Health Data Ethics: Understanding Big Data, Algorithmic Bias, and GenAI

Taught By: Shannon Peifer and Dipa Desai
Digital Integration Teaching Initiative (DITI)
PHIL 1165
Professor Shorey
Spring 2026

Northeastern University
NULab for Digital Humanities and Computational Social Science

Workshop Objectives:

- Understand the ways in which data reflects and reinforces cultural, social, and political biases.
- Explore ways of interpreting and effectively utilizing data-based evidence in written arguments.
- Explore moral and social considerations for working with big data, algorithms, and AI in healthcare spaces.
- Explore the ways in which these questions and methods are influencing how humanists and social scientists do research.

Handouts and slides are available at: https://bit.ly/sp26-shorey-data_ethics

Questions to consider:

- When have you seen healthcare data **represented in your daily lives**?
- Where is health data coming from and how is it being **collected**?
- How is our data being used within the **healthcare industry**? For what purposes?
- What are some potential **uses and benefits**?
- What are some potential **ethical concerns**?

What is "Big Data"?

Defining Big Data

Companies, governments, and other groups **collect vast amounts of data from vast numbers of users** and analyze that data quickly for a variety of purposes—including advertising, marketing, surveillance, building profiles, etc.

The goal of big data is **to predict individual user behavior based on patterns from the user as well as patterns from "similar" users** (based on demographic information, behavioral patterns, etc).

We're living in an era of "surveillance capitalism"—**our information is a valuable *product*.**

Why care about Big Data?

- Big data is **omnipresent**—its **sources** include: digitized records, internet activity, and even sensors from the physical environment.
- Big data is often **privately owned** and it is hard to ensure oversight over how it is developed, used, and controlled.
- The **scale** of big data enables those who use, develop, and control it to **magnify** their influence.
- Some websites **monetize** data in a "data exploitation market," selling their users' personal information.
- Big data can be used to (inadvertently or purposefully) **entrench stereotypes** or **reproduce results** that harm individuals and communities.

Big Data in the healthcare industry

- Healthcare organizations are attempting to generate and analyze big data to "revolutionize the medical therapies and personalized medicine." ([Dash, et al](#).)
- Common sources include: electronic health records (EHRs), electronic medical records (EMRs), personal health records (PHRs), and data produced by wearables and mobile health apps ([Coursera](#))
- Analyzed with hopes of improving patient outcomes, predicting disease outbreaks, and decreasing health care costs ([Coursera](#))
- Challenge of ensuring patient privacy and data security through anonymization and encryption ([Abouelmehdi, et al](#).)

Healthcare Data Ethics: Big Data, Online Presence, & Data Privacy

Data Privacy

- It's important to pay attention to data privacy when using digital resources
- At its simplest, **data privacy** is a person's ability to control what of their personal information is shared and with whom.
- To help you make informed decisions about interacting with digital tools in ways that honor your boundaries with your data and/or personal information, the DITI has prepared a **Data Privacy handout**.

How do we contribute to Big Data?

- Entertainment media
- Healthcare and medical services
- Shopping and marketing
- News and information

- Social media
- Travel and transportation
- Education and employment
- Public policy and safety

Personal information can be collected by digital platforms and tools, causing data privacy issues as well as concerns of using personal data for company profit.

How does Big Data impact our daily lives?

- One example of big data related to health is wearable tech, such as smartwatches and fitness trackers, which are marketed as helpful tools to monitor health and fitness. They've also been linked to proactive patient care and improved clinical outcomes ([Mayo Clinic](#)).
- But in the workplace, wellness programs that capture this data by rolling out these devices with employees raises some ethical and privacy concerns. While wearable tech may collect useful information on employee health and well-being, a lack of transparency and insecure data storage can normalize over surveillance and increase workplace stress ([Inc. Magazine](#)).

Wearable Tech Data Collection

- *Nature,* "Privacy in consumer wearable technologies: a living systematic analysis of data policies across leading manufacturers"
- *Maryland Healthcare Commission,* "Data Privacy When Using Wearable Health and Fitness Devices"
- *Brown University Center for Digital Health,* "Privacy Data Ethics of Wearable Digital Health Technology"

Issues in Big Data: Who gets represented?

Technology is Not Neutral

- Information systems, data collection and analysis platforms, and algorithms for working with big data are **not neutral**.
- They can **reinforce** systemic, political, and cultural **biases**.
- They are **affected by input data**, the way that data is presented, how the data is interpreted by machines, and more.
- This means **we also have the ability to challenge these biases**, norms, and forms of discrimination.

"Greatest Authors of All Time"

Open Google's search engine and type in "Greatest authors of all time."

- What are some of the results? What do you notice about these results?
- Where do you think these results came from?
- How many authors on this list have you read? Do you agree with the list?
- What do these results suggest to you in terms of defining "greatest" and "authors"?

"Greatest [blank] Authors of All Time"
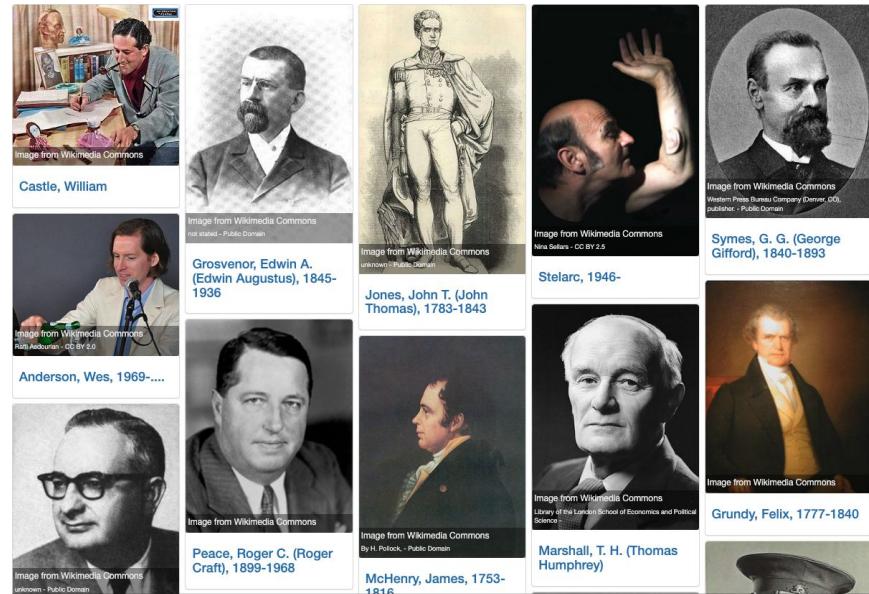
Now try these searches:

- Greatest women authors
- Greatest Black women authors
- Greatest Black authors
- Greatest white authors

"Black" leads to substantial results, while "white" does not. Why do you think this might be?

Representation in digital archives

- [SNAC](#) (Social Networks and Archival Context) is an online resource that compiles historical information.
- Click on the link above to explore SNAC and try refreshing the page a few times. What do you observe about the kinds of people who appear each time you reload the page?

[SNAC](#)

Representation in healthcare research

- Usefulness of social media in public health research is limited by demographic bias in "who chooses to use each platform and what they choose to share" ([WebSi '19](#))
- Typically relies on electronic health record (EHR) data which presents a challenge with "bias from incomplete data, poor data quality, lack of representation from people who are medically underserved, and implicit bias in EHR design" ([DH '18](#))
- To promote health equity, researchers must consider strategies to mitigate bias and ensure accurate diagnosis and treatment plan

Issues in Big Data: Algorithmic Bias

Algorithms: Introduction

- An algorithm is a set of instructions given to a computer, typically to solve a problem, perform a computation, or to make a decision.
    - Computers are given an input, then they follow an algorithm, which leads to some kind of output.
- We all follow sets of predefined actions to achieve desired results.
    - For example, think about the specific steps you follow to make coffee. That's also an algorithm! Algorithms are everywhere, even when we don't realize it. They're just a set of rules or steps that get done over and over again (very quickly, if done on a computer).

Algorithmic Bias

- Algorithms are *not neutral*. **People create algorithms**.
  - Algorithmic processes—and even the data itself—reflect societal biases.
- When an algorithm is written or trained using data that misrepresents the actual population, this produces **algorithmic bias.**
- Similarly, **when data reflects biased realities**, the algorithm will continue to reproduce outcomes based on those biases (despite their harm to—or erasure of—other groups).
- Algorithms reflect social inequalities, and can even exacerbate them.
- Read this [Vox article](#) for more information on algorithmic bias.

Algorithmic Injustice: Healthcare

- Algorithms are used in public health systems to inform decisions on who should receive preventative care and medical treatment.
- They predict the likelihood of specific conditions, e.g. cardiovascular risk, or of general ill-health among different demographic groups.

Information Source: "Algorithmic Bias in Health Care Exacerbates Social Inequities - How to Prevent It"

Image Source: "AI Could Worsen Health Disparities"

Algorithmic Injustice: Healthcare Data

- The data used to make these predictions is often collected from white patients, which makes risk scores far less accurate for Black or other non-white patients.
  - The Harvard School of Public Health estimates that Caucasians make up 80 percent of collected data in the fields of genomics and genetics.

"We found that a category of algorithms that influences health care decisions for over a hundred million Americans shows significant racial bias."
Sendhil Mullainathan, Chicago Booth University

Source: Katherine Igoe, Harvard TH Chan School of Public Health, 2021

Northeastern University
NULab for Digital Humanities and Computational Social Science

Alleviating Injustice: Development

- When we look at the data used to train an algorithm, we must ask **what kinds of data** are being counted, and what kinds of data are being **overlooked**, **ignored**, and **excluded**?

> **"Algorithms by themselves are neither good nor bad. It is merely a question of taking care in how they are built."** Sendhil Mullainathan, Chicago Booth University

Alleviating Injustice: Counting

- What are the consequences of counting and not counting different kinds of data on various populations, especially marginalized groups?
- Ask: will the data-driven solution **reduce** human bias or **amplify** it?

> **"Counting and measuring do not always have to be tools of oppression. We can also use them to hold power accountable, to reclaim overlooked histories, and to build collectivity and solidarity."** Catherine D'Ignazio & Lauren Klein, *Data Feminism*, 2020

Issues in Big Data: Generative AI

Generative AI: Introduction

AIs like ChatGPT and Bing Chat (which generate text) and DALL-E and Midjourney (which generate images) work by "training" on very large datasets of information and then remixing that training data to produce "new" outputs.

Generative AI: How an AI Works

- AIs are fundamentally based on statistics. A text-generation AI like ChatGPT generates words that "sound right" based on the context you provide and what it's seen before.
- As a result, AI outputs are strongly shaped by bias from what was included in the training data.
  - They are trained mostly on non-paywalled internet content, which is often different from paid journalism, scholarly articles, and books.

Examples of Non-Paywalled Content

- [CBC](#)
- [BBC](#)
- [The Guardian](#)

Here's an article to understand a little more about paywalls and news sources from *Columbia Journalism Review*, "[In paywall age, free content remains king for newspaper sites](#)"

Generative AI: Training Ethics

- To reduce the chance that AIs will generate reprehensible content, AI companies typically rely on a secondary training phase in which human workers identify such content.
  - For example, OpenAI (the maker of ChatGPT) outsourced this task to Sama, a company which claims to produce "ethical AI" by paying workers in Kenya, Uganda, and India to label violent, pornographic, and discriminatory content.

Generative AI and "Truth"

- Text-generation AIs aim to produce text that is grammatically correct and linguistically probable.
  - They do not understand "facts," only patterns of word use.
- They can generate truthful text, but also frequently create falsehoods.
  - When asked to generate citations, they can generate plausible-looking but fake sources, known as [confabulations](#) or hallucinations.
  - They may link real but irrelevant sites as sources for made-up facts. They may also invent URLs that do not work and have never worked.

Generative AI: "Originality" Ethics

- Some argue that all AI-generated output constitutes plagiarism and copyright infringement, since it is remixing training data that was scraped from the internet without permission from the original creators.
  - Many AI companies are facing lawsuits from people whose content was used as training data without their consent.
  - Some publication venues, like the *Science* journals, have made it an official policy that AI does not meet the standard of originality for authorship and that AI-generated text is unacceptable.

Generative AI & Healthcare

- Generative AI models may be applied for clinical decision support and alleviate administrative burdens (for example, AI scribes). But issues of accuracy and reliability of AI-driven decisions and ethical considerations over patient data are critical concerns as these AI-enabled healthcare solutions are rolled out (Zhang, et al., 2023).

- Despite these and other ethical, legal, and privacy issues, generative AI is pitched as a tool to improve healthcare research, decision-making, outcomes (Okonji et al, 2024; Sai et al., 2024).

Generative AI Impact on Health

- The data centers powering generative AI models cause widespread pollution, creating costly public health issues across the US ([Han et al., 2025](#)).
- Similarly, data centers use substantial amounts of freshwater and reduces available water for human consumption ([Yañez-Barnuevo, 2025](#)).
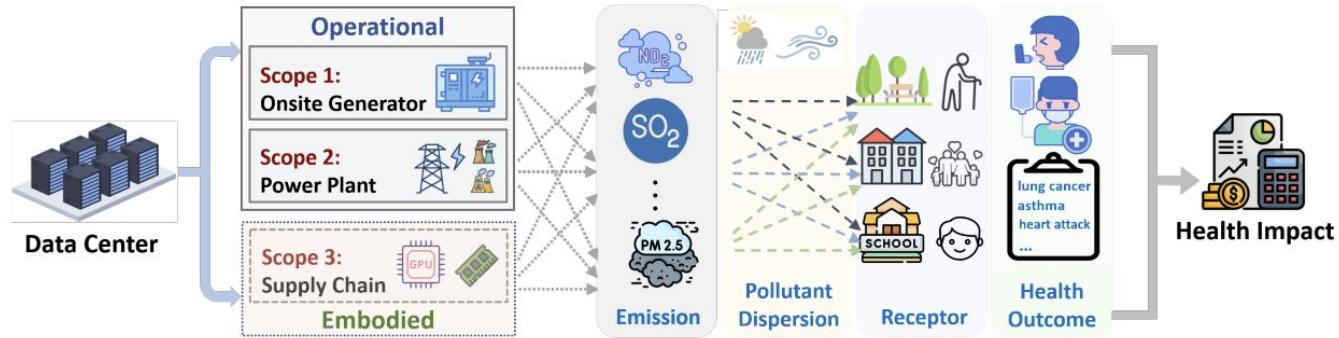
Image source: [Han et al., 2025](#)

NULab Faculty Research on Data, Algorithms, & AI

- "[John Wihbey and Christo Wilson on Tiktok Data Espionage Concerns](#)"

- "[Alan Mislove Co-Authors Research on Discriminatory Ad Algorithms](#)"

- "[John Wihbey Weighs In On AI's Potential to Impact the 2024 Presidential Election](#)"

- "[Tina Eliassi-Rad Co-Creates New AI Model that Predicts Human Lifespan](#)"

- "[Nabeel Gillani Interviewed by Tech Talk Podcast on AI and Education](#)"

Moving Forward

How can we use data responsibly?

Questions Researchers Must Ask

- What **information** is being collected and from where? To whom does this data **belong**?

- How is it being **collected**? Do **participants** know that it is collected, how it will be collected, and how will it be used?

- **How** will the data be analyzed? What **biases** and **ideologies** may be implicit in this analysis?

- Who will this research impact? Who will it **benefit**? Who will it potentially **harm**?

Activity: Apply Research Questions

- Let's apply the research questions from the previous slide to a recent example of data privacy ethics.

- Ancestry.com sold the DNA and pre-existing condition data of 18 million people to a private equity firm that also owns health sciences businesses.

- Read through [this article](#) and try answering some of the questions with one or two people near you.

Collecting Data

Which of these is likely to collect more accurate and representative data about users' genders?
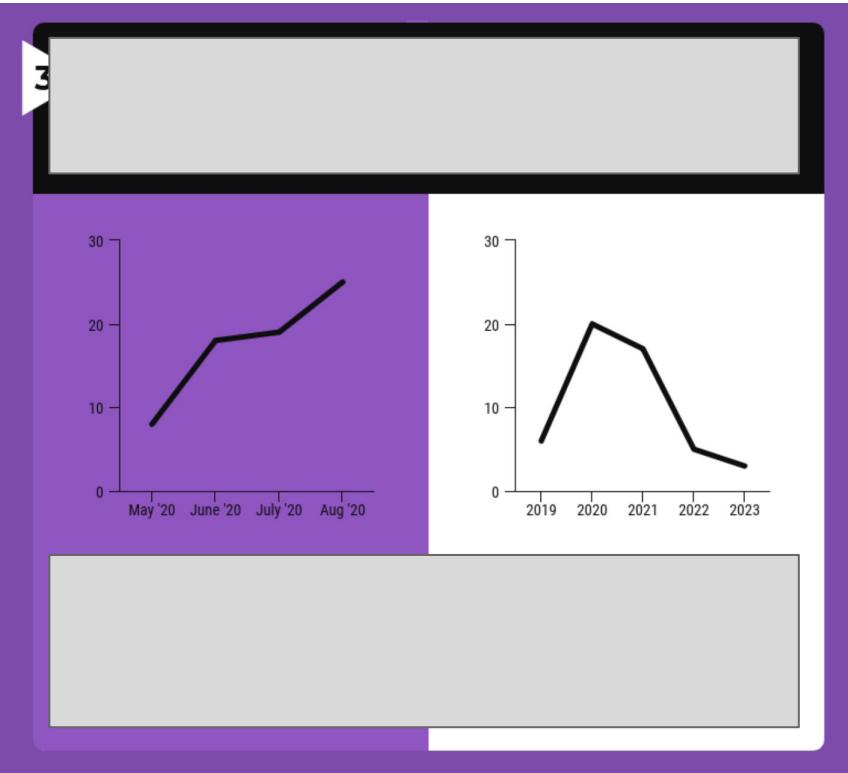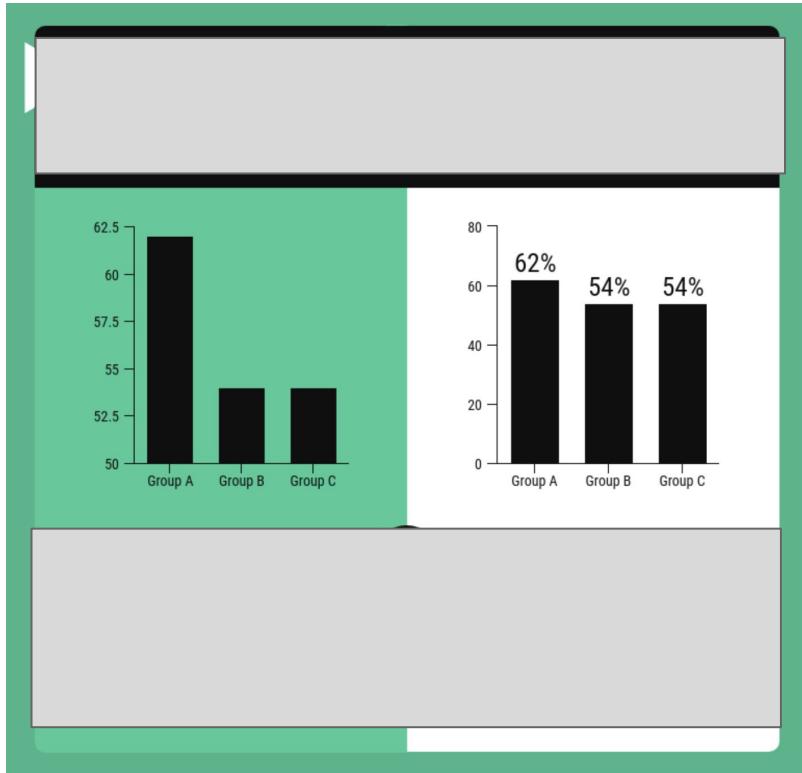

Source: Positive Voices survey from Public Health England, published in D'Ignazio & Klein, *Data Feminism*, 2020

Source: Facebook's account creation page circa 2018, published in D'Ignazio & Klein, *Data Feminism*, 2020

Responsibly Using Health Data

- Be **data-literate**—turn a critical eye to studies that use big data, evaluate the sources of that data, and carefully examine the conclusions authors draw from their sources.
- Be **thoughtful** and **intentional** as you incorporate big data or conclusions drawn from big data sources in your work. Think:
  - Could this evidence be interpreted in a different way?
  - Is this the strongest evidence I could use to support my claim?
  - Is the way I'm presenting this information accurate, or could it be considered in any way **misleading**?
  - Does the data accurately represent issues across patient groups?

Be Mindful of Infographics and Data Visualizations

Image Credits: Venngage.com

For Further Exploration

[Copyright and fair use handout](#)

[Data Ethics handout](#)

[Data Privacy handout](#)

Thank you!
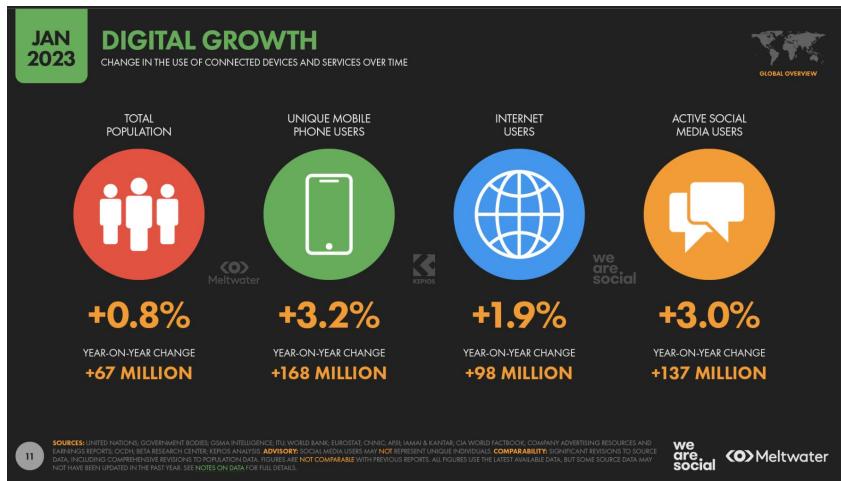
—**Developed by:** Shannon Peifer, Dipa Desai, Claire Tratnyek, Vaishali Kushwaha, Yana Mommadova, Colleen Nugent, Tieanna Graphenreed, Javier Rosario, Ana Abraham & Chris McNulty

- For more information on DITI, please see: https://bit.ly/diti-about
- Schedule an appointment with us! https://bit.ly/diti-meeting
- If you have any questions, contact us at: nulab.info@gmail.com
- We'd love your feedback! https://bit.ly/diti-feedback
- Handouts and slides: https://bit.ly/sp26-shorey-data_ethics

Appendix

Big Data is getting bigger: Internet Use

- Internet usage is constantly increasing: 64.4% of the world's population regularly uses the internet, and 59.4% of the population uses **social media**.

- All sorts of data is collected about online audiences and their activities.
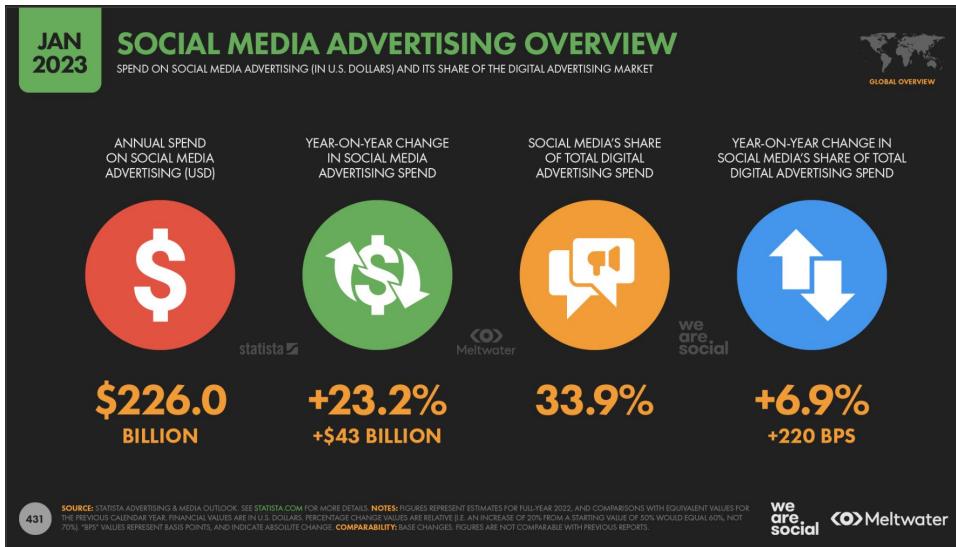
Source: DataReportal

Big Data is getting bigger: Advertising

- The big data collected allows advertisers to **target** users.
- **Spending** on social media advertising is also increasing: $226bn was spent globally in 2022.

Source: DataReportal

Example of Big Data in our daily lives

AWARENESS | SCIENCE & TECH | AUG 3, 2019 AT 11:08 AM.

Google's File on You is 10 Times Bigger Than Facebook's — Here's How to View It
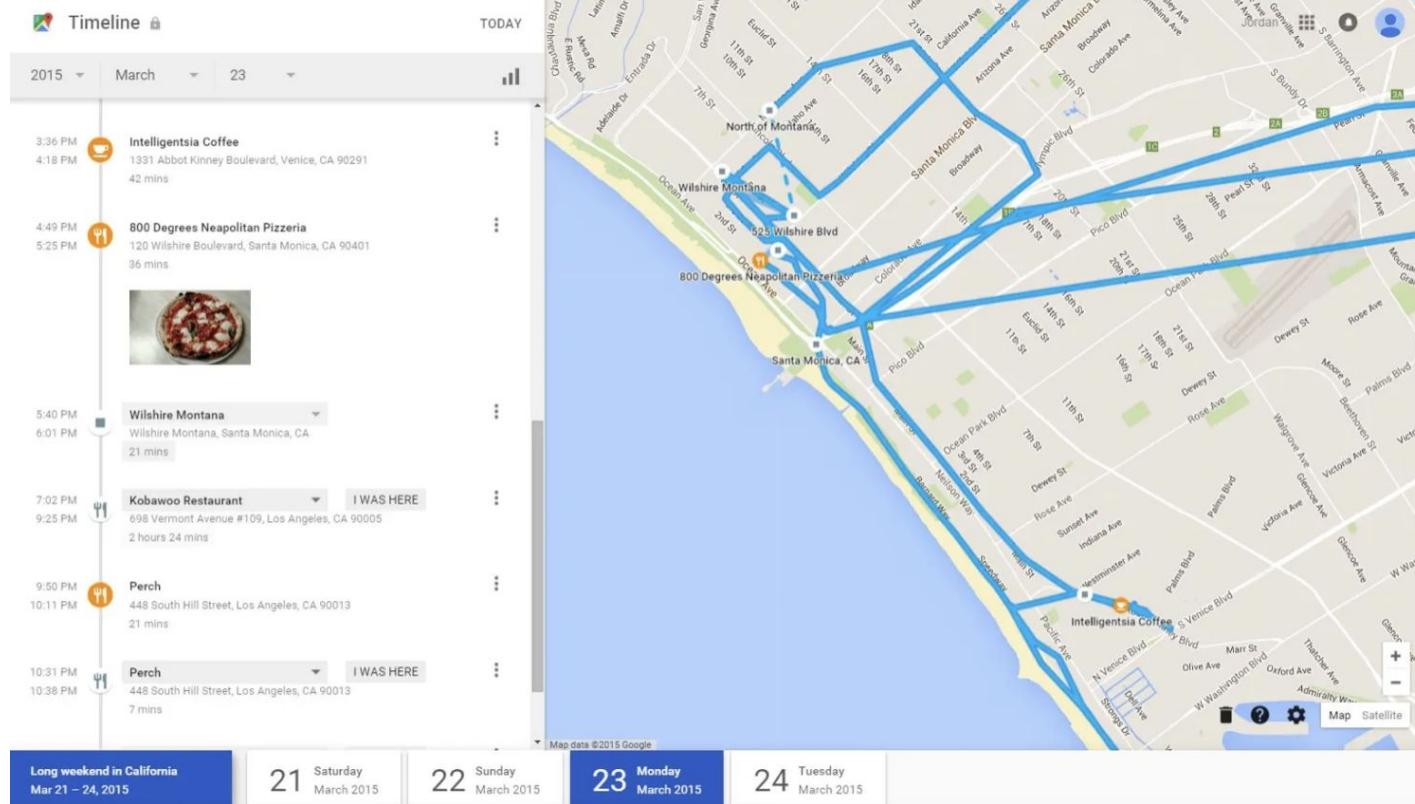
Google, Amazon, Apple, and Microsoft are all central players in "surveillance capitalism" and prey on our data.

THE ANTI-MEDIA

If you have **location services** turned on for Google (for instance, if you use Google maps), Google can track your every move. Go to:

https://www.google.com/maps/timeline

Example of Google Maps' Timeline

How Do We Know When We Are Being Tracked?

- There are ways to identify what information websites collect about their users.
- Be sure to access the Terms and Conditions offered by a website to see what the site is disclosing about how they track you.
- [Blacklight](#) is a "real-time website privacy inspector" developed by *The Markup*, a nonprofit publication that investigates data misconduct. You can use it to scan and reveal the specific user-tracking technologies on any site.

Sample Terms & Conditions:

Categories of information we collect

Identifiers and similar information such as your name and password, phone number, and address, as well as unique identifiers tied to the browser, application, or device you're using. Some Google services like YouTube Studio provide the option to submit a valid ID (such as a passport or driver's license) to verify your identity to use additional features.

Demographic information, such as your age, gender and language. If you choose to use optional features like YouTube Creator Demographics, you can also provide additional information, like your gender identity or race and ethnicity.

Commercial information such as your payment information and a history of purchases you make on Google's services.

Biometric information if you choose to provide it, such as fingerprints in Google's product development studies.

Internet, network, and other activity information such as your search terms; views and interactions with content and ads; Chrome browsing history you've synced with your Google Account; information about the interaction of your apps, browsers, and devices with our services (like IP address, crash reports, and system activity); and activity on third-party sites and apps that use our services. You can review and control activity data stored in your Google Account in My Activity.

Geolocation data, such as may be determined by GPS, IP address, and other data from sensors on or around your device, depending in part on your device and account settings. Depending on these settings, this may include **precise location data**, for example GPS data for Android features like navigation or finding your phone. Learn more about Google's use of location information.

Audio, electronic, visual, and similar information, such as voice and audio information.

Communications data, such as emails, if you use our services to send and receive messages.

Health information if you choose to provide it, such as your medical history, vital signs and health metrics (like blood glucose levels), and other similar information related to your physical or mental health, in the course of using Google services that offer health-related features, such as the Google Health Studies app, Fitbit, or Pixel devices..

Downloading Your Data & Tightening your Privacy

- **Facebook**: Settings >Settings & Privacy > Account Center > Your Information> Download your Information
- **Google**: [How to download your data](#)
- **Instagram**: Settings > Settings & Privacy > Account Center > Your Information and Permissions > Download your information
- **TikTok**: Profile > 3-line Menu icon (top right) > Settings & Privacy > Account > Download your data

Algorithms & Big Data: *What gets counted counts*

D'Ignazio and Klein identify several problematic data practices that cause real harm:
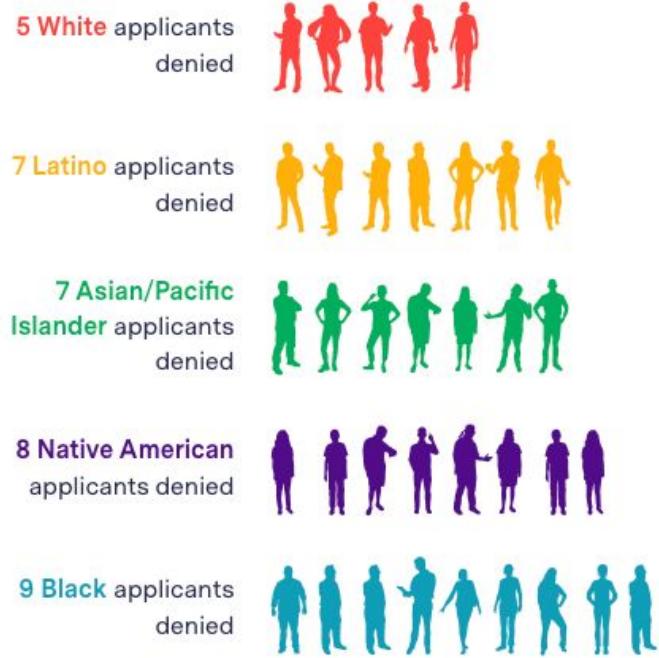
- Lack of quantitative research on maternal mortality masks systemic problems.
- Undocumented immigrants are often (sometimes voluntarily) absent from census data, which determines levels of federal funding: a "paradox of exposure."
- TSA scanning machines binarize bodies to attempt to uncover concealments, but can thereby mistakenly assign risk alerts.

"What is counted—like being a man or a woman—often becomes the basis for policymaking and resource allocation. By contrast, what is not counted—like being nonbinary—becomes invisible..."
Catherine D'Ignazio & Lauren Klein, *Data Feminism*, 2020

Algorithmic Injustice: Mortgages

- Mortgage approval algorithms can gather and use data in ways that express a racial bias.
- On Fannie & Freddie, which buys about half of all mortgages in America:
  "This algorithm was developed from data from the 1990s and is more than 15 years old. It's widely considered detrimental to people of color because it rewards traditional credit, to which white Americans have more access."

5 White applicants denied

7 Latino applicants denied

7 Asian/Pacific Islander applicants denied

8 Native American applicants denied

9 Black applicants denied

Source: Emmanuel Martinez & Lauren Kirchner, *The Markup,* 2021

Northeastern University
NULab for Digital Humanities and Computational Social Science

AI and Automated Plagiarism Checkers

- Some companies sell tools that claim to identify whether text is AI-generated or human-generated.
  - They do this by calculating how statistically "predictable" each word in the text is. If a text consistently uses the most predictable next word based on the preceding words, it is labeled as "probably AI."

Plagiarism Checkers: False Positives

- No matter how advanced they become, these tools will always have the potential for "false positives" (identifying human texts as AI), which can be catastrophic for students.
  - There is no way to check whether a text is "predictable" because an AI wrote it, or because a human writer wasn't very original. Sometimes humans just *are* predictable!

Plagiarism Checkers: Biases

- False positives are especially likely for novice writers, for highly-formulaic genres of writing, and for writing that mainly reiterates established facts and definitions (in other words, a lot of university writing).

- False positives are also very common for texts by writers for whom English is not their first language.
  - To make things worse, these writers can often *reduce* their "AI score" by using an AI to reword their essays!

(Cont.) NULab Faculty Research on Data, Algorithms, & AI

- "John Wihbey Participates in a Panel on Content Moderation"
- "John Wihbey Comments on Google's New 'AI Overview'"
- "John Wihbey on the Politics of AI"
- "John Wihbey Interviewed on AI and Epistemic Risk"
- "Malik Haddad on the Regulation of AI"

Finding and Using Non-Traditional Sources

Some kinds of non-traditional and/or non-academic sources:

- Public Media (written/broadcast journalism)
- Crowdsourced projects (including Wikipedia, aggregate reviews, etc.)
- Multimedia sources (including social media and blog posts)
- Prof. Eunsong Kim's *The Politics of Trending*
- Oral histories and interviews
- Indigenous forms of knowledge

Vetting and Citing Non-Traditional Sources

Regardless of the type of source you're using, you should always:

1) Try to **verify the information** presented in the source by finding other (independent) sources that support it.
2) Be clear in your writing about what kind of source it is, where you found it, and how you're using it (be explicit about your **process** and the source's **purpose**).
3) **Cite your source** appropriately so that any reader can find it.