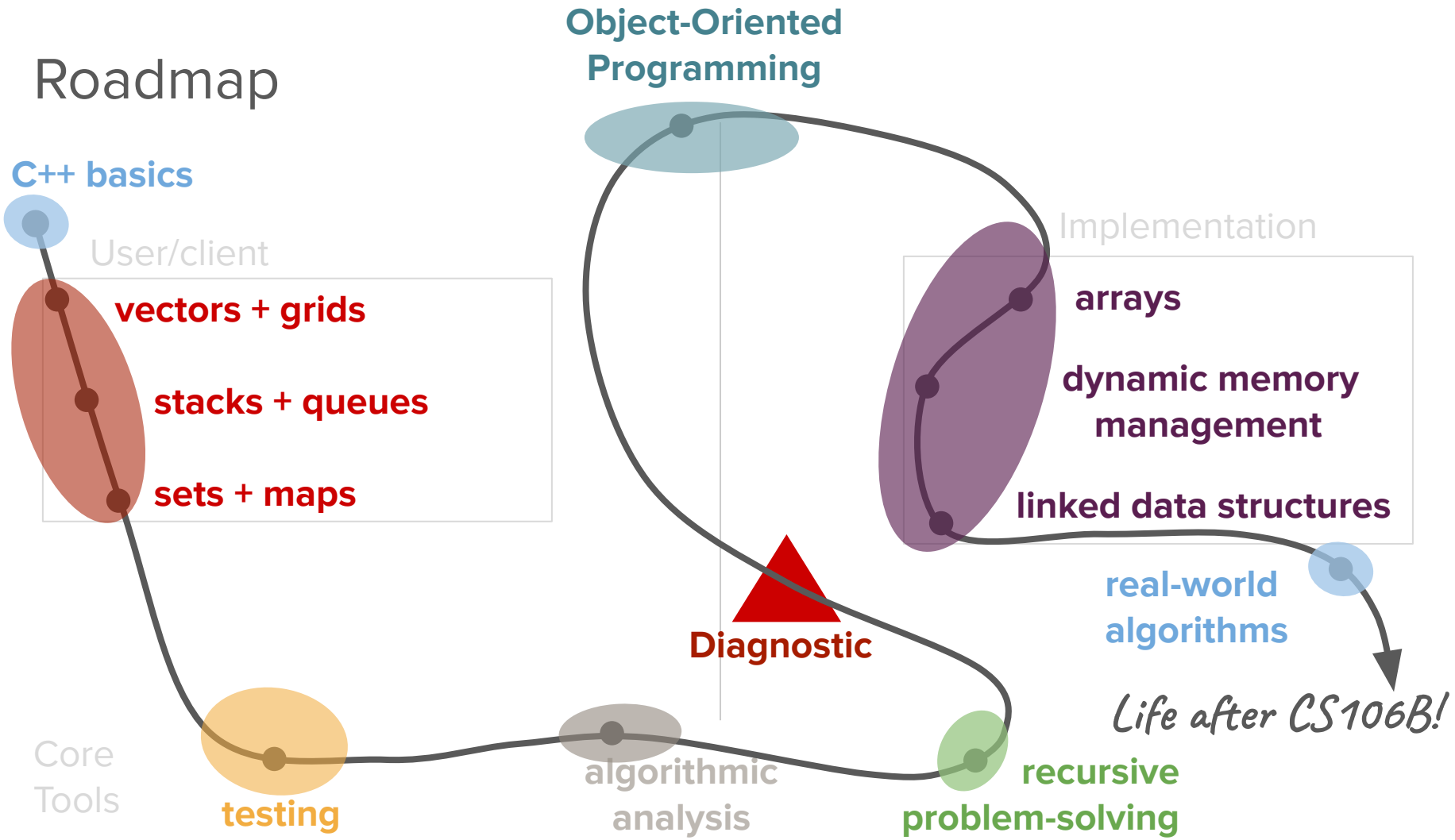
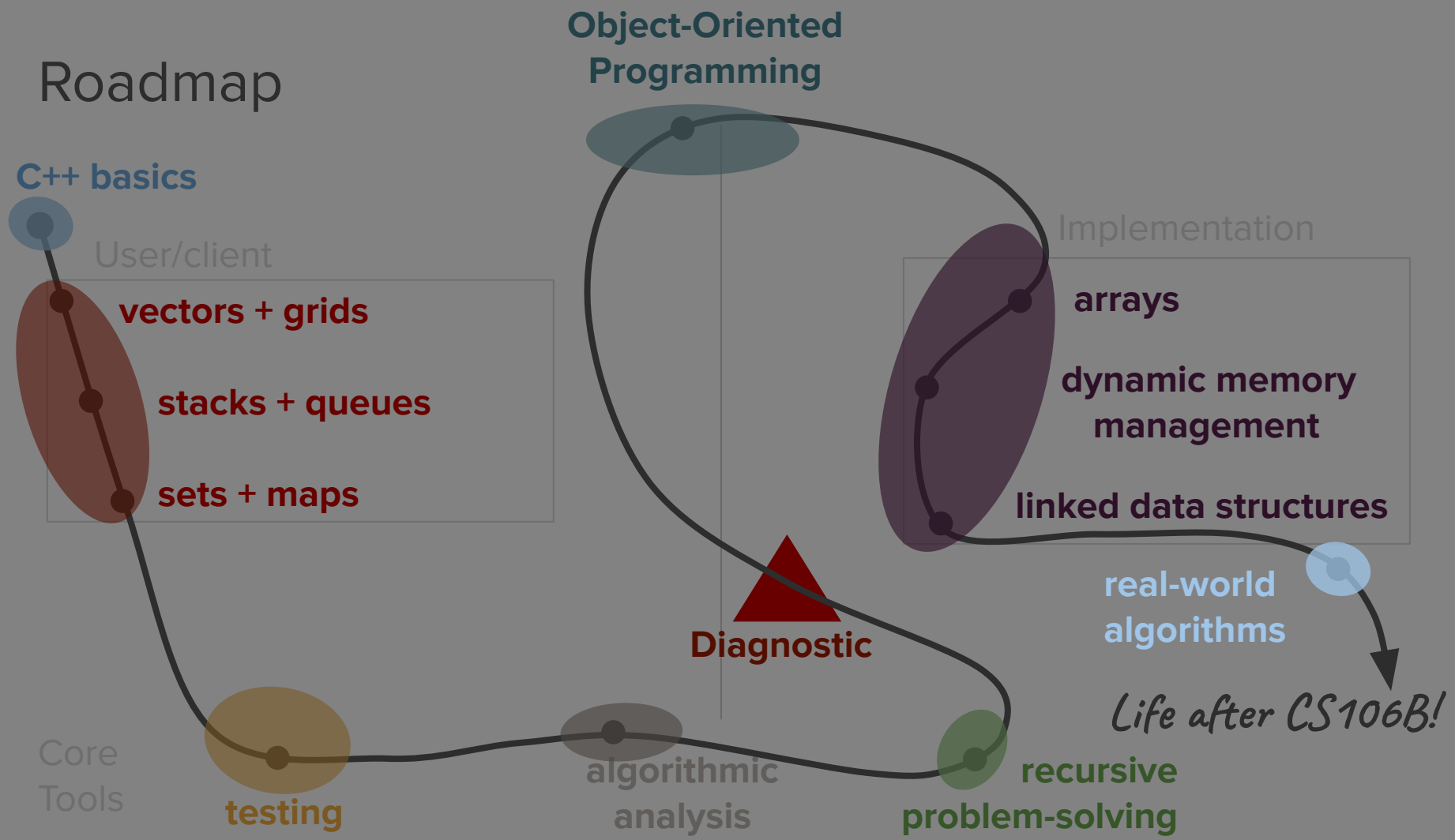
Ethics Roundtable

Roadmap


Object-Oriented Programming
C++ basics
User/client
vectors + grids
stacks + queues
sets + maps
Implementation
arrays
dynamic memory management
linked data structures
real-world algorithms
Life after CS106B!
Diagnostic
Core Tools
testing
algorithmic analysis
recursive problem-solving

Roadmap


Object-Oriented Programming
C++ basics
User/client
vectors + grids
stacks + queues
sets + maps
Implementation
arrays
dynamic memory management
linked data structures
real-world algorithms
Life after CS106B!
Diagnostic
Core Tools
testing
algorithmic analysis
recursive problem-solving

Today's questions

Which problems *should* be solved with technological solutions?

What do responsible programmers do when designing technology that impacts real people?

Today's topics

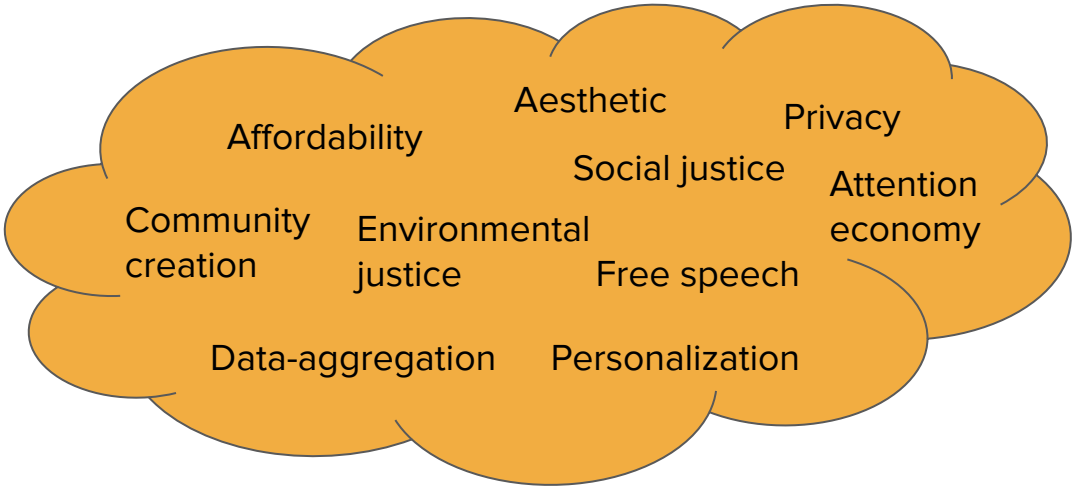
1. Bias in AI/ML
2. Data collection and the attention economy
3. Cryptocurrency and environmental impacts of computing
4. Open Q&A

Ethics in CS106B

Why do we care about ethics in CS106B?

- Technology is not unbiased or neutral – it reflects a set of “design” choices based on the people who created it.

- These design choices encode a set of values....



Ethics in CS106B

- Throughout the quarter, we've covered a number of interesting technical topics and relevant associated ethical discussions.

- Today's class is all about continuing this exploration and allowing you to explore relevant issues that align with your interests!

- Outline
  - Introduction to the three main topics for today
  - Choose a topic/breakout room discussion to participate in (facilitated by Kylie/Nick/Katie)
  - Come together as a class and discuss questions/takeaways from targeted discussions
  - Open Q&A with Nick and Katie

Topics you requested

Bias in Artificial Intelligence/Machine Learning

Facebook challenges legitimacy of some Native names

By BBC Trending
What's popular and why

3 March 2015

Dana Lone Hill had trouble using her real name on Facebook

When Lance Browneyes of the Oglala Lakota community in South Dakota was blocked from Facebook for using a "fake" name, he submitted proof of his identification. Facebook then changed his name to Lance Brown.

Twitter offers $3,500 reward to people who can help fix apparent racial and gender bias in its image-cropping algorithm

Zahra Tayeb Aug 1, 2021, 5:44 AM

Data collection and the attention economy

Much of the digital world runs on a simple formula:

Collect Data -> Improved Recommendation Algorithms -> Increased Attention -> Advertising $$

What are its costs?

- Data collection
- Data sale
- Surveillance
- Tech addiction?

11-06-17 | PLATFORM WARS

Netflix CEO Reed Hastings: Sleep Is Our Competition

For Netflix, the battle for domination goes far beyond which TV remote to pick up.

FIREFOX

The Attention Economy to the Addiction Economy

JULY 27, 2018 MICHAELA SMILEY (THAYER)

Much has been said about the attention economy. Gamification. Tech addiction. happens when those things combine and form an addiction economy online?

Our ever-connected life has provided a lot of value to me. I'm better informed at world and current events, I can use my smart-home to set my thermostat on my and check in on my dog, and I can talk daily with friends far away on any numbe messaging or video-chat services.

The internet and incentives

Cryptocurrency and environmental impacts

Cryptocurrency environmental impact

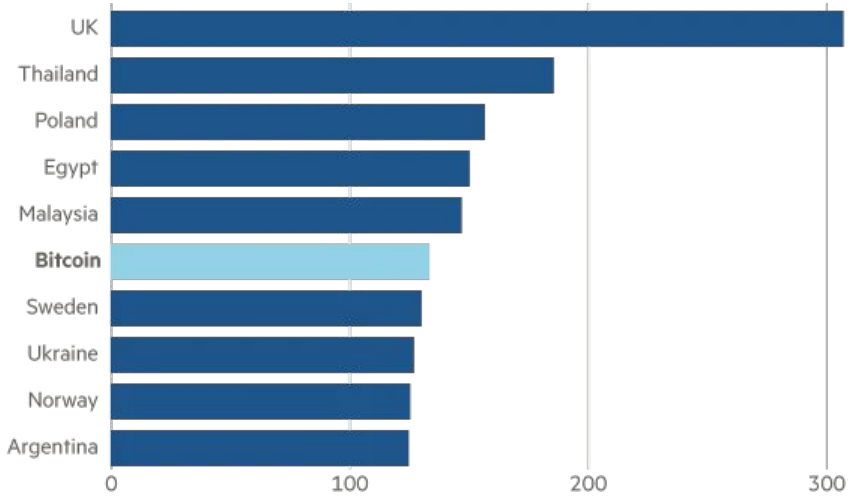
Could we discuss the environmental impact of computing? Specifically, the energy usage by clusters and servers, and the heat produced during their operation.

Current events on ethical computing e.g. bitcoin

Cryptocurrency and environmental impacts

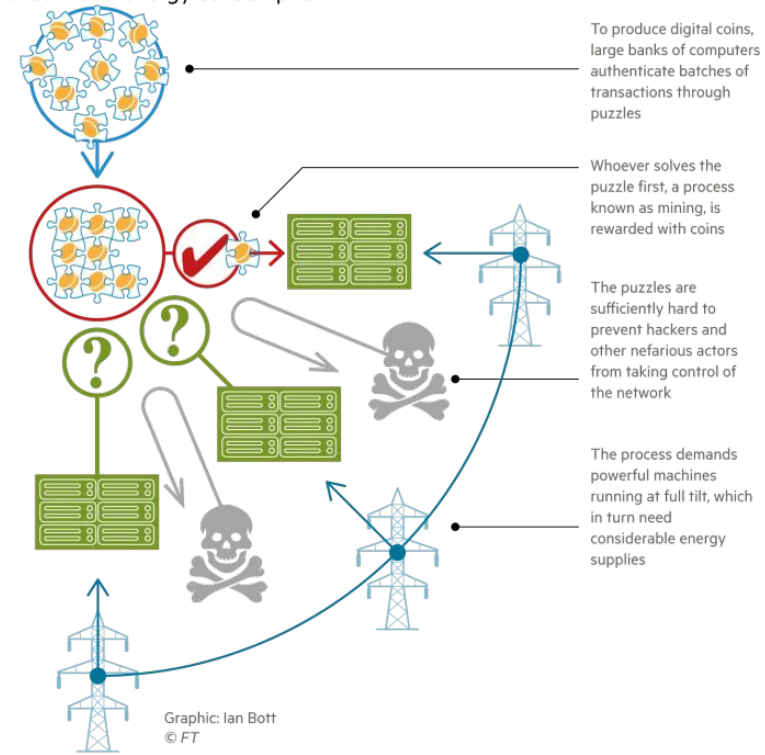
Bitcoin consumes around half as much electricity as the UK

Annualised consumption of bitcoin, compared with countries' readings (TWh)



Using assumption that electricity cost paid by miners globally is $0.05 per kWh. Bitcoin data from May 19 2021.
Country data from 2019 or most recent year
Source: Cambridge Bitcoin Electricity Consumption index
© FT

Bitcoin and energy consumption



Graphic: Ian Bott
© FT

Time for breakouts!

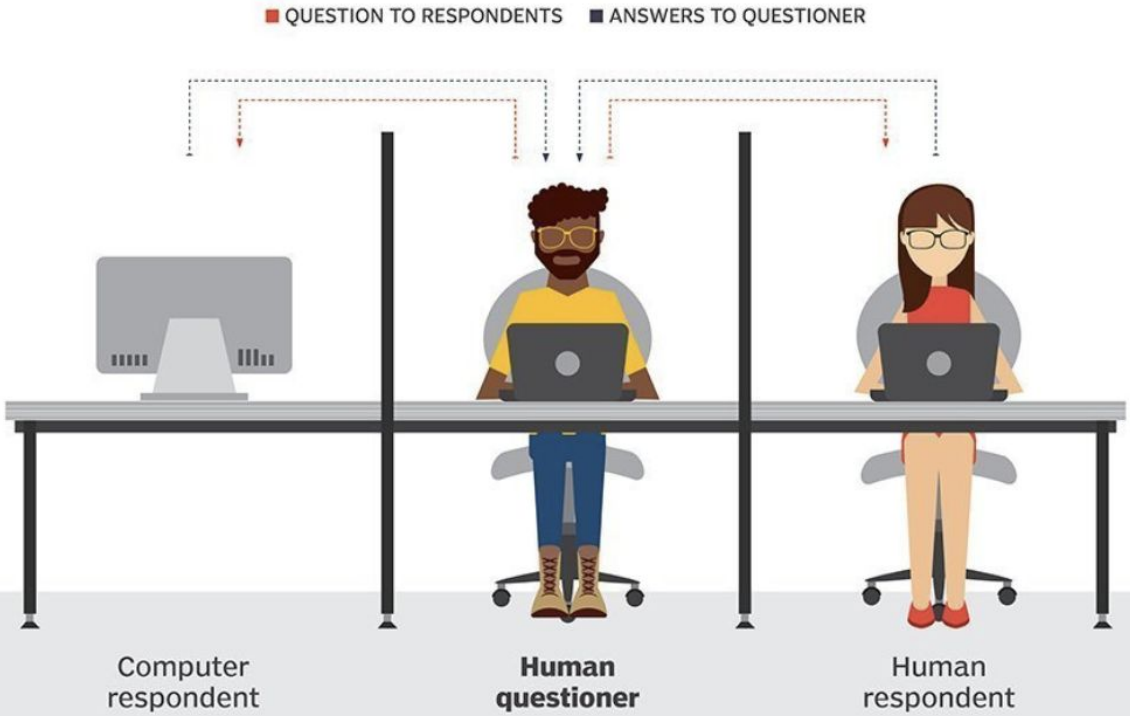
Ethics Redux Remix Reloaded

Bias in Artificial Intelligence / Machine Learning

What traits of a human being does a robot need to have “human intelligence”?

Turing test

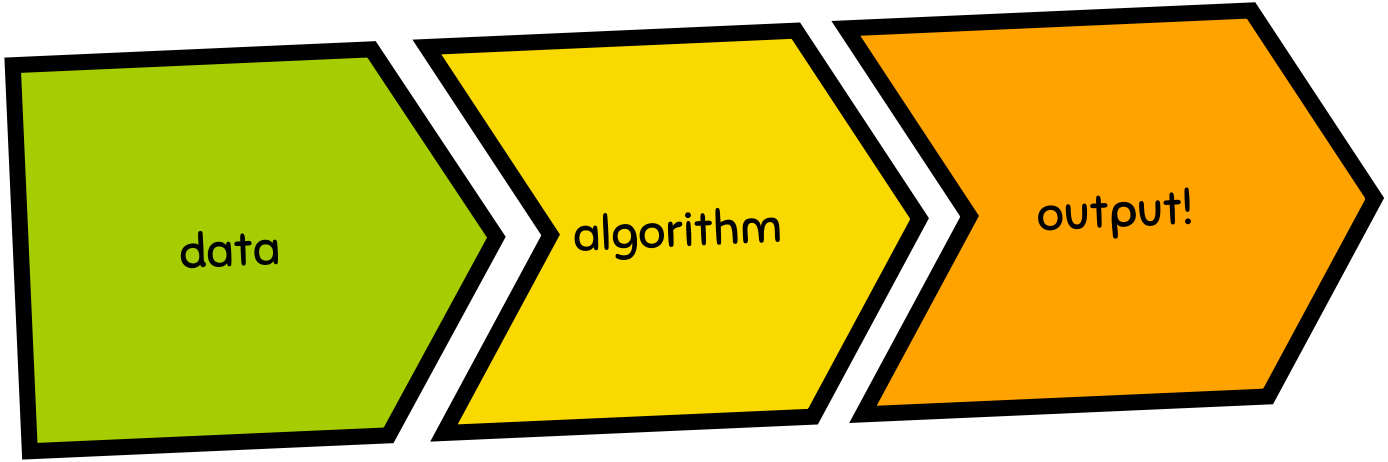
During the Turing test, the human questioner asks a series of questions to both respondents. After the specified time, the questioner tries to decide which terminal is operated by the human respondent and which terminal is operated by the computer.



How we define AI today

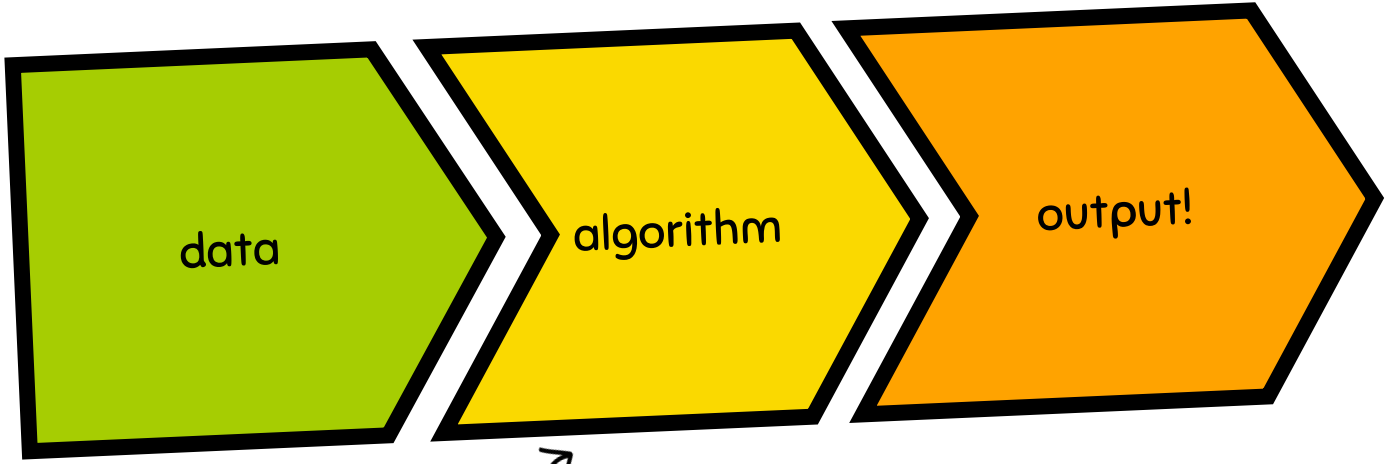
The goal of artificial intelligence (AI) is to “mimic the perception, learning, problem-solving, and decision-making capabilities of the human mind.” (IBM)

AI has three parts



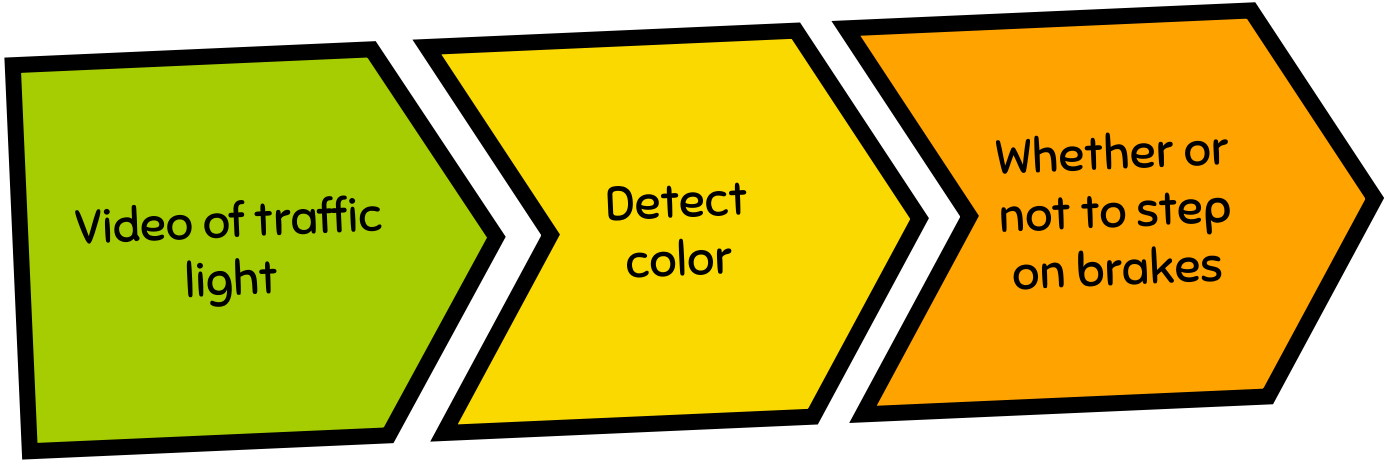
(Thanks to MIT Media Lab for this framework!)

AI has three parts

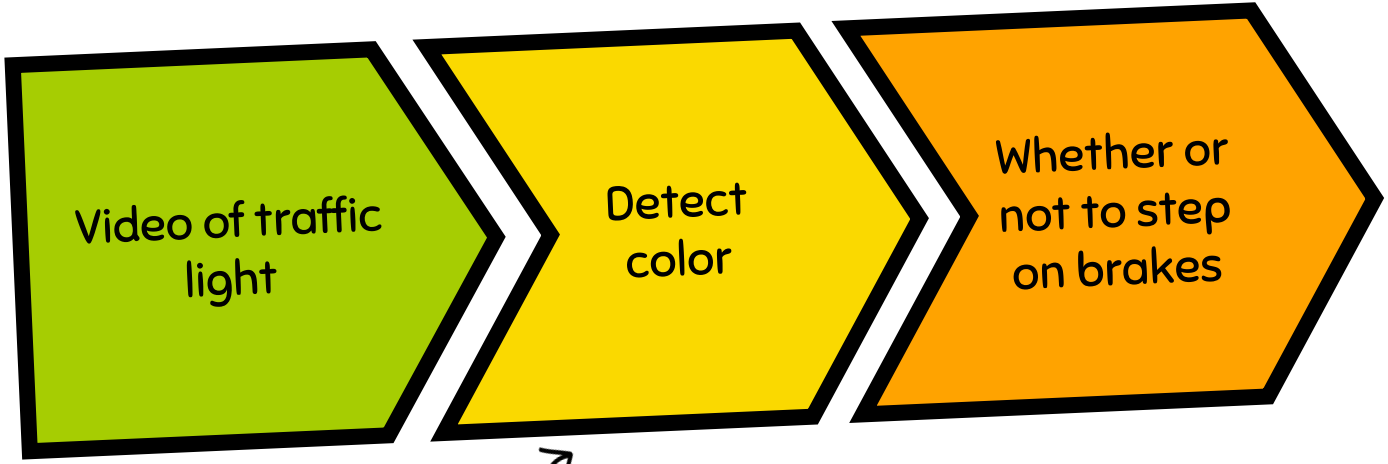


Often a set of code-based "rules" to determine the output

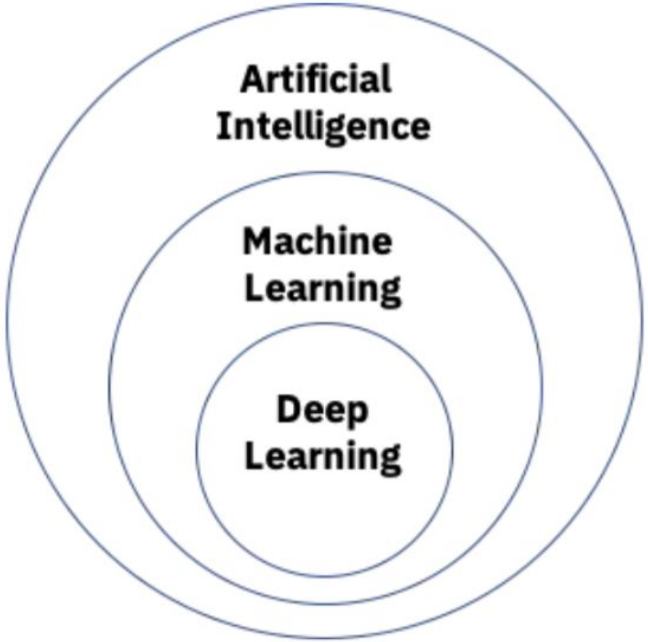
An example


Video of traffic light
Detect color
Whether or not to step on brakes

An example

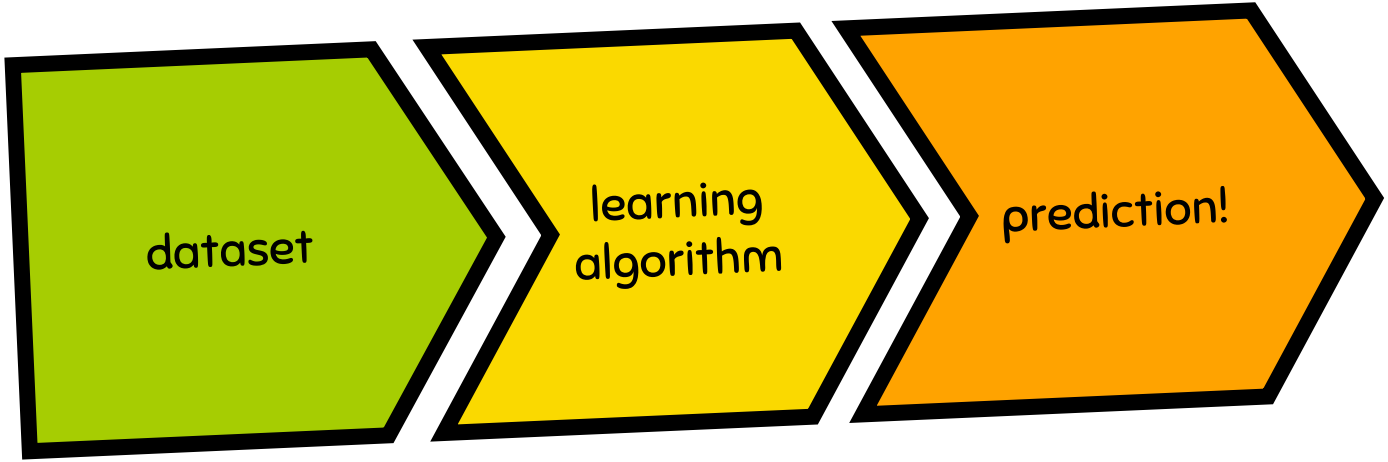


If green, press gas. If red, trigger brakes. (etc.)

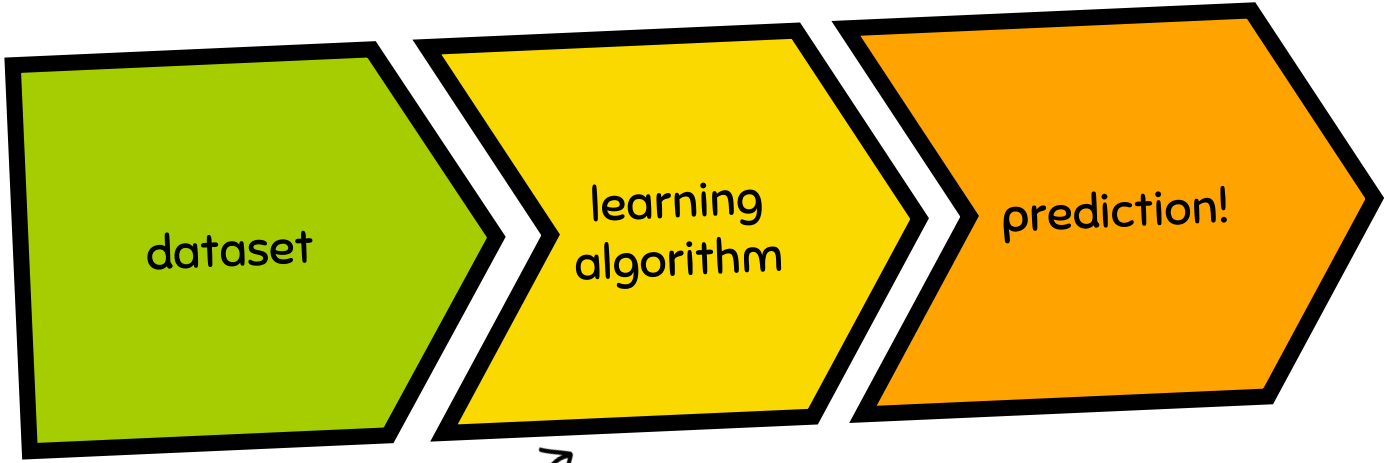


Machine learning (ML) is a subset of artificial intelligence (AI).

Machine learning has three parts

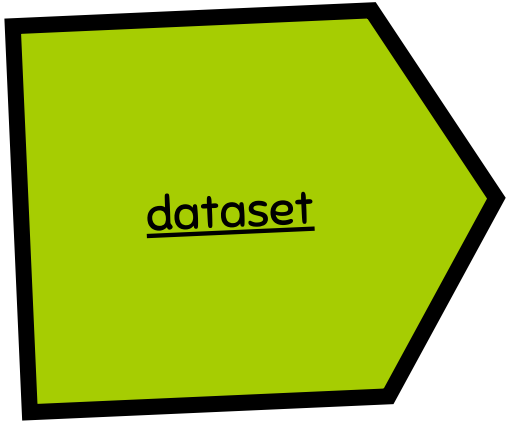


Machine learning has three parts



In the case of machine learning, the **algorithm itself is learned and improves over time.**

ML has three parts

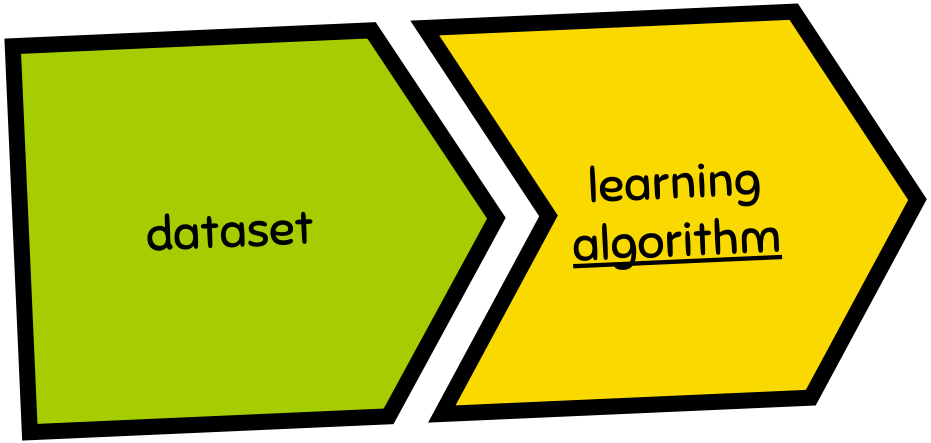


Datasets

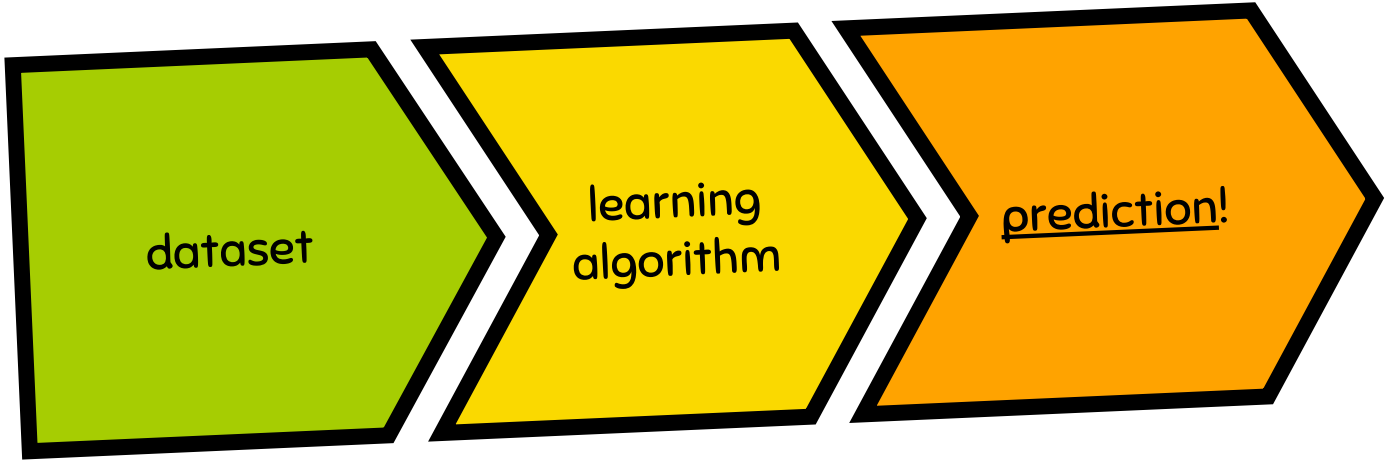
A <u>dataset</u> is a collection of curated data

- Images
- Measurements (time, views, inches, etc)
- Text
- Video recordings!

ML has three parts

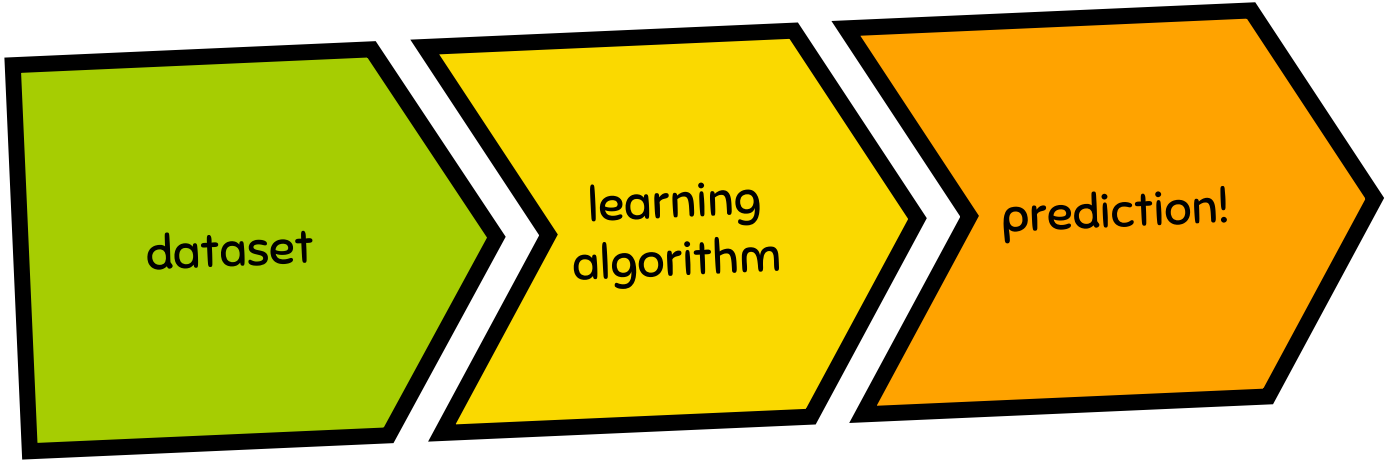


ML has three parts



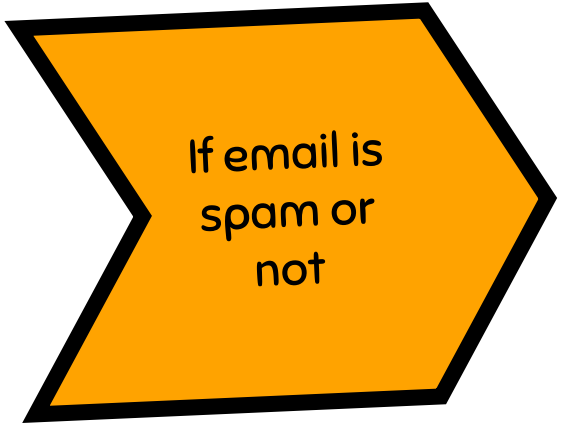
An example: Spam filtering

Email spam filter

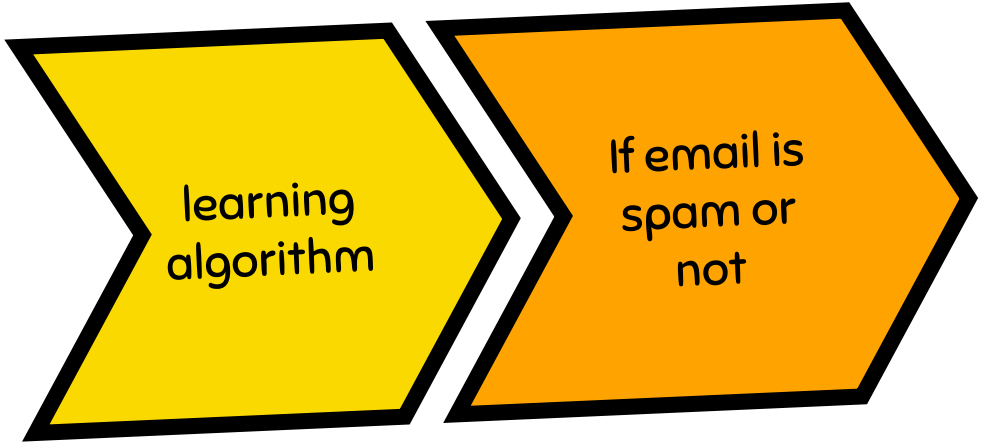


Email spam filter

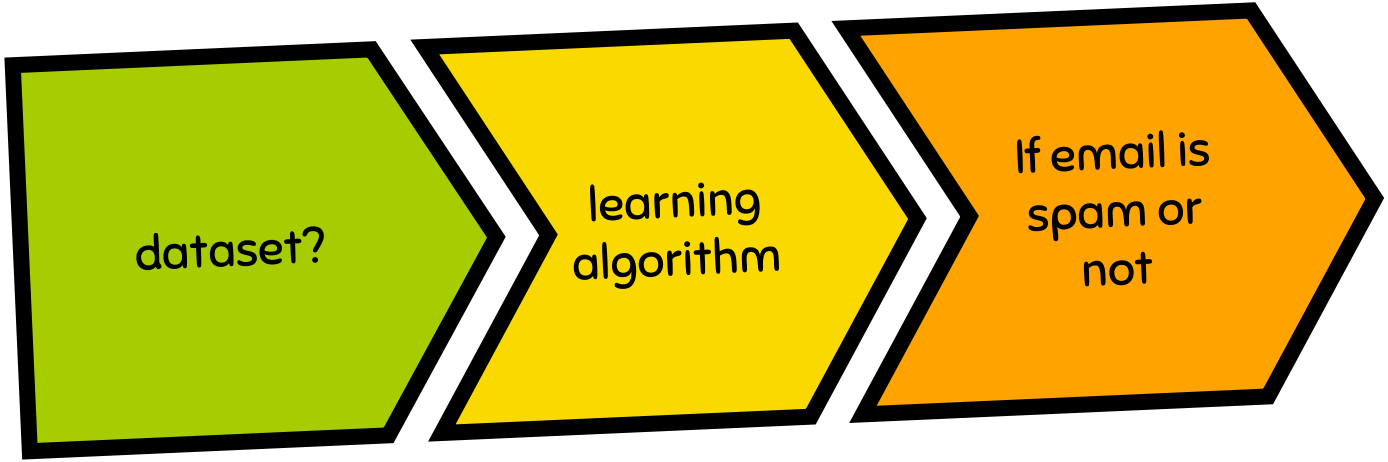
Email spam filter



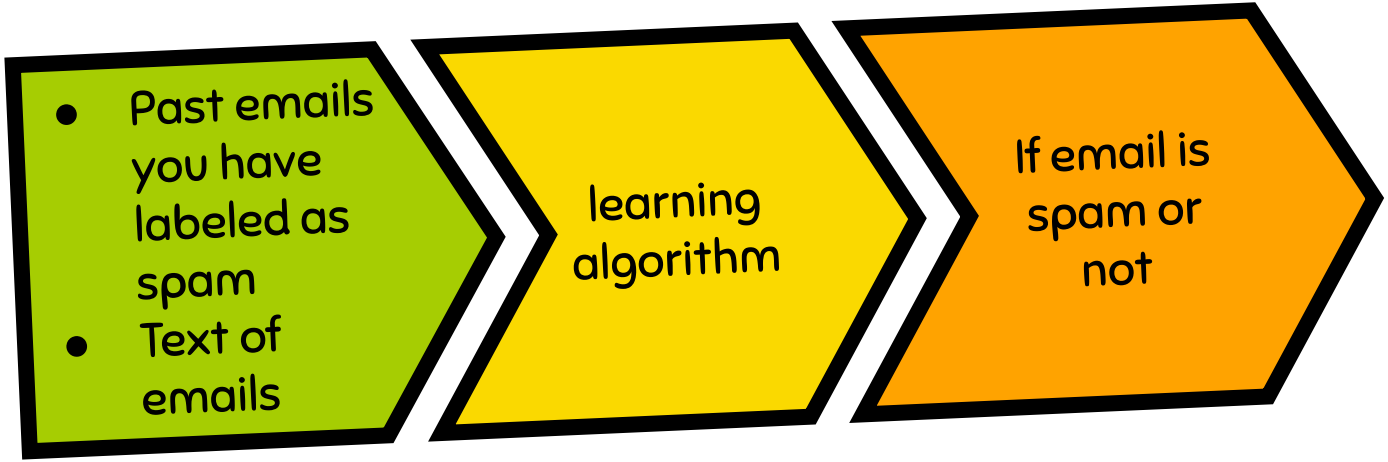
Email spam filter



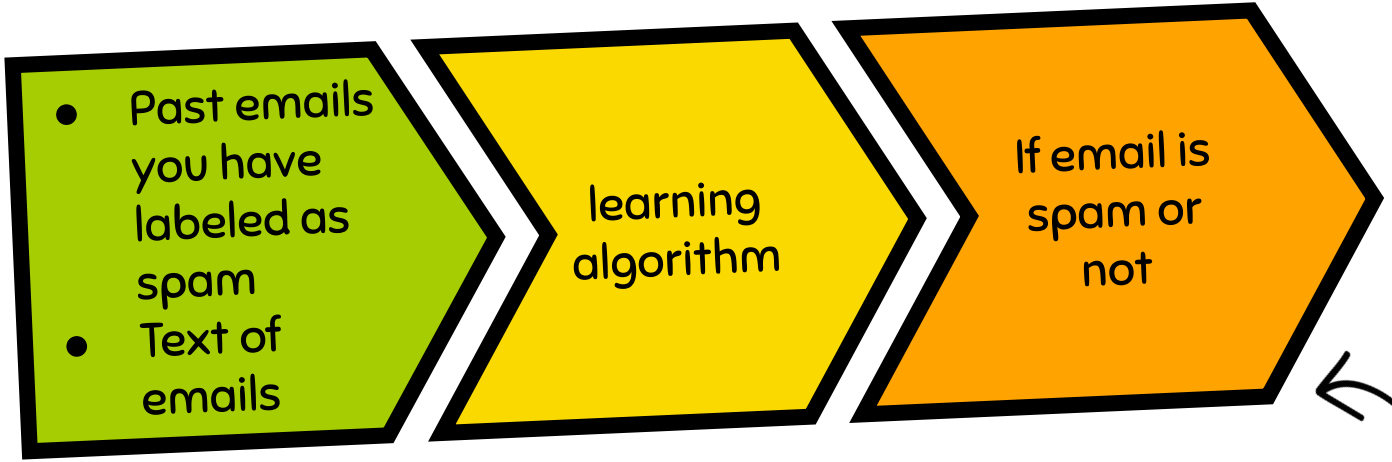
Email spam filter



Email spam filter



Email spam filter



Note that the prediction happens for emails (i.e. data) that we've never encountered before.

Machine learning for recognizing “names”

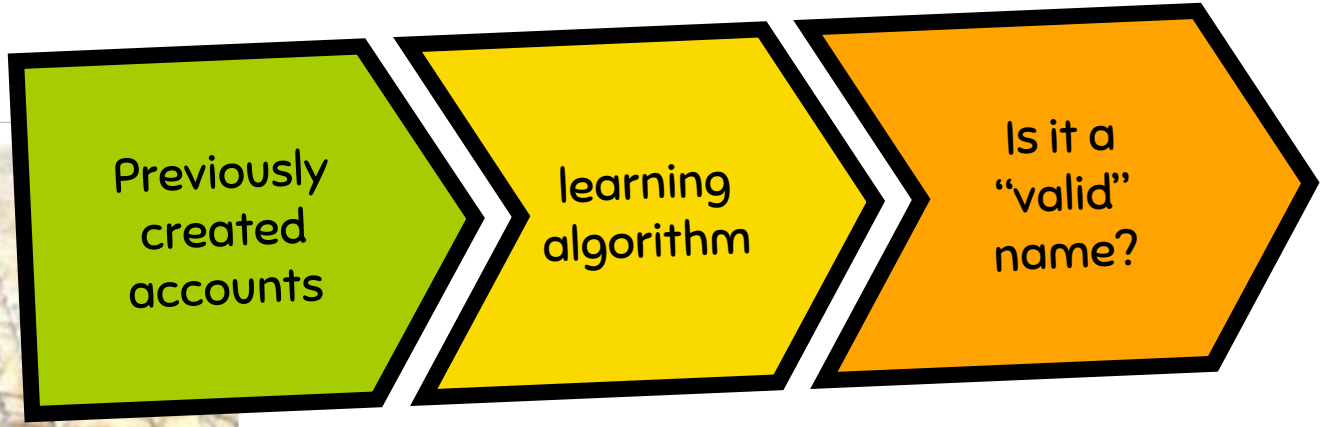
Facebook challenges legitimacy of some Native names

By BBC Trending
What's popular and why

3 March 2015

Dana Lone Hill had trouble using her real name on Facebook

When Lance Browneyes of the Oglala Lakota community in South Dakota was blocked from Facebook for using a "fake" name, he submitted proof of his identification. Facebook then changed his name to Lance Brown.



https://www.bbc.com/news/blogs-trending-31699618

What are some “rules” that may have been in place in the name recognition code?

- No nouns allowed as names?
- Checking for certain prohibited words (ex: Robin Kills The Enemy)
- Checking for combinations of words (e.g. 2 nouns joined together)
- Checking for numbers/symbols

Machine learning for recognizing “names”

Twitter offers $3,500 reward to people who can help fix apparent racial and gender bias in its image-cropping algorithm

Zahra Tayeb Aug 1, 2021, 5:44 AM

Previously cropped images/ images in posts → learning algorithm → How to crop an image

https://blog.twitter.com/engineering/en_us/topics/insights/2021/sharing-learnings-about-our-image-cropping-algorithm
https://www.businessinsider.com/twitter-offers-cash-bounty-racial-gender-bias-algorithm-2021-8

What would you do to try to “fix” the cropping algorithm?

- Add more diverse pictures to the dataset so the learning algorithm can adapt
- Study the algorithm and maybe weight color less than facial features?
-

Takeaways + terminology

- **Overfitting** is when you end up creating a ML classifier that only works for the known test data.
  - We want the classifier have high accuracy even when tested on completely new data.

- **Bias** can be embedded into our input data and can cause overfitting.
  - E.g. having many more example cats than dogs when classifying cats vs. dogs, not including diverse photos of a particular prediction class, etc.

- Try out Google's Teachable Machine to try creating a machine learning classifier given your own dataset and prediction categories – are you able to create an unbiased dataset?

Questions to discuss with the larger group

Data collection and the attention economy

Machine Learning, Data, & the “Attention Economy”

Collect Data	Train Algorithms	Increased Attention	Advertising Money
Companies collect data on the activities of users on their platform. Some also buy third party data that can be associated with individual users.	Data is used to train algorithms that learn in more fine-grained detail what users like and what content increases their time on the platform.	Users spend more time on the platform or app and more time attending to the content (and also to the advertisements)	Advertisers have more opportunity to show ads or otherwise monetize the attention.

Articles to Skim

1. Our Competitor is Sleep:

https://www.fastcompany.com/40491939/netflix-ceo-reed-hastings-sleep-is-our-competition

2. Tik Tok's Algorithm:

https://www.forbes.com/sites/michellegreenwald/2021/04/01/audience-algorithm-and-virality-why-tiktok-will-continue-to-shape-culture-in-2021/

How TikTok Holds our Attention (longer article if you have time -- skim to the middle):

https://www.newyorker.com/magazine/2019/09/30/how-tiktok-holds-our-attention

Questions

What are the positive and negative effects of staying longer on a website, app, or platform?

- To the platform?
- To the individual person/user?
- To society?

Some software mandates a maximum length of time that one can spend per day on a game or app. Is this a good idea? Should it become mandatory?

The attention economy is not the only economy. What apps, platforms, and services exist outside of the attention economy & don’t rely on advertising money?

Privacy and Trust

Privacy is sometimes described as being a tradeoff with other important values like security or efficacy.

It’s not so simple.

Respect for privacy can also build justified trust in an institution or product, which can lead to citizens or users cooperation.

Possible harms of surveillance and data sharing:

Aggregation — joining together small pieces of information that together reveal information a user might not want to share. Especially common in “secondary use” of data/data resale. Solove’s example: cancer book + wig = undergoing chemotherapy.

Inhibition — surveillance can inhibit activity or speech, even without further threats, just because the user or citizen knows the activity will be observed or monitored.

Exclusion — the citizen or user is excluded from seeing their own data and therefore does not know how it is being used and is not able to correct errors in the data.

Distortion — parties with access to data (companies or governments) may draw incorrect conclusions from the (necessarily incomplete) data they have.

Broken Trust — Users or citizens may place less trust in companies or governments that collect and share their data. This can make behavior modification less effective.

1: Opt-in/out Location Data for iMoveItMoveIt

A behavior modification app, iMoveItMoveIt, wants help users to move more by sending “nudge” notifications throughout the day. How should the company handle its location data? Discuss the options below with reference to the possible harms of surveillance and data sharing and to possible benefits of having the data.

- OPT-IN: Turn off all location tracking by default. Users will be sent standardized, research-based nudges at pre-determined times throughout the day unless they opt-in to location tracking, which turns on more personalized recommendations.

- OPT-OUT: The standard app includes location tracking, but users have an option to turn off all location tracking. The two versions of the app are the same as in option 1, but the default is different.

For both options, how should users be notified that the choice exists?

2: Where is iMoveItMoveIt Location Data Stored?

A behavior modification app, iMoveItMoveIt, wants help users to move more by sending “nudge” notifications throughout the day. How should the company handle its location data? Discuss the options below with reference to the possible harms of surveillance and data sharing and to possible benefits of having the data.

- KEEP DATA LOCAL: Access location data within the app to generate reports, but do not send any location data to iMoveItMoveIt. Instead, share reports about what nudges increased or decreased movement.

- SEND DATA: Send all location data to iMoveItMoveIt in hopes that fine-grained data analytics and patterns in location data can be used to improve nudges.

For both options, if data is stored, how should the company control who has access to the data and what happens to the data if the company is sold?

3: Is iMoveItMoveIt Data Sold?

A behavior modification app, iMoveItMoveIt, wants help users to move more by sending “nudge” notifications throughout the day. How should the company handle its location data? Discuss the options below with reference to the possible harms of surveillance and data sharing.

- PROTECT DATA: Access location data within the app to generate reports, but do not send any location data to iMoveItMoveIt. Instead, share reports about what nudges increased or decreased movement.

- SELL DATA: Sell all location data generated by iMoveItMoveIt users in order to monetize the app.

For both options, if data is stored, how should the company control who has access to the data and what happens to the data if the company is sold?

4: Data Disclosure

A behavior modification app, iMoveItMoveIt, wants help users to move more by sending “nudge” notifications throughout the day. Assume the company is currently collecting location data. What should be disclosed to users? Discuss the options below with reference to the possible harms of surveillance and data sharing.

- SILENT COLLECTION: Location data of iMoveItMoveIt users is silently collected - the only place to find out what is collected is in the long legal “consent” form to which the users initially agreed.

- ACTIVE DISCLOSURE: iMoveItMoveIt informs users about what kind of data is collected periodically with a click-through graphic.

How would your preferred option affect user trust in the app? What are the other stakes of this decision?

Student Contributions/Ideas

Responses to Questions

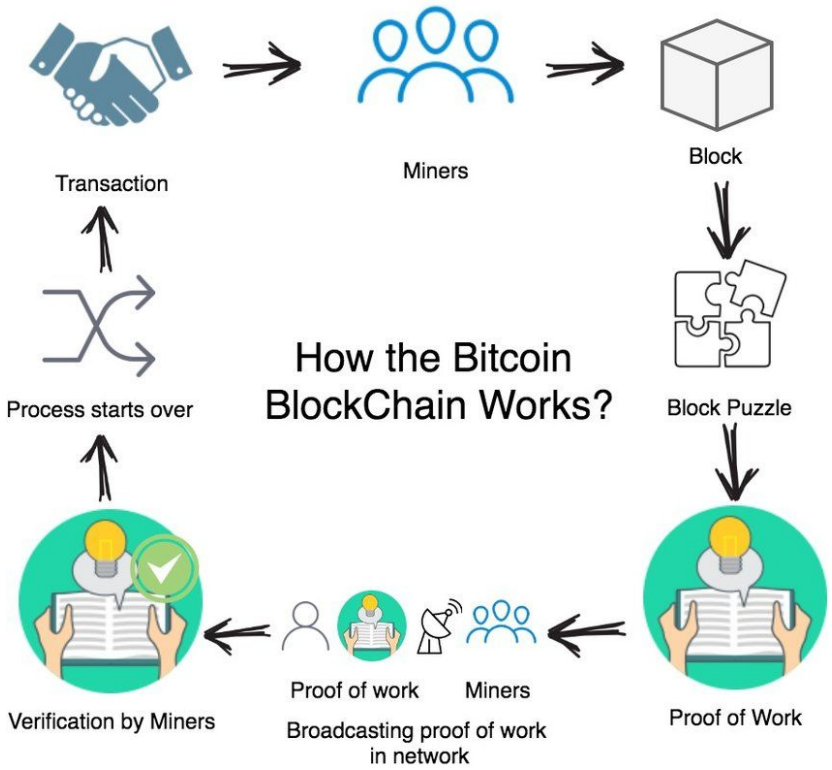
- Content influences opinions -- what about political opinions? How would we know what influence was occurring?
- Mandates for length? Or a series of warnings for users that express the risk?
- What if time limits don't actually change behavior but just move between games? Some games request that users spend a minimum amount of time.

New Questions to Ask

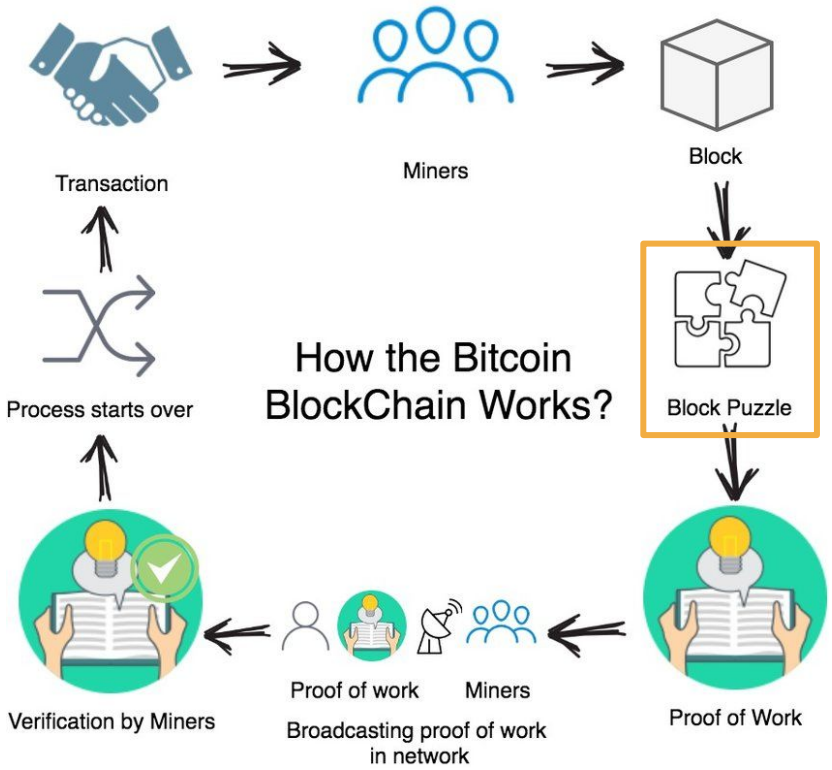
- How do we improve the situation? What levers do we have for change?
- To what extent should companies be held accountable when their products promote harmful behaviors/habits? Should the consumer be held accountable instead?

Cryptocurrency and environmental impacts of computing

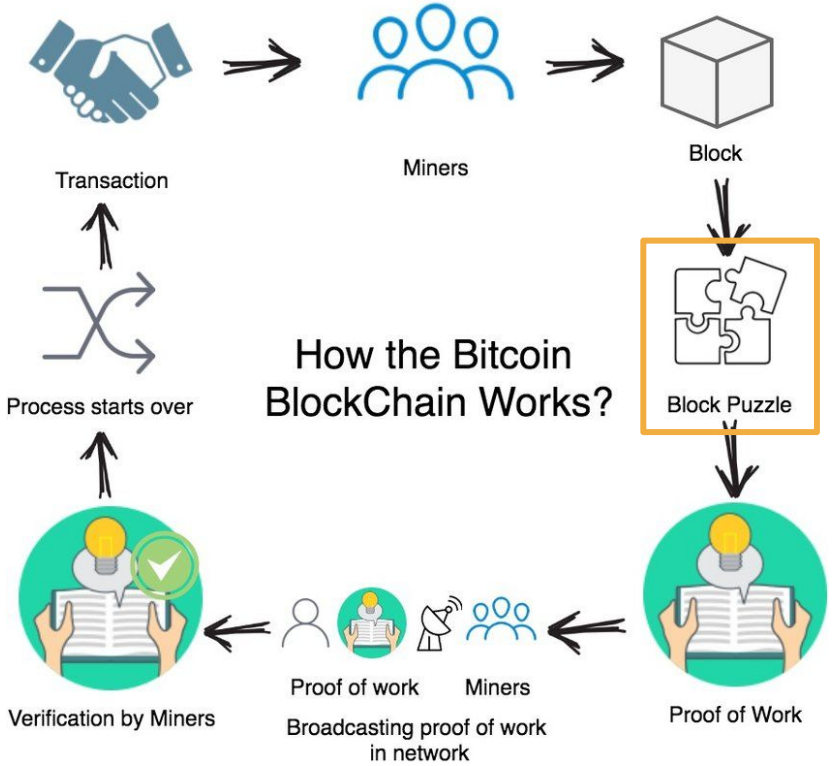
Cryptocurrency Background

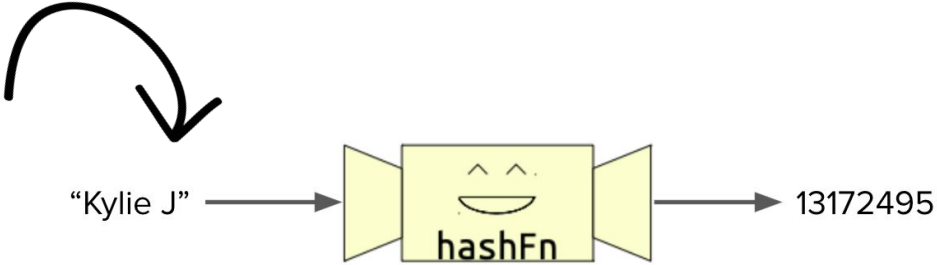


Cryptocurrency Background

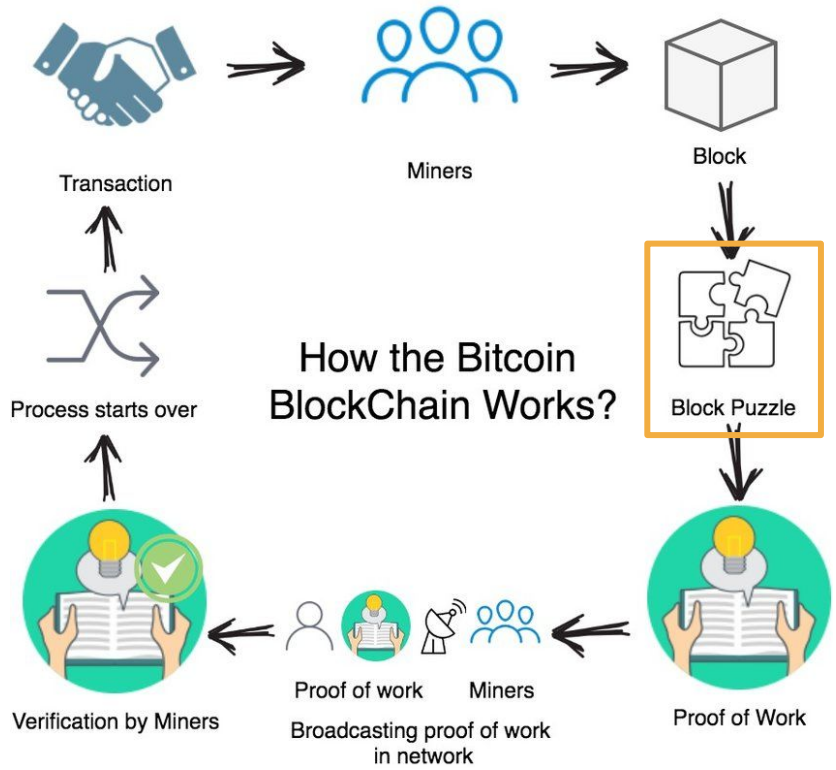


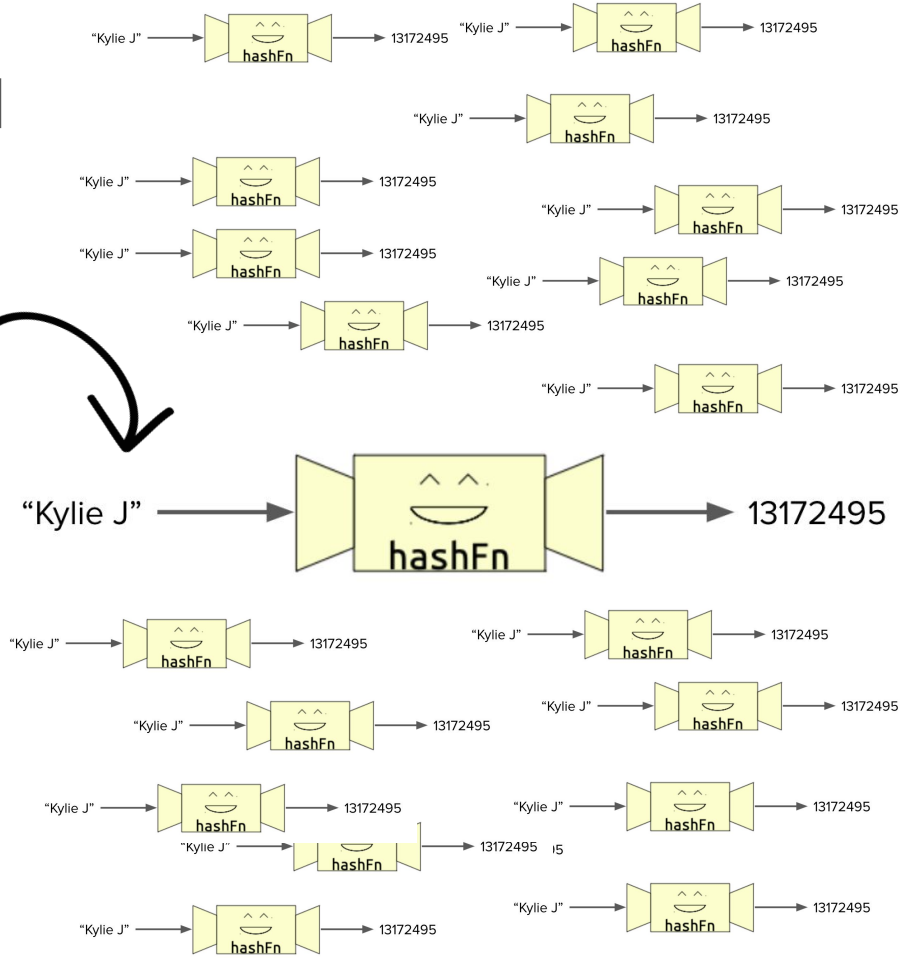
Cryptocurrency Background



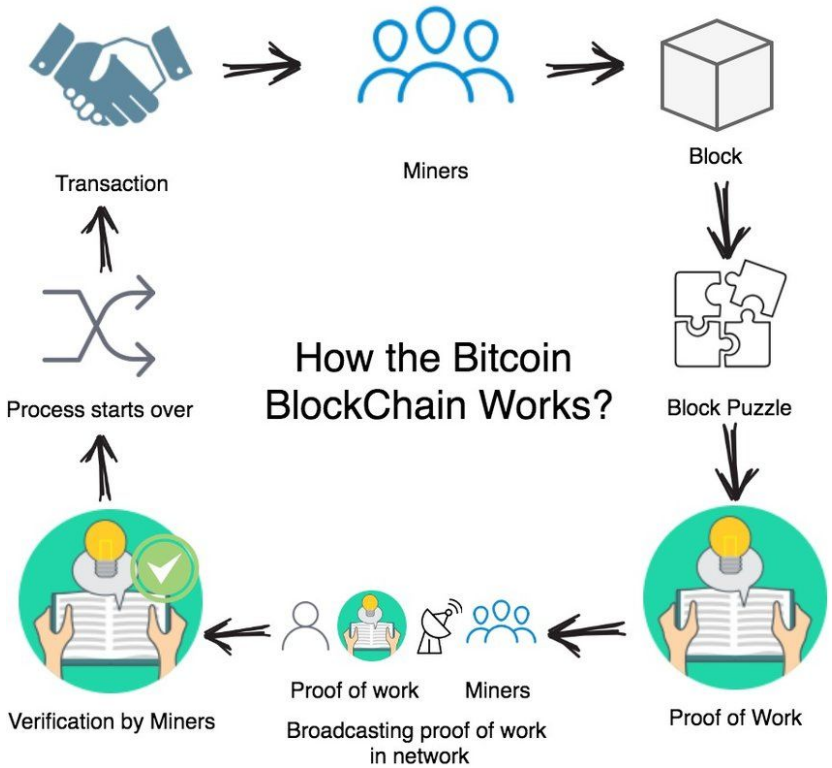


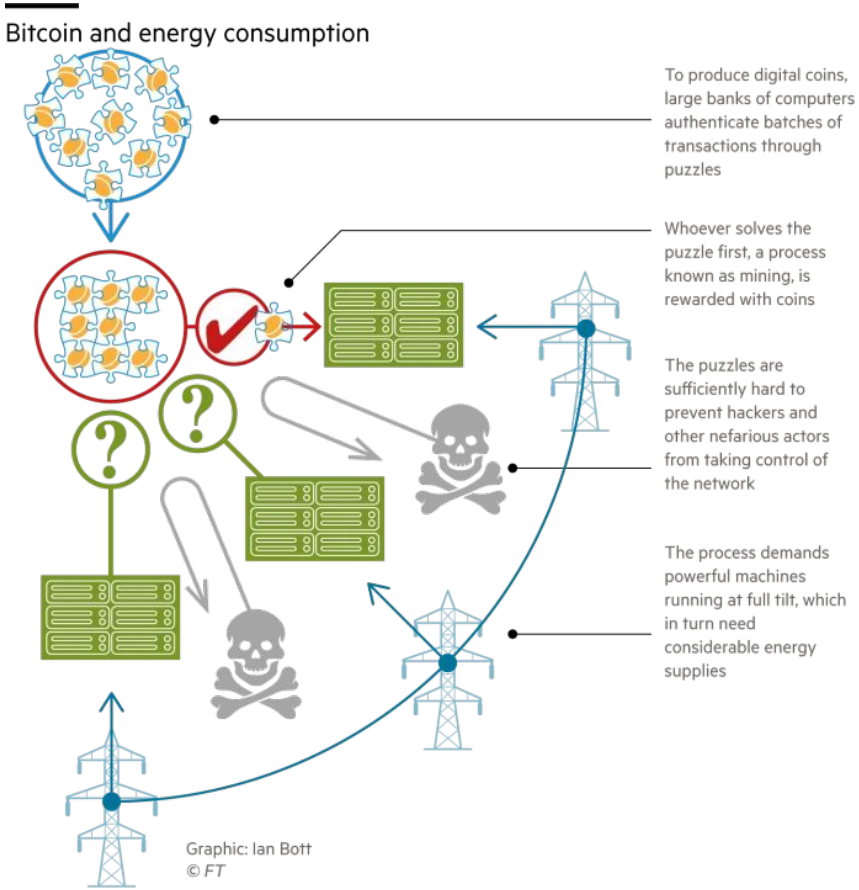
Cryptocurrency Background



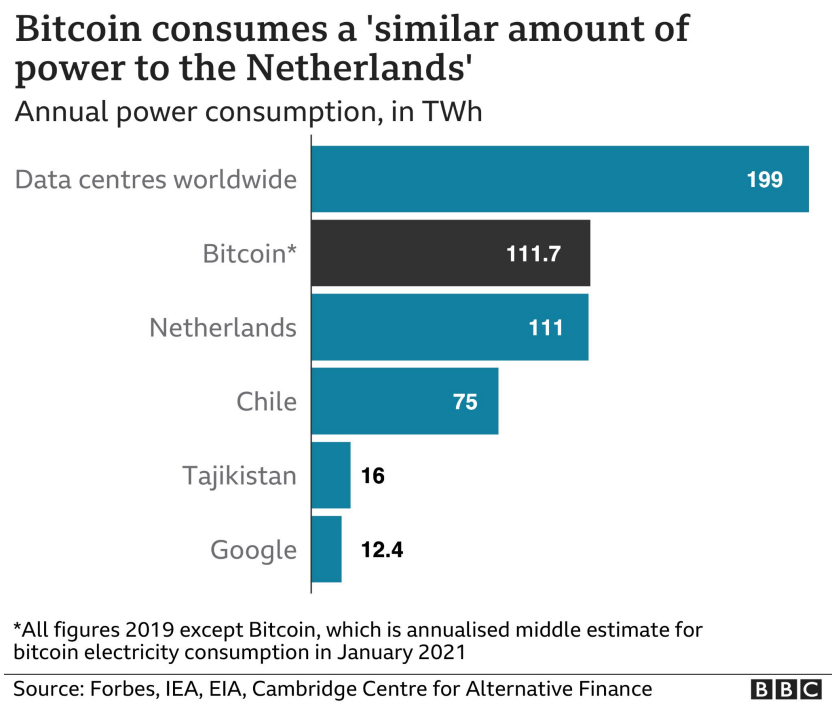


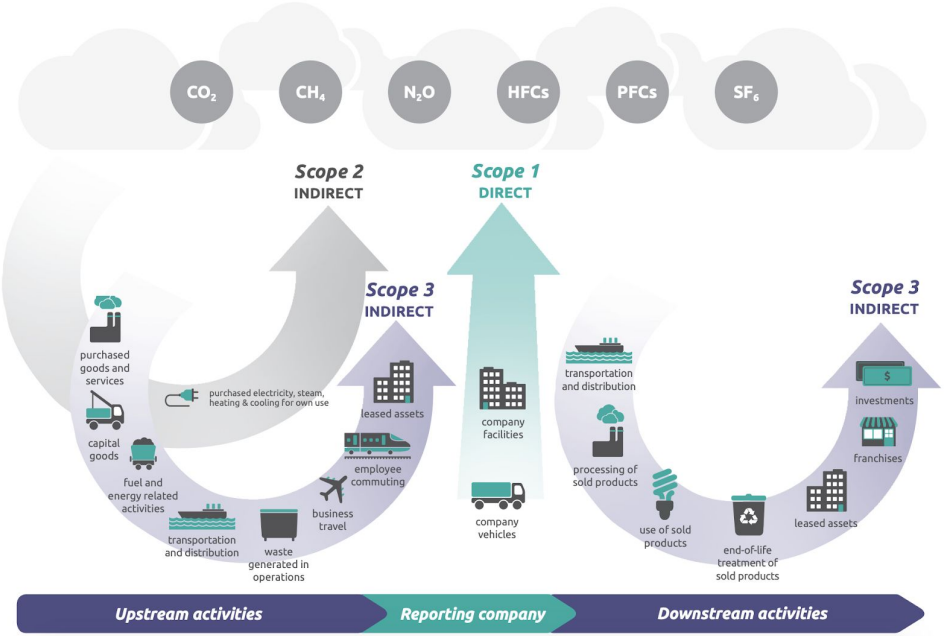
Cryptocurrency Background





Environmental Impacts of Computing





Discussion Activity

- Take 5 minutes to read this article: https://qz.com/1204842/bitcoin-mining-should-use-renewable-energy-if-we-want-cryptocurrencies-to-be-ethical/

- When we come back together, we'll discuss the following questions:
  - What benefits, if any, do cryptocurrencies provide in relation to traditional monetary systems?
  - What environmental sacrifices should we be willing to make in order to obtain these benefits? Should we make environment sacrifices at all?

- What questions/insights would you add to this discourse?

Student Contributions/Ideas

Responses to Questions

- At some point, systems might "self-regulate" when the cost of mining exceeds the value gained by people using the system
- Benefits of cryptocurrencies: better than unstable national currencies in some countries, more accessible/easier to transact
- Put cryptocurrency gains into clean energy investment

New Questions to Ask

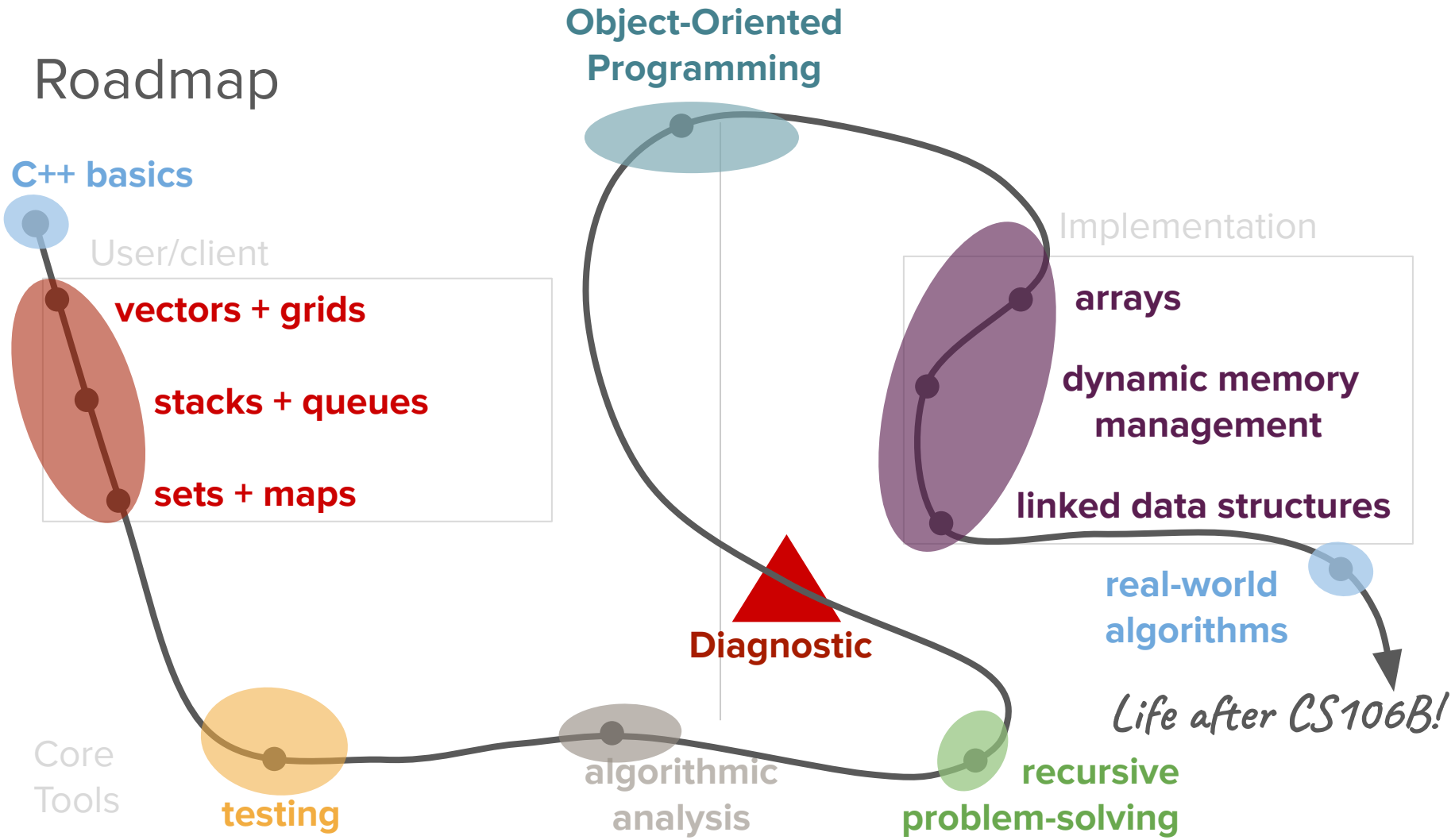
- What if there was an incentive system to encourage miners to work together to solve the block puzzle and avoid wasted computation?

Open Q&A/Discussion

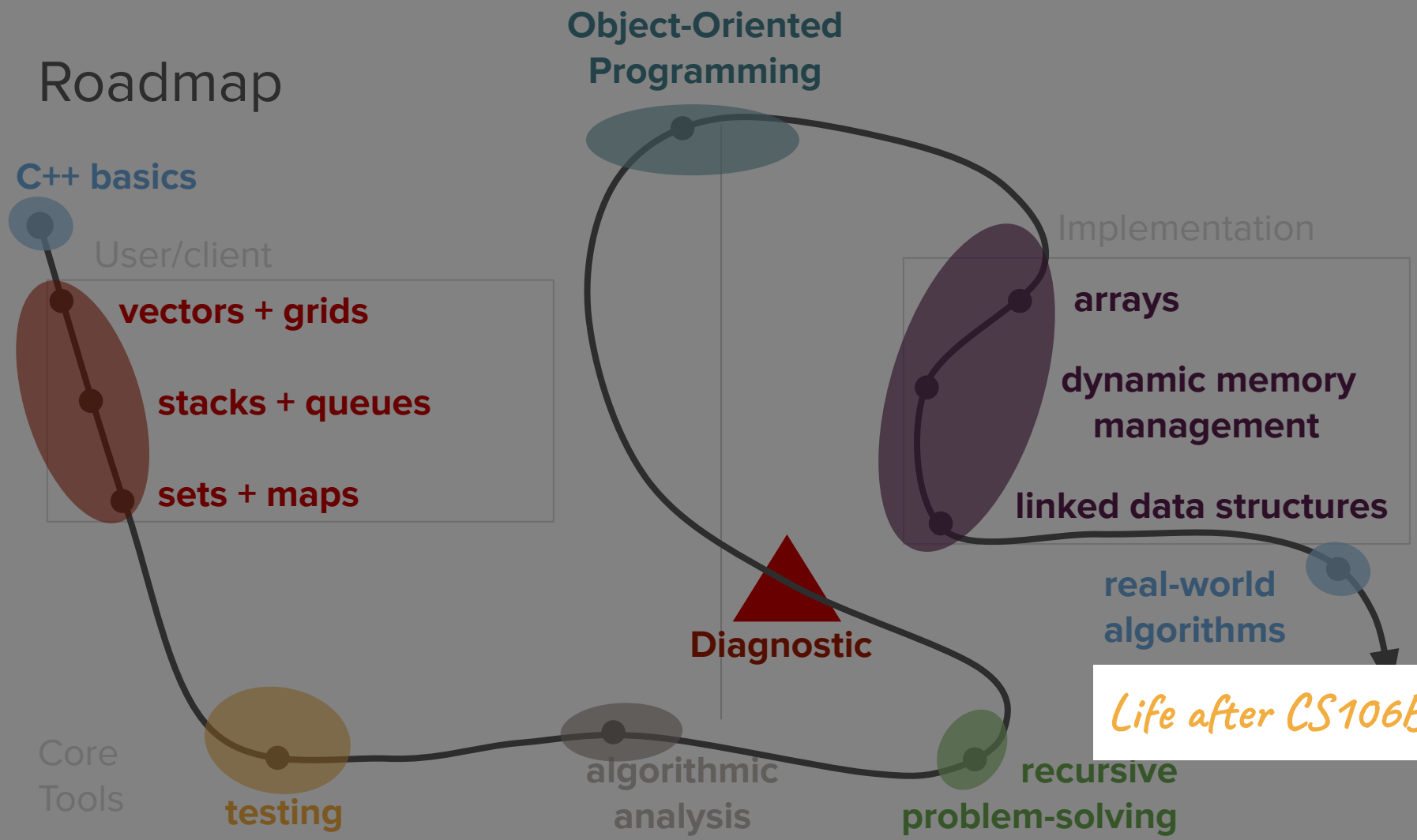
What questions do you have relating to ethics in computing?

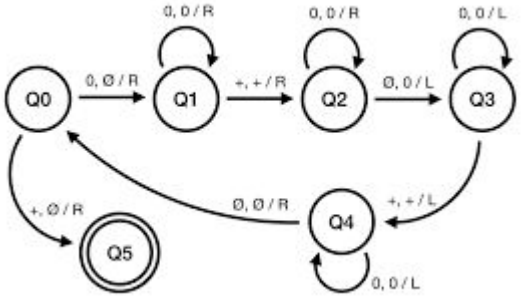
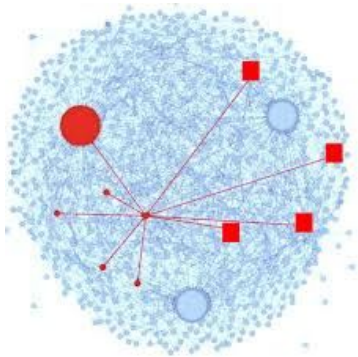
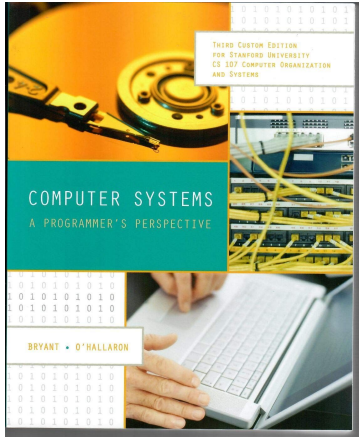
What’s next?

Roadmap


Object-Oriented Programming
C++ basics
User/client
vectors + grids
stacks + queues
sets + maps
Implementation
arrays
dynamic memory management
linked data structures
real-world algorithms
Life after CS106B!
Diagnostic
Core Tools
testing
algorithmic analysis
recursive problem-solving

Roadmap


Object-Oriented Programming
C++ basics
User/client
vectors + grids
stacks + queues
sets + maps
Implementation
arrays
dynamic memory management
linked data structures
real-world algorithms
Diagnostic
Life after CS106B!
Core Tools
testing
algorithmic analysis
recursive problem-solving

Life after CS106B

During Friday's lecture, we will be having an "Ask Us Anything" component, where we will open up the floor to you all to ask us any questions you have about computer science, Stanford, or life in general. **We'll be collecting questions in advance as well – if you have any burning inquiries on your mind, go ahead and fill out this Google form!**