The Politics of Algorithmic Justice

Kathy L. Powers

University of New Mexico

Department of Africana Studies

Department of Political Science

Kpowers1@unm.edu

September 27, 2023

Acknowledgements

- New Mexico State Legislature Committee on Courts, Corrections; and Justice
- UNM Algorithmic Justice Project
- Santa Fe Institute's Interdisciplinary Working Group on Algorithmic Justice
- UNM Department of Africana Studies
- UNM Department of Political Science

What is an Algorithm?

- A set of instructions (inputs) that leads to an outcome (output).
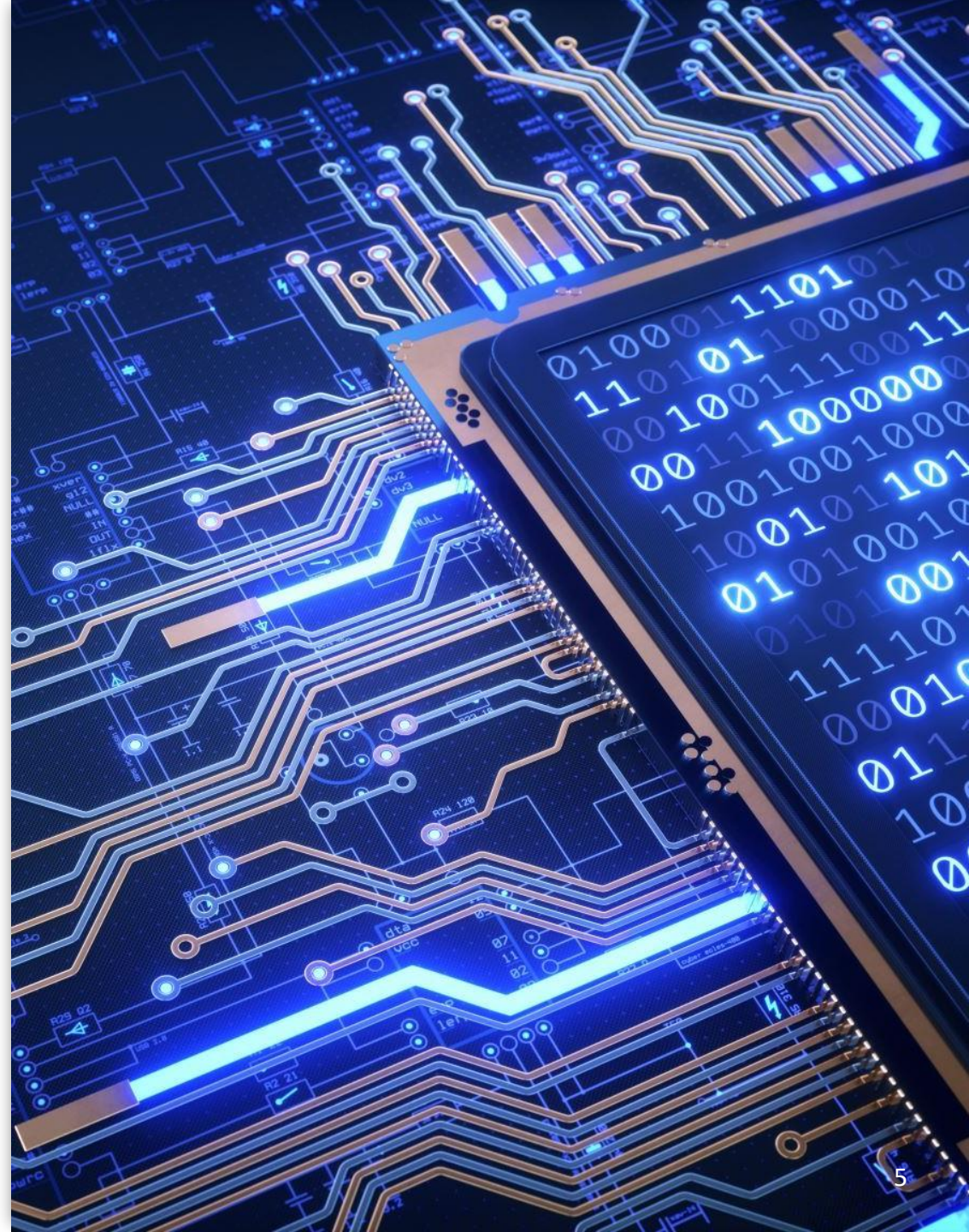
The Politics of Algorithm Design

Who has the power to decide...

- When to use an algorithm and for what purpose?

- What inputs go into an algorithm?

- Who chooses the training data that shapes the algorithm's view of the world? Algorithms make decisions for the future based on past data often embedded with bias.

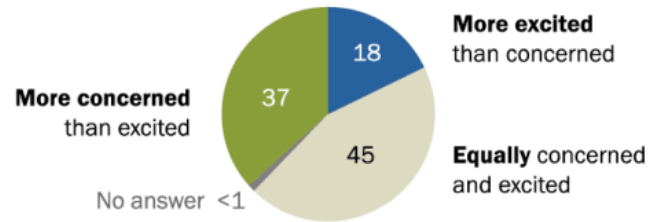- Who has property rights and over what stage of algorithm development?

For what purpose do we use Algorithms?

- To compute
- To problem solve
- To decide
- To help humans
- To monitor and control humans
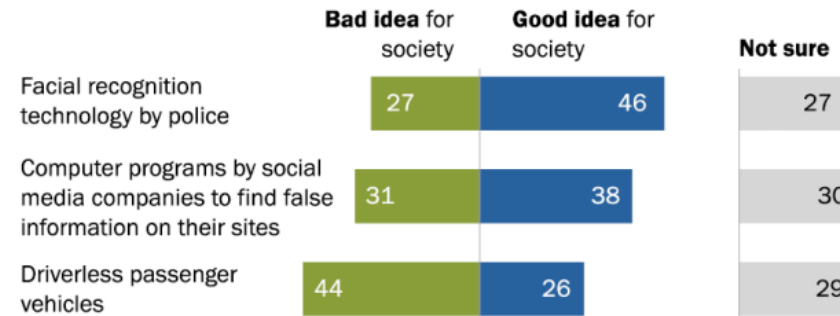- To be more than human

Americans lean toward concern over excitement when it comes to the increased use of AI in daily life ...

% of U.S. adults who say that overall, the increased use of artificial intelligence computer programs in daily life makes them feel ...

More excited than concerned: 18
Equally concerned and excited: 45
More concerned than excited: 37
No answer: <1

And public views are varied when it comes to three specific AI applications

% of U.S. adults who think widespread use of each of the following has been/would be a ...

	Bad idea for society	Good idea for society	Not sure
Facial recognition technology by police	27	46	27
Computer programs by social media companies to find false information on their sites	31	38	30
Driverless passenger vehicles	44	26	29

Note: Respondents who did not give an answer are not shown. Figures may not add up to 100% due to rounding.
Source: Survey conducted Nov. 1-7, 2021.
"AI and Human Enhancement: Americans' Openness Is Tempered by a Range of Concerns"
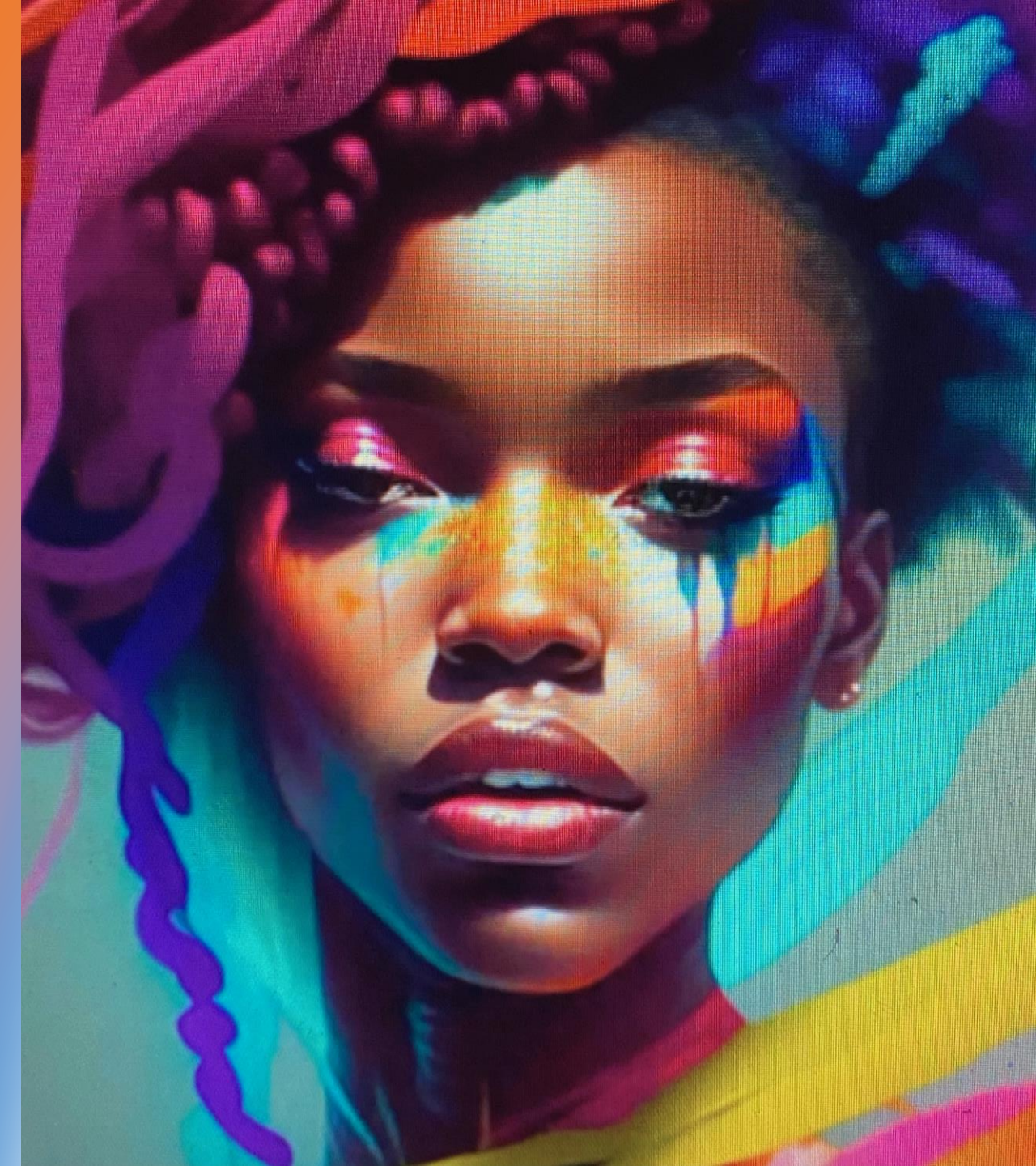
PEW RESEARCH CENTER

What is Generative AI?

- A category of AI algorithms that generate *new outputs* based on the data they have been trained on.

- Algorithms and models that can produce new content including
  - Text
  - Images
  - Audio
  - Video

- Based on patterns and relationships learned from existing datasets

ChatGPT and Disinformation: Australian Mayor Sues Open AI

- An Australian mayor said he may sue OpenAI
  - if the company does not correct ***ChatGPT's false claims*** that he was convicted of paying bribes and sentenced to 30 months in jail

- Mayor Brian Hood was never charged with any crime and he was actually the whistleblower who helped uncover an international bribery scandal linked to the Reserve Bank of Australia in the early 2000s.

- False information about an elected official

- First defamation lawsuit against Open AI

- Can ChatGPT compromise free and fair elections?

- Who can be held legally accountable when a chatbot disseminates false information?

Generative AI and ChatGPT

- Currently, ChatGPT is the most well-known generative AI model developed by OpenAI that has been trained on big datasets to answer questions in a human-like way.

- The most famous generative AI application is ChatGPT
  - a chatbot that Microsoft-backed OpenAI released in 2022

- The AI powering it is known as a *large language model*
  - because it takes in a text prompt and from that writes a human-like response.

- GPT4 is "multimodal" because it can perceive not only text but images as well.

ChatGPT and Disinformation

Generative AI responds to a text prompt with new content. Is the content produced real or fake?

ChatGPT and others also have a tendency to make things up entirely.

Creates legal issues for Open AI

A defamation case that has global implications may ensue.

ChatGPT, Disinformation, and Democracy:

- ChatGPT can spread information quickly and efficiently with the appearance of credibility.

  - When ChatGPT gives false information, is it intentional or unintentional?

  - Can it be designed to give false information and/or does it make mistakes?

  - Are the hallucinations an artifact of the algorithm?

  - Who is accountable? Who has legal personality to be accountable?

  - What are the consequences for democracy and elections in the U.S. and around the globe?
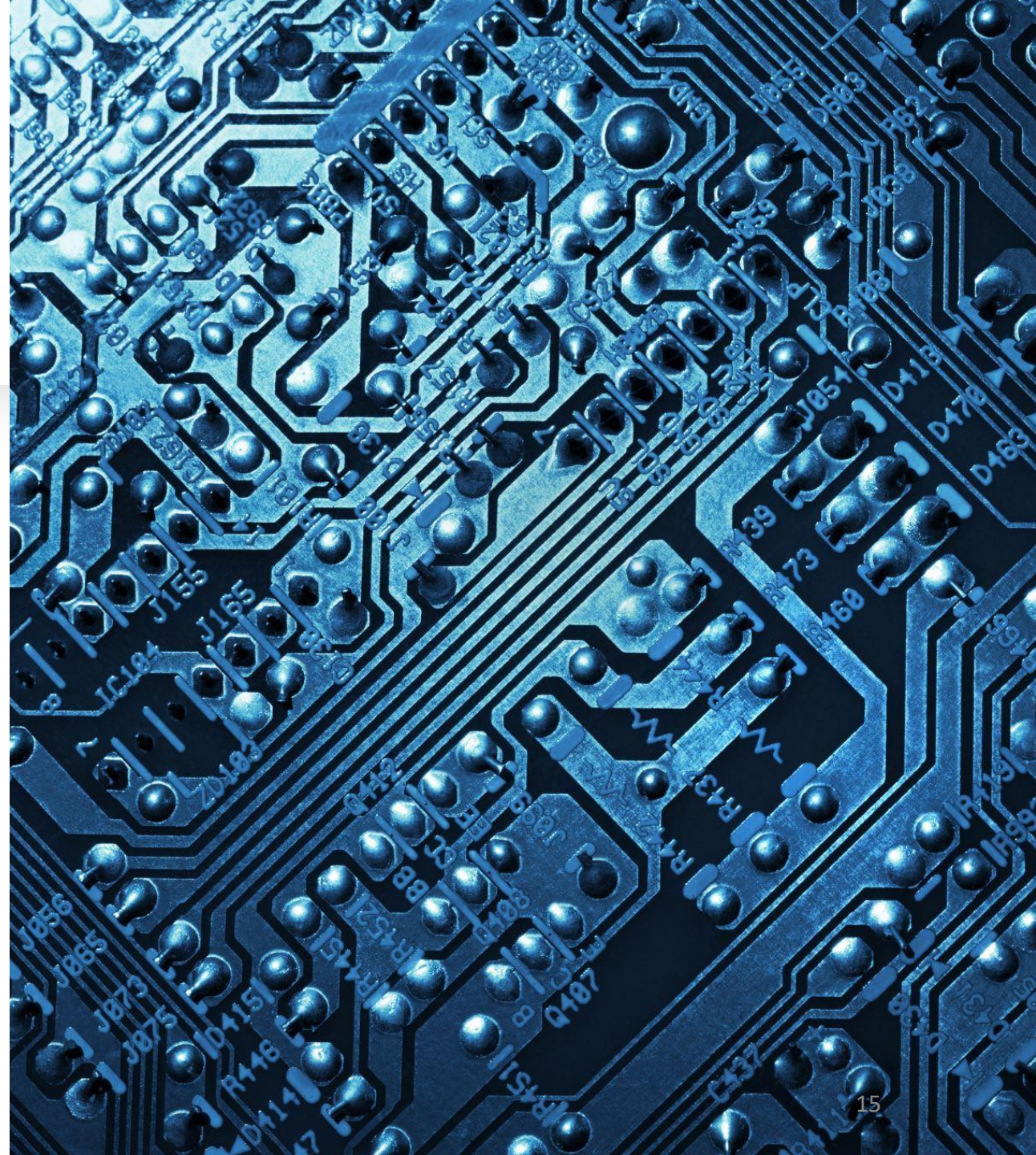
SAG-AFRA Strike, Lawsuits, and AI

- *Actors*
  - Ownership of the following in perpetuity
    - Image/physical Likeness
    - Voice

- *Screen Writers*
  - Who earns the line credit?
    - Human or computer
  - ChatGPT as a tool of screen writing or a screen writer

- *Movie companies*
  - Own the technology to digitize and own the intellectual property rights

What is Algorithmic Justice?

- Technology should benefit all of us, not the privileged few.

- How do we achieve equitable and accountable AI?
  - How are AI systems developed?
  - How to actively prevent the harmful use of AI systems.

- How do we empower communities and galvanize decision makers to take action that mitigates the harms and biases of AI?

Equitable AI: Agency and Control

- Equitable AI requires that people have agency and control over how they interact with an AI system.

- To have agency, people must first be aware of how these systems are used all around them
    - airports, stadiums, schools, hospitals and in hiring and housing

- Who is involved in creating the system?
    - business, government and academia

- The risks and potential harms.

- Equitable AI requires securing **affirmative consent** from people on how or whether they interact with an AI system.

- The idea here is that people understand exactly how their data will be used and if consent is given then their data is limited only to that permitted use.

- The defaults for affirmative consent are "opt-in," and if people elect not to opt-in, affirmative consent requires that they will not suffer any penalty or denial of access to platforms or services as a result.

- Unlike the terms of service that tech companies require people to click through to use their platforms, affirmative consent for AI cannot be coerced.
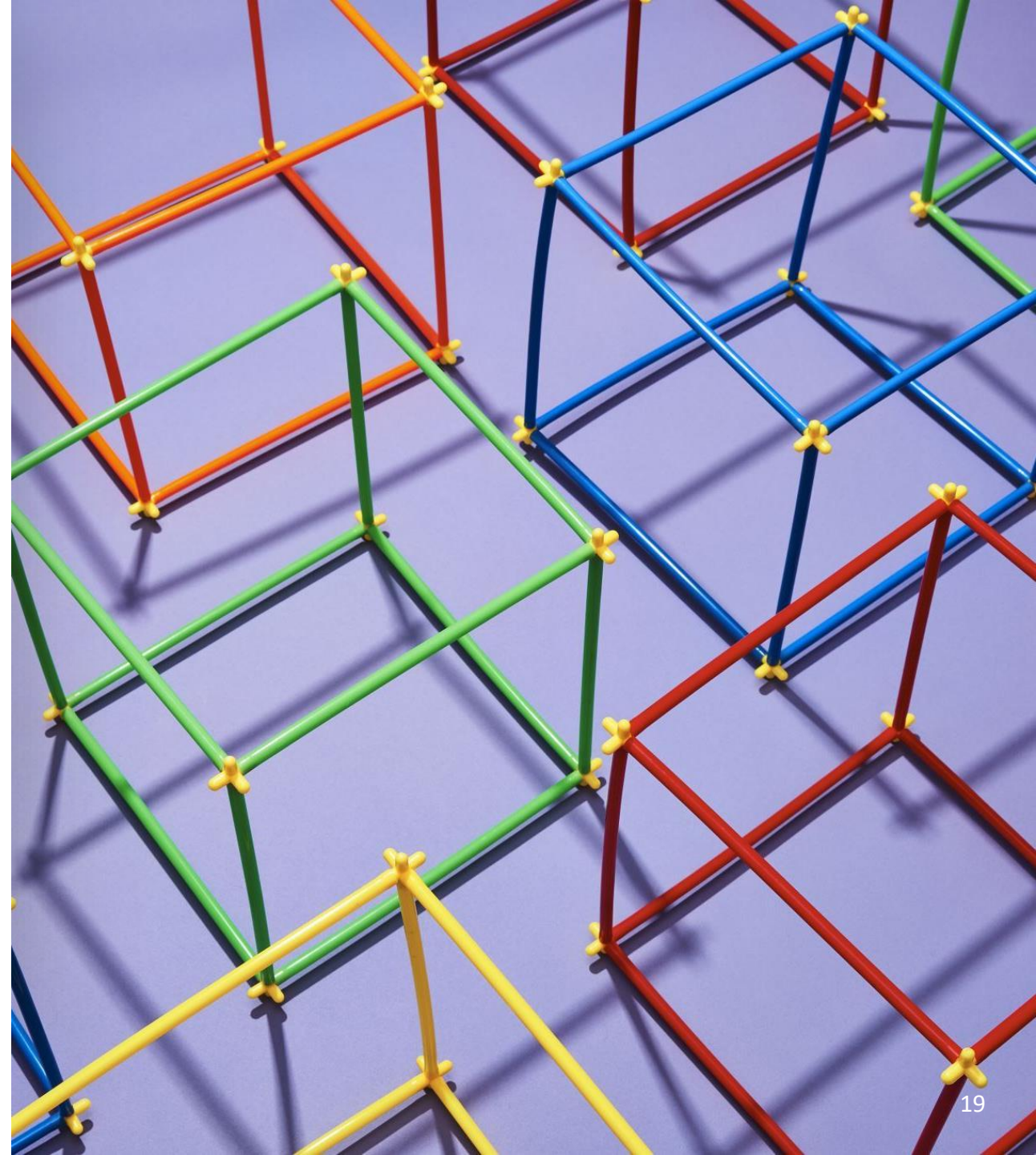
Equitable AI: Affirmative Consent

Centering Justice by Focusing on Impermissible Use

Equitable AI respects human life, dignity and right:

"…prohibiting government and commercial uses of AI. This includes AI that would identify targets when use of lethal force is an option (for example, when being stopped by the police) and use in government and citizen-led surveillance. Justice requires that we prevent AI from being used by those with power to increase their absolute level of control, particularly where it would automate long-standing patterns of injustice such as racial profiling in law enforcement, gender bias in hiring and over policing immigrant communities. Justice means protecting those that would be targeted by these systems from their misuse."- The Algorithmic Justice League

Accountable AI: Continuous Oversight

- AI systems are constantly evolving.

- Accountable AI requires continuous oversight by independent third parties.

- To support continuous oversight there must be laws that require companies and government agencies deploying AI to meet minimum requirements, for example: maintaining on-going documentation, submitting to audit requirements, and allowing access to civil society organizations for assessment and review.

Accountable AI: Redress Harms

- Accountable AI provides people who have been harmed with access to remedy, meaning that there is a working pathway for people to **contest and correct** a harmful decision made by artificial intelligence.

- For example, if an AI tool incorrectly denied a welfare benefits check, remedy would entail an easy way for the recipient to call attention to this error and receive payment plus interest for the lost time.

- If an <u>AI system was suspected of disqualifying a job applicant based on gender or race</u>, remedy would allow the applicant to discover how the decision was made and provide a basis for challenging the decision in court.

Accountability in AI: Meaningful Transparency

- How does the system work?

- How was it designed, and for what specific purpose?

- Allows people to clearly understand the intended capabilities and known limitations of the AI.

- Companies and governments must share information about how AI is being used in their own decision-making processes and sold to others.

.

- Suggestions and Questions?