

# Data Justice, Data Rights and AI Ethics

# This is us....

Farah Bennani [fbennani@usdla.org](mailto:fbennani@usdla.org)  
Twitter@drfbennani

Kae Novak [novak.kae@gmail.com](mailto:novak.kae@gmail.com)  
Twitter @kzenovka

# AI Ethics but we're still dealing with concerns about

- Data Bias
- Data Privacy
- Intellectual Property



**BIG DATA**

 <https://www.nytimes.com/wirecutter/blog/data-harvesting-by-companies>



Illustration: Wirecutter

PUBLISHED MAY 28, 2020



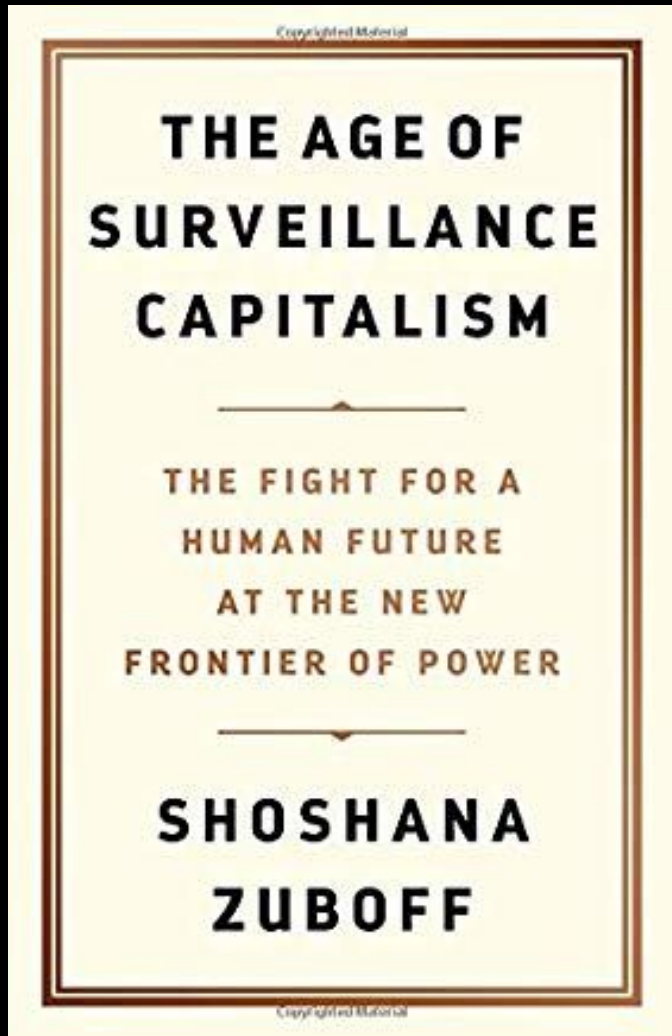
**Thorin Klosowski**

Share this post



It's a surreal experience to see all the data you've given a company in one place. Online and off, nearly every life choice you've made, every item you've purchased, or every website you've visited has been logged, categorized, and then entered in a spreadsheet to be sold off. Once it's laid out in front of you, it may make you rethink how you share information in the future.

At the start of this year, California enacted its new privacy law, the California Consumer Privacy Act (CCPA), which gives people the right to see what data



- Introduces the idea of a behavioral futures market where individual (and group data) is bought and sold based on predicted/expected earnings, success and failures (Zuboff, 2019)

## Microsoft invests \$1 billion in OpenAI in new multiyear partnership

Darrell Etherington @etherington / 7:37 AM MDT • July 22, 2019

 Comment

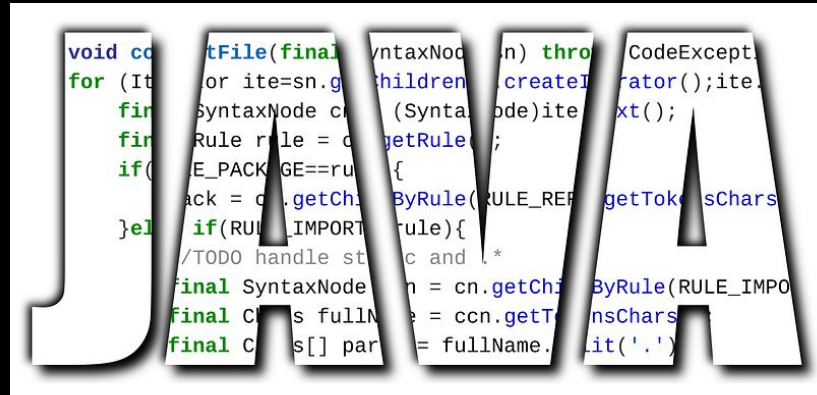


OpenAI has an unusual corporate structure, operating under a “capped-profit” model that limits backers’ returns to 100 times their investment — or possibly less in the future.

Microsoft purchased an exclusive license to the tech behind GPT-3 in 2020

# Algorithm

- Instructions that direct a computer to carry out a function or task
  - Built around AND, OR, and NOT logics



# Large Language Model

- Deep learning **algorithm** that can perform a variety of natural language processing (NLP) tasks

why are black women so



why are black women so angry

why are black women so loud

why are black women so mean

why are black women so attractive

why are black women so lazy

why are black women so annoying

why are black women so confident

why are black women so sassy

why are black women so insecure

# ALGORITHMS OF OPPRESSION

HOW SEARCH ENGINES  
REINFORCE RACISM

SAFIYA UMOJA NOBLE

- Are the proprietary algorithms, systems, and processes used by our vendors biased?
- Are aggregate results compounding stereotype and other biases?
- How do we protect students from unanticipated outcomes that end up being harmful?
- Are there equity gaps in our LMS?

# Algorithmic Accountability Act of 2019

- Would require large companies to audit their machine learning and artificial intelligence tools for bias
- Large company
  - > \$50 million per year
  - Holds data on over 1 million people or devices



# Can data be biased?



# The Right to Sanctuary

## Fourth Amendment Concerns of IoT

- Protection from unreasonable search and seizure



**The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years.**

**The regulation will fundamentally reshape the way in which data is handled across every sector, from healthcare to banking and beyond.**

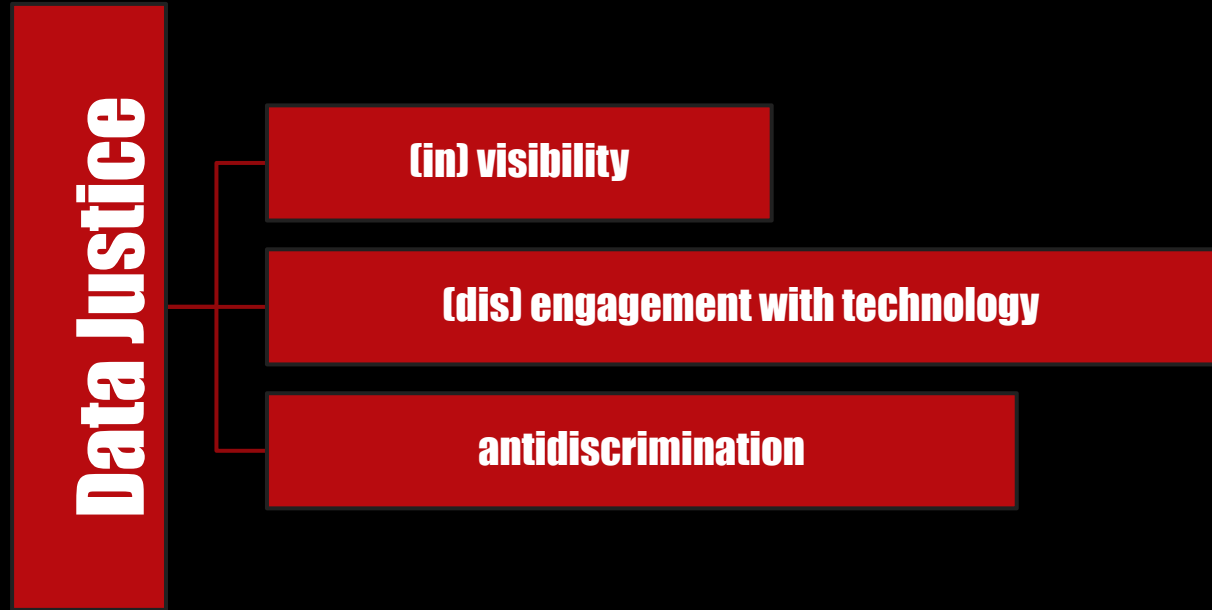


# California Consumer Privacy Act (CCPA)

<https://oag.ca.gov/privacy/ccpa>

# DATA JUSTICE

- “FAIRNESS IN THE WAY PEOPLE ARE MADE VISIBLE, REPRESENTED, AND TREATED AS A RESULT OF THEIR PRODUCTION OF DIGITAL DATA” (TAYLOR, 2017)



Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. Big Data & Society. <https://doi.org/10.1177/2053951717736335>

# Data Justice

“fairness in the way people are  
made visible, represented and  
treated as a result of their  
production of digital data”

(Taylor, 2017, p. 1)

**(In)visibility**

**Antidiscrimination**

**(Dis)engagement**

# (In) Visibility

“Despite the data being anonymous in the sense of being de-individualized, groups are becoming more transparent” (Zwitter, 2014, p. 4).



# Right to Oblivion/ Right to be Forgotten

- Do you have the right to have your information, photos, and videos deleted or archived to not show up in search engines?



# #MMIW 5100+



## Murdered & Missing Indigenous Women

# The Prison to School Pipeline of EdTech

**Thursday  
Keynote**

**Shea Swauger**  
Auraria Library, University of Colorado Denver  
*"The Prison to School Pipeline to  
EdTech"*  
**12:00PM**  
Twitter: @SheaSwauger

A portrait of Shea Swauger, a smiling person with short brown hair and glasses, wearing a light-colored shirt with a small pattern. They have visible tattoos on their forearms.

A small video feed in the top right corner showing Karin Kvedel, an older woman with short grey hair, wearing a patterned top.

# @douglevin

## The K-12 Cybersecurity Resource Center

[Home](#) [K-12 Cyber Incident Map](#) [2018 Year in Review](#) [Cybersecurity Resources](#) [Blog](#) [About](#)

[Newsletter](#) [Support](#)

[Home](#) [Blog](#) [Commentary](#) [Hacking the ISTE18 Smart Badge](#)

### Hacking the ISTE18 Smart Badge

June 15, 2018 [Written by Doug Levin](#)

**UPDATE: If you liked this post, be sure to read [Hacking the ISTE18 Smart Badge, Part II](#).**

I was curious how the ISTE 2018 Conference Smart Badge worked (and I saw other questions on Twitter speculating on the same question), so I spent a couple of hours trying to hack it. I don't claim to be any kind of expert here, but what I found (so far) was pretty interesting.

The ISTE Smart Badge is enclosed in a flat plastic case. It measures about 1" x 1/2", is about as deep as a quarter, and looks like this:

*Figure 1*



[Buy me a coffee!](#)

[f](#) [t](#) [v](#) [in](#) [p](#) [r](#) [y](#)

[+](#) [e](#) [s](#) [s](#)

#### Newsletter Sign Up

The best way to keep up-to-date with the K-12-specific cybersecurity news you need.

[Subscribe](#)

#### K-12 Cybersecurity: 2018 Year in Review

THE STATE OF K-12  
CYBERSECURITY:  
2018 YEAR IN REVIEW





**The Spook who sat by the Ring Doorbell**  @hypervisible · Oct 22 

And here a librarian—who is speaking up for students who don't have reliable internet—is afraid to provide a name because the company has a reputation of coming after people. This is textbook intimidation.

technology. Petitions with thousands of signatures have called it ableist and discriminatory, intrusive, unsafe, inaccessible, and huge invasion of privacy. Members of UBC's population were vocally opposed to Proctorio throughout the summer, in both an open letter and UBC's subreddit. (Olsen noted that "We have about 3.6 million active weekly



1



9



41





**The Spook who sat by the Ring Doorbell** ✓ @hypervisible · 23h

And what does the company say to a student who is having trouble connecting?

"Have you tried connecting to better WiFi?" 🙄 🧑 🏠

Martin had to take her first chemistry test in a Starbucks parking lot because she didn't have the bandwidth at home to accommodate Proctorio. "Support said 'Have you tried connecting to better Wi-Fi?'" Martin said. "Like that was, in any way, going to help me."



1



3



39



# Digital Sovereignty

“the autonomy of citizens in their roles as employees, consumers, and users of digital technologies and services” (Pohle & Theil, 2020, para 24).

# *Datenbewusstsein*

Critical Consciousness of how your data is being used.

(Bundesministerium für Bildung und Forschung, 2019)

# Safeguard academic integrity with an AI writing detection solution built for educators

Our advanced AI writing detection technology is highly reliable and proficient in distinguishing between AI- and human-written text and is specialized for student writing. What's more, it's integrated into your workflow for a seamless experience.

Learn more →



# Turnitin

 10 languages ▾

Contents [hide]

**(Top)**

## ▼ Functionality

- Student-paper database
- Classroom integration
- Admissions applications

## ▼ Reception

- Privacy
- Copyright-violation concerns
- Presumption of guilt
- WriteCheck

## Litigation

## ▼ Flaws

- Ad hoc encodings, fonts and text representation
- Automated paraphrasing
- Turnitin's response

## Further criticism

See also

References

External links

Article Talk

Read Edit View history Tools ▾

From Wikipedia, the free encyclopedia

**Turnitin** (stylized as **turnitin**) is an Internet-based [similarity detection](#) service run by the American company Turnitin, LLC, a subsidiary of [Advance Publications](#).

Founded in 1998, it sells its licenses to universities and high schools who then use the [software as a service](#) (SaaS) website to check submitted documents against its database and the content of other websites with the aim of identifying [plagiarism](#). Results can identify similarities with existing sources and can also be used in [formative assessment](#) to help students learn to avoid plagiarism and improve their writing. <sup>[*citation needed*]</sup>

Students may be required to submit work to Turnitin as a requirement of taking a certain course or class. The software has been a source of controversy, with some students refusing to submit, arguing that requiring submission implies a [presumption of guilt](#). Some critics have alleged that use of this [proprietary software](#) violates educational [privacy](#) as well as international [intellectual-property](#) laws, and exploits students' works for commercial purposes by permanently storing them in Turnitin's privately held [database](#).<sup>[1]</sup>

Turnitin, LLC also runs the informational website plagiarism.org and offers a similar plagiarism-detection service for newspaper editors and book and magazine publishers called [iThenticate](#). Other tools included with the Turnitin [suite](#) are GradeMark (online grading and [corrective feedback](#)) and PeerMark (student peer-review service).

In March 2019, [Advance Publications](#) acquired Turnitin, LLC for US\$1.75 billion.<sup>[2]</sup>

In the UK the service is supported and promoted by [JISC](#) as 'Plagiarism Detection Service Turnitin UK'. The Service is operated by iParadigms, in conjunction with Northumbria Learning, the European reseller of the Service.<sup>[3]</sup>

## Functionality [edit]

The Turnitin software checks for potentially unoriginal content by comparing submitted papers to several databases using a proprietary [algorithm](#). It scans its own databases and also has licensing agreements with large academic proprietary databases.

### Student-paper database [edit]

The essays submitted by students are stored in a database used to check for plagiarism. This prevents one student from using another student's paper, by identifying matching text between papers. In addition to student papers, the database contains a copy of the publicly accessible Internet, with the company using a [web crawler](#) to continually add content to Turnitin's archive. It also contains commercial and/or copyrighted pages from books,

### Turnitin



<b>Type of business</b>	Subsidiary, Privately held company
<b>Type of site</b>	Online SaaS editor
<b>Founded</b>	1998
<b>Headquarters</b>	2101 Webster Street Suite 1800 Oakland, California 94612, United States
<b>Area served</b>	Worldwide
<b>Industry</b>	Education
<b>Parent</b>	<a href="#">Advance Publications</a>
<b>URL</b>	<a href="#">turnitin.com</a> 
<b>Commercial</b>	Yes
<b>Registration</b>	Yes
<b>Users</b>	30M+ students (15,000 institutions)
<b>Content licence</b>	Proprietary



# (Dis) Engagement with Technology

“Are we going to be working for a smart machine, or will we have smart people around the machine? “ (Zuboff, 2019)



Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York, NY: Hachette Book Group

# Chihuahua or Muffin?



Source: [Chihuahua Or Muffin? Comparing Computer Vision APIs - Metamaven](#)

# Lawsuits against OpenAI

June 28, 2023: ChatGPT maker OpenAI faces a lawsuit over how it used people's data

[ChatGPT maker OpenAI faces class action lawsuit over data to train AI - The Washington Post](#)

A California-based law firm launched a class-action lawsuit against OpenAI, alleging the artificial-intelligence company that created popular chatbot ChatGPT massively violated the copyrights and privacy of countless people when it used data [scraped from the internet](#) to train its

# Lawsuits against OpenAI

September 6, 2023: OpenAI, Microsoft hit with new US consumer privacy class action.

[OpenAI, Microsoft hit with new US consumer privacy class action | Reuters](#)

OpenAI and its main backer Microsoft are facing at least their second class action lawsuit in San Francisco federal court for allegedly breaking several privacy laws in developing OpenAI's popular chatbot ChatGPT and other generative artificial intelligence systems.

The [complaint](#), filed on Tuesday on behalf of two unnamed software engineers who use ChatGPT, accuses the companies of training their fast-growing AI technology using stolen personal information from hundreds of millions of internet users.

# Lawsuits against OpenAI

November 6, 2023: OpenAI To Pay Legal Fees Of Business Users Hit With Copyright Lawsuits

[OpenAI To Pay Legal Fees Of Business Users Hit With Copyright Lawsuits \(forbes.com\)](https://www.forbes.com/sites/stevecox/2023/11/06/openai-to-pay-legal-fees-of-business-users-hit-with-copyright-lawsuits/)

OpenAI announced on Monday it would pay the legal fees of any business customer who was sued over copyright claims. The pledge does not apply to anyone using the free version of ChatGPT.

“OpenAI is committed to protecting our customers with built-in copyright safeguards in our systems.

Today, we’re going one step further and introducing **Copyright Shield**—we will now step in and defend our customers, and pay the costs incurred, if you face legal claims around copyright infringement,” the company said in a statement on its website as part of its first ever developer’s conference on Monday.

# Lawsuits against OpenAI

September 21, 2023: key issues in writers' case against OpenAI

[Key issues in writers' case against OpenAI explained – Harvard Gazette](#)

September 20, 2023: 'Game of Thrones' author and others accuse ChatGPT maker of 'theft' in lawsuit.

[George R.R. Martin, Jodi Picoult and other writers sue OpenAI - The Washington Post](#)

September 20, 2023: Franzen, Grisham and Other Prominent Authors Sue OpenAI

[Franzen, Grisham and Other Prominent Authors Sue OpenAI - The New York Times \(nytimes.com\)](#)

The suit, filed with the Authors Guild, accuses the A.I. company of infringing on authors' copyrights, claiming it used their books to train its ChatGPT chatbot.

# President Biden Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence



The Executive Order establishes new standards for AI safety and security, protects Americans' privacy, advances equity and civil rights, stands up for consumers and workers, promotes innovation and competition, advances American leadership around the world, and more.

# New Standards for AI Safety and Security

Require that developers of the most powerful AI systems share their safety test results and other critical information with the U.S. government.

Develop standards, tools, and tests to help ensure that AI systems are safe, secure, and trustworthy.

Protect against the risks of using AI to engineer dangerous biological materials by developing strong new standards for biological synthesis screening.

Protect Americans from AI-enabled fraud and deception by establishing standards and best practices for detecting AI-generated content and authenticating official content.

Establish an advanced cybersecurity program to develop AI tools to find and fix vulnerabilities in critical software.

# Protecting Americans' Privacy

Protect Americans' privacy by prioritizing federal support for accelerating the development and use of privacy-preserving techniques.

Strengthen privacy-preserving research and technologies, such as cryptographic tools that preserve individuals' privacy.

Evaluate how agencies collect and use commercially available information—including information they procure from data brokers—and strengthen privacy guidance for federal agencies to account for AI risks.

Develop guidelines for federal agencies to evaluate the effectiveness of privacy-preserving techniques, including those used in AI systems.

# Advancing Equity and Civil Rights

- Provide clear guidance to landlords, Federal benefits programs, and federal contractors to keep AI algorithms from being used to exacerbate discrimination.
- Address algorithmic discrimination
- Ensure fairness throughout the criminal justice system



# Standing Up for Consumers, Patients, and Students

Advance the responsible use of AI in healthcare and the development of affordable and life-saving drugs.

Shape AI's potential to transform education by creating resources to support educators deploying AI-enabled educational tools, such as personalized tutoring in schools.



# Supporting Workers

Develop principles and best practices to mitigate the harms and maximize the benefits of AI for workers.

Produce a report on AI's potential labor-market impacts, and study and identify options for strengthening federal support for workers facing labor disruptions, including from AI.



# Advancing American Leadership Abroad

Expand bilateral, multilateral, and multistakeholder engagements to collaborate on AI.

Accelerate development and implementation of vital AI standards with international partners and in standards organizations, ensuring that the technology is safe, secure, trustworthy, and interoperable.

Promote the safe, responsible, and rights-affirming development and deployment of AI abroad to solve global challenges.

# Critics of the executive order

The fear is that, without proper safeguards and regulations, advanced AI could be used for malicious purposes, from cyberattacks to the misuse of automation in critical infrastructure. The idea is not to make "Terminator style killer machines," but rather to ensure that AI development is guided by safety precautions and ethical considerations to prevent unintended catastrophic outcomes.

AI as an existential threat is a topic of debate, and proponents of this view often advocate for responsible AI research and development to mitigate potential risks. So, while it might not be a "doomsday scenario" in the Hollywood sense, it's a valid concern that should be addressed carefully. This is called "Safety Lobotomy".

Feel the Fear! AI Turns Deadly, Data Disappears, Criminals Clone Voices, and more

<https://www.deeplearning.ai/the-batch/issue-220/>

# Critics of the executive order

Making laws and regulations based on some imaginary future scenario is very dangerous

This fuels AI Doomers who are coming up with extremely bizarre doomsday scenarios like AI turning human bodies into bitcoin mining systems .

These crazy theories are in turn amplified by normies who hate AI because they believe it will take away their jobs.

The doomers are using the fear to gin up engagement and stay relevant while the GPU rich are using it for regulatory capture and to build AI monopolies

Who or how should we be looking at concerns...



# Engineered Inequity

- Are Robots racists?
- Are AI technologies inheriting their creators' biases?
- Is Race currently used as a form of technology embodied in robots?
- Is default discrimination a systemic glitch or a feature?
- How can we fix the coded inequity and discriminatory design?

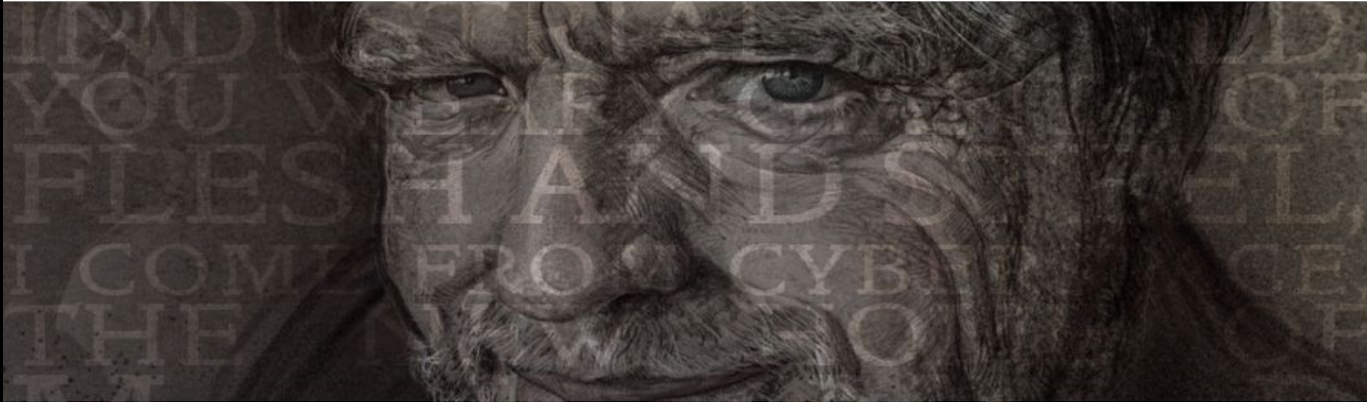
It is certainly NOT by the “I Tinker, Therefore I am” approach.

Is there a normalization of  
whiteness  
and masculinity in data  
collection and large  
language models?

# Declaration of the Independence of Cyberspace



[About](#) [Issues](#) [Our Work](#) [Take Action](#) [Tools](#) [Donate](#) [Q](#)



◀ **JOHN PERRY BARLOW  
LIBRARY**  
A DECLARATION OF THE  
INDEPENDENCE OF CYBERSPACE

A NOT TERRIBLY BRIEF HISTORY OF  
THE ELECTRONIC FRONTIER  
FOUNDATION

A PLAIN TEXT ON CRYPTO POLICY

A PRETTY BAD PROBLEM

ACROSS THE ELECTRONIC  
FRONTIER

## A Declaration of the Independence of Cyberspace

by John Perry Barlow

Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.

# Algorithm Justice League



[Algorithmic Justice League - Unmasking AI harms and biases \(ajl.org\)](https://ajl.org)

The Algorithmic Justice League's mission is to raise awareness about the impacts of AI, equip advocates with empirical research, build the voice and choice of the most impacted communities, and galvanize researchers, policy makers, and industry practitioners to mitigate AI harms and biases. We're building a movement to shift the AI ecosystem towards [equitable and accountable AI](#).

# Equitable AI



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- for example, at airports, stadiums, schools, hospitals and in hiring and housing- who are involved in creating the system- from business, government and academia- and the risks and potential harms. Equitable AI requires securing affirmative consent from people on how or whether they interact with an AI system.

# ACCOUNTABLE AI

## MEANINGFUL TRANSPARENCY

For an AI system to demonstrate meaningful transparency it must provide an explanation of how the system works, how it was designed, and for what specific purpose. Critically, meaningful transparency allows people to clearly understand the intended capabilities and known limitations of the AI.



# CONTINUOUS OVERSIGHT



AI systems are constantly evolving. As a result, 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.

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

# Data for Black Lives <https://d4bl.org/>

Annual D4BL Conference

2020:

Updates on D4BL III coming soon...

2019:



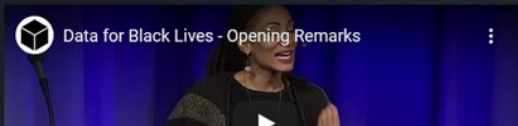
Prev



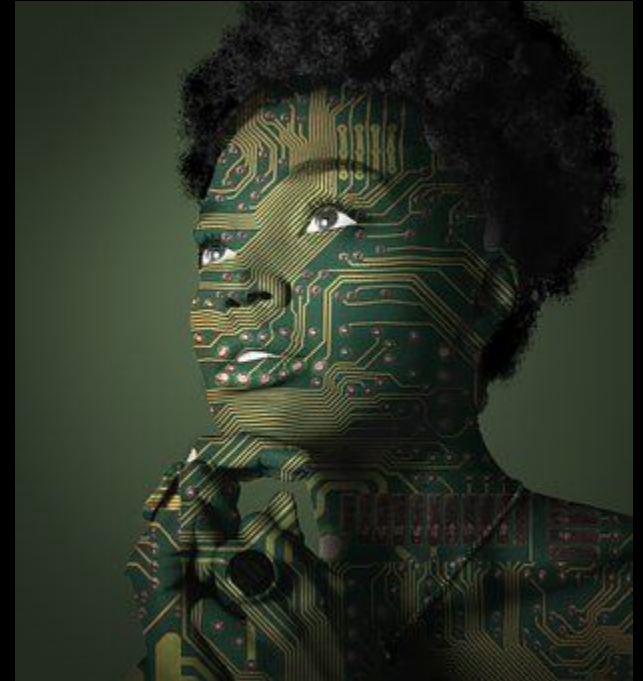
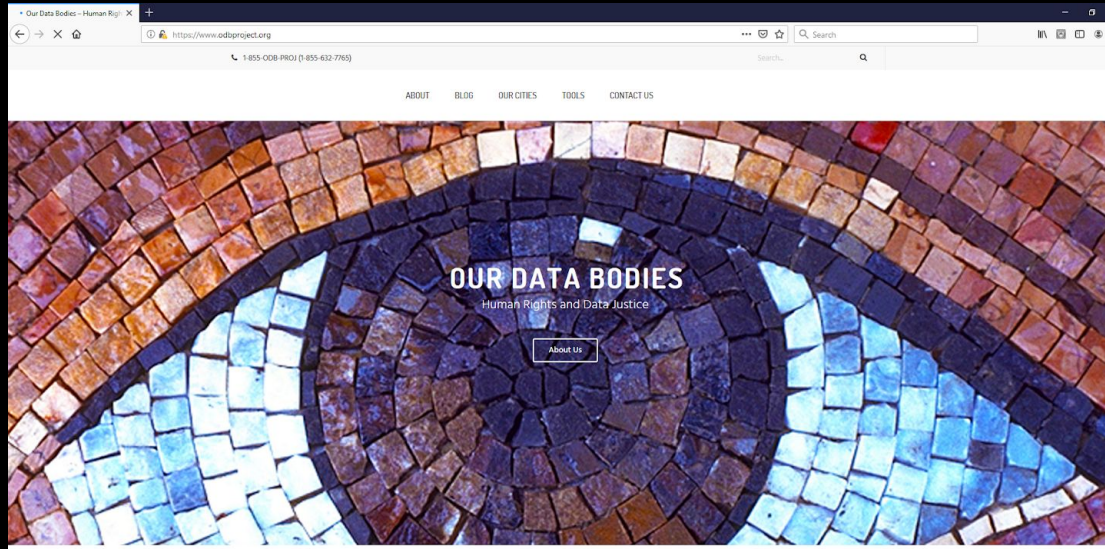
next

PROGRAM

2018:



# Our Data Bodies <https://www.odbproject.org/>





Community Power Tools  
for Reclaiming Data

Our Data Bodies: <https://www.odbproject.org/tools/>



Worried about the Working Conditions  
not the Technology

# Contact Us!

Farah Bennani [fbennani@usdla.org](mailto:fbennani@usdla.org)  
Twitter@drfbennani

Kae Novak [novak.kae@gmail.com](mailto:novak.kae@gmail.com)  
Twitter @kzenovka

**Keep the Internet weird.**



**Defend net neutrality.**