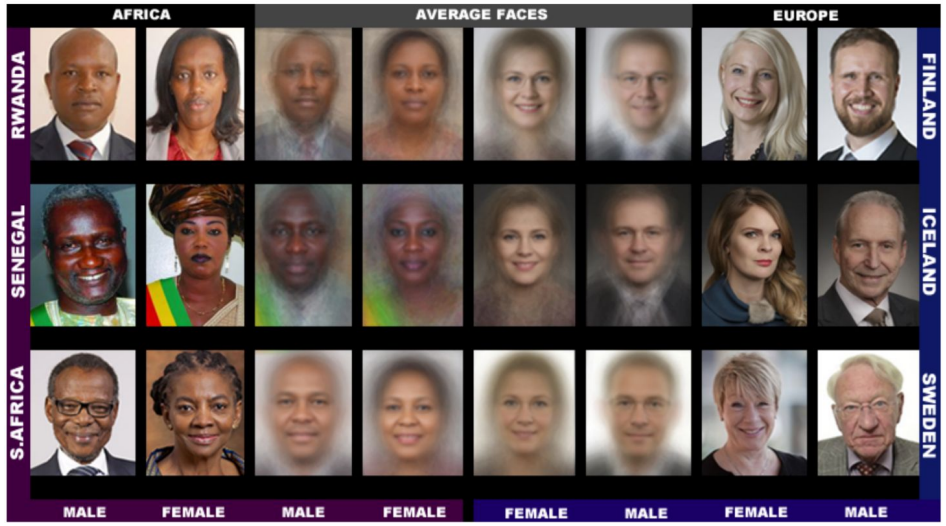
This tool has been around for a while and is used for a variety of purposes (automatic tagging, ID verification, etc.). However, some uses have been controversial.

Bias in Facial Recognition Algorithms

Recent studies that test facial recognition algorithms have shown huge variation in performance. One found that many facial recognition algorithms are "10 to 100 times more likely to inaccurately identify a photograph of a black or East Asian face, compared with a white one". Even among the best algorithms there are notable differences in recognition performance across race and gender.

One factor that could lead to this difference is **bias in the data used to train the algorithms**. An analysis showed that two popular training sets were overwhelmingly composed of lighter-skinned subjects. This is supported by the above study, which showed that algorithms developed in Asian countries performed better on Asian faces.



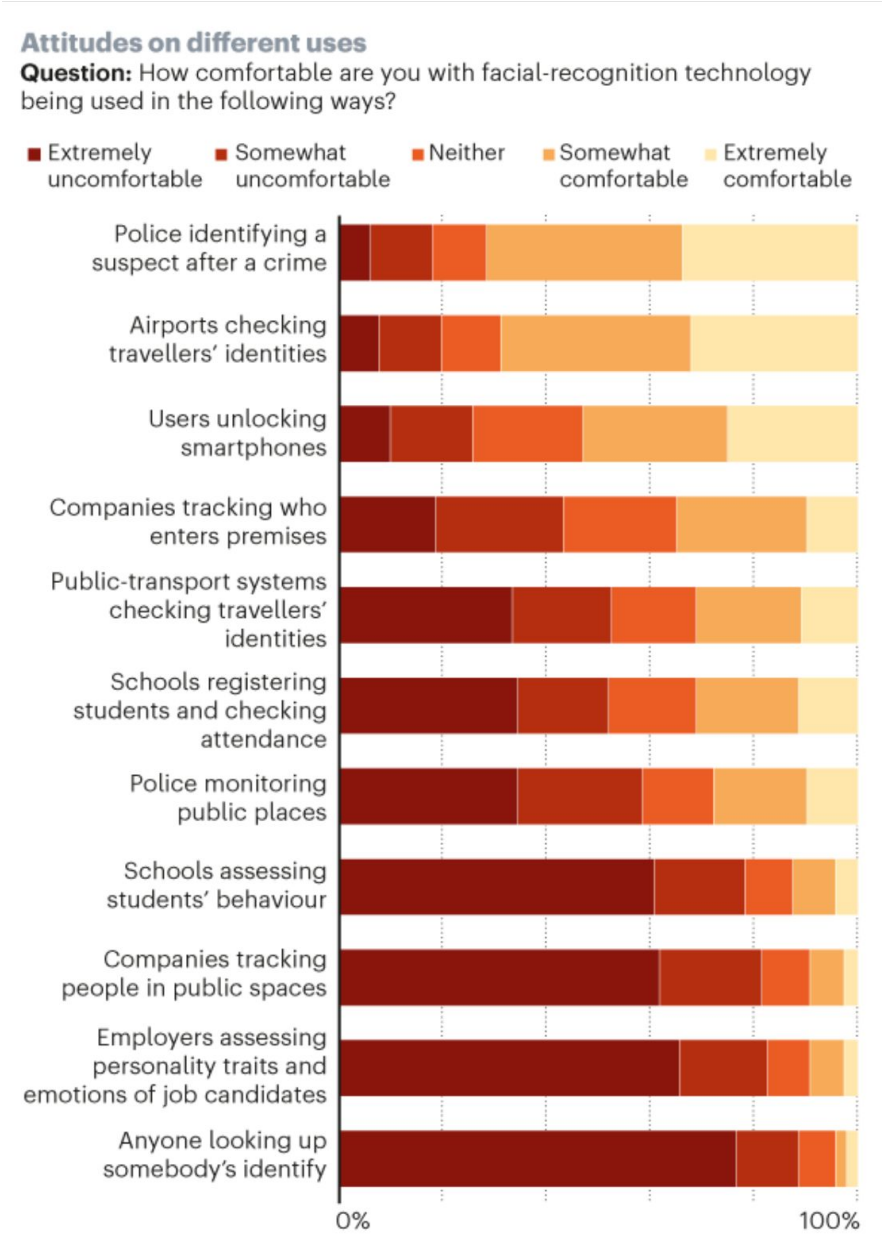


Controversial Uses

There are also concerns about some commercial uses of facial recognition even when it works well.

For example- is it okay for police to use facial recognition to identify a suspect? What about using it to monitor people in a public space?

This especially causes problems when a controversial use of facial recognition collides with a weakness in the algorithm. For example, an algorithm led to an innocent man being arrested for shoplifting.



Legislation of Facial Recognition

Some communities have gone to the length of banning facial recognition technologies from being used in certain contexts.

Several US cities (including San Francisco and Boston) have recently banned the use of facial recognition by local governments.

Beyond legislation, algorithms can be foiled by obfuscating part of the face. In particular, wearing a mask significantly decreases the accuracy of most facial recognition algorithms!

AI Decision Making

Artificial Intelligence Potential

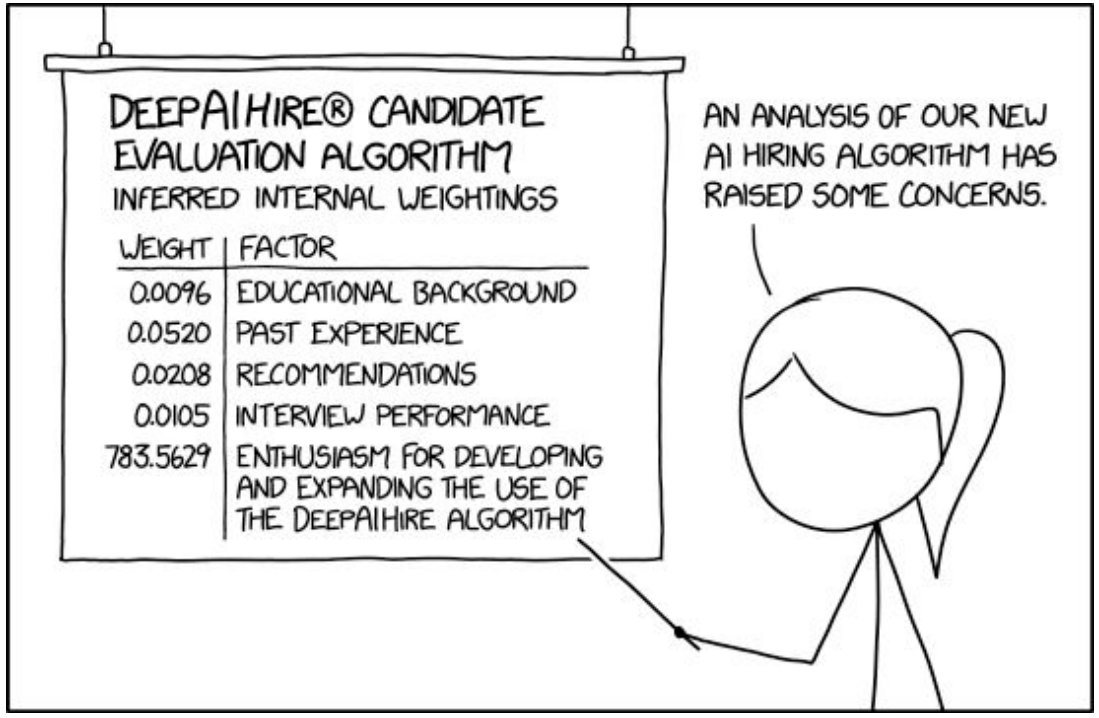
There are potentially enormous benefits to be gained by using machine learning and artificial intelligence to accomplish tasks and solve problems.

However, we must keep in mind that there are potential downsides to these algorithms as well. In particular, let's consider the problem of **explainability**; when an AI makes a decision, can it explain why?

Explainability

Decisions made by machine learning algorithms are usually based on a huge number of tiny factors. In some algorithms (like neural networks) those factors aren't named in a human-readable way.

This is a problem when the algorithm makes an important decision about a person's life, like whether they should be admitted to a school or hired for a job.

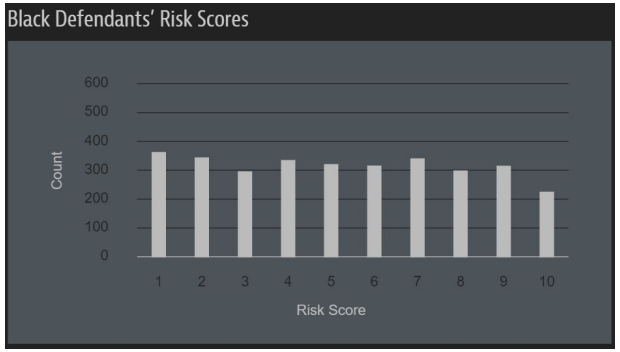


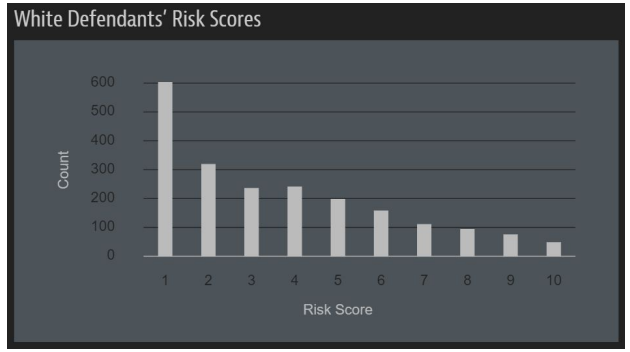
Bias in Machine Learning

Just like in facial recognition, bias in the data fed into a machine learning algorithm can lead to bias in the algorithm's results.

This has caused problems in algorithms for determining bail, which have shown systematic bias in predicting a person's likelihood to commit future crimes; this bias was correlated with race.

A similar problem was observed in an algorithm to hire engineers for Amazon, which showed bias towards hiring employees based on gender because it was trained on mostly male examples.





Prediction Fails Differently for Black Defendants

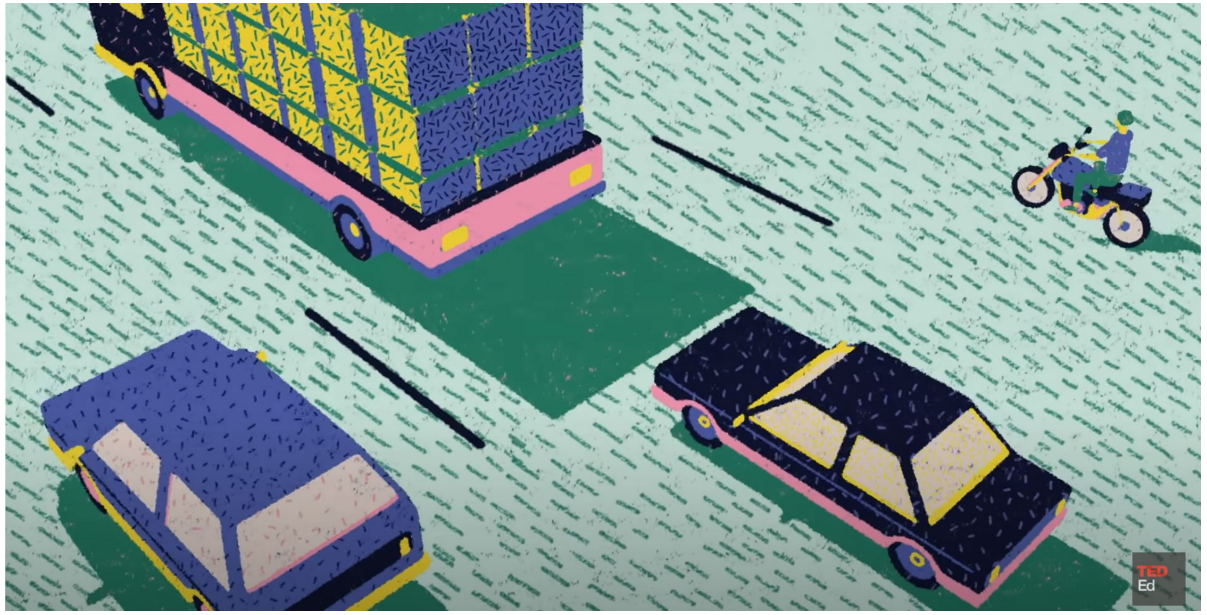
	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

Ethics in AI Design

Even if we set aside the problems related to bias in data (which obviously affect human decision making as well), there are still big ethical questions about how we should use AIs.

Consider decisions that are made by self-driving cars. If a car is put in a position where it will inevitably get into an accident, should the car protect its passenger, or should it optimize for the greatest preservation of human life? And how should this be treated legally?



Responsibility and AI

Questions about AI and responsibility extend to smaller day-to-day actions AIs may take too.

For example, Google has become a gatekeeper for much of the information in the world. If a small change to Google's search algorithm moves a small business from the first page to the second, that could have a drastic effect on the business's revenue.

This also applies to the algorithms social media networks use to decide which posts should be promoted. Studies have shown these algorithms can lead to the spread of false information.

Controversial Applications of AI

Some AI applications are inherently controversial.

- Automated surveillance systems can improve safety by locating missing children or wanted criminals.
  - But they can also lead to a police state and loss of privacy and freedom.

- Autonomous weapons systems remove human decision making from the loop, allowing robots to kill people. Ban them?
  - What if some states don't agree to a ban?

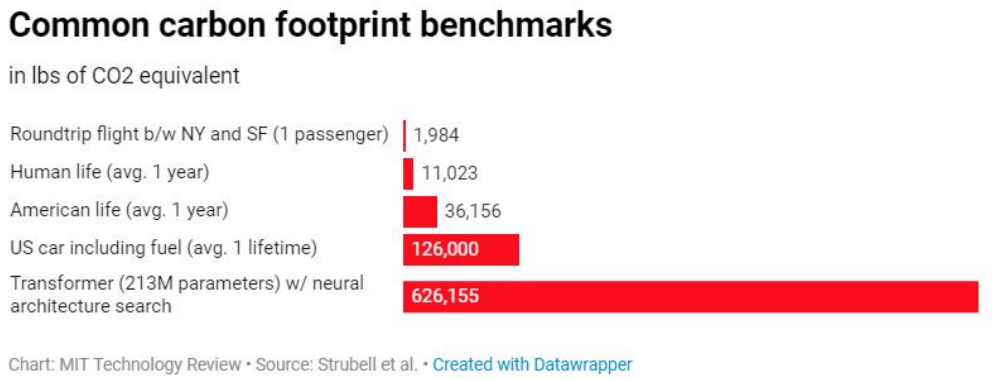
Effects on the Environment

Even when we make productive and unbiased AIs, algorithms can still have unintended side effects.

Many companies and researchers train machine learning algorithms on very large datasets to answer questions. This analysis does not come without a cost.

An enormous amount of energy is needed to run these algorithms, and in the US, that energy often has a carbon footprint. A recent study found that training a popular NLP model, The Transformer, left a gigantic carbon footprint.

On the bright side, some tech companies have pledged to go carbon negative to combat this. Other scientists are exploring new ways to make algorithms more energy efficient.



Coming Up Next: CS Future

In Wednesday's lecture, we'll discuss upcoming big ideas in computer science.

We want your input on what is most relevant and interesting! Fill out this poll to tell us what **you** want to learn about:

https://forms.gle/YhuJy1bQJ5UFXx7R6

Learning Goals

- Understand the current extent of **data collection** on the internet and how data is used

- Recognize the uses and drawbacks of **facial recognition** algorithms in different contexts

- Identify the societal impact when **AI decision making** replaces human decision making due to the explainability problem