Data Ethics: Understanding Big Data and Algorithmic Bias

Digital Integration Teaching Initiative



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 the ways in which privacy and security are being redefined through the use of big data, algorithms, and policy.
- Explore the ways in which these questions and methods are influencing how humanists and social scientists do research.

For more information, please see: https://bit.ly/handout-data-ethics

What is "Big Data"?



Questions to consider:

- How are we being represented online?
- Where is data about our lives coming from, and how is it being collected?
- Who is using our data and for what purposes?
- How might our data be used in the future?
- How does "big data" impact our daily lives?

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*.



Volume, Variety, Velocity, and Veracity

- Big data is large in volume, with thousands, millions, or even billions of observations.
- The second component of big data is **variety**. Big data is often collected from many sources and combined together, providing new and more comprehensive datasets.
- The third component is **velocity**. Big data is usually collected in real-time, without people necessarily knowing that they are being observed.
 - For example, transportation-related big data could be the result of people across
 Boston swiping their CharlieCard to board busses and trains.
- The fourth component is **veracity**. We should always be wary of the truthfulness and generalizability of big data because it is not necessarily representative. We need to understand the limits of what big data is and what it is telling us.

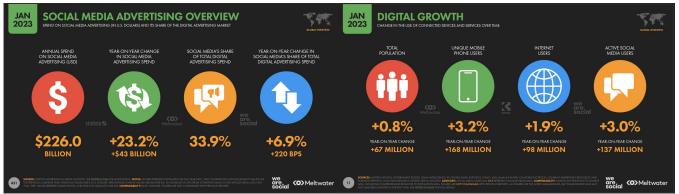


Source:

Big Data is getting bigger



- 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.
- 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



Why should we 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, Online Presence, & Data Privacy

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

How does Big Data impact 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

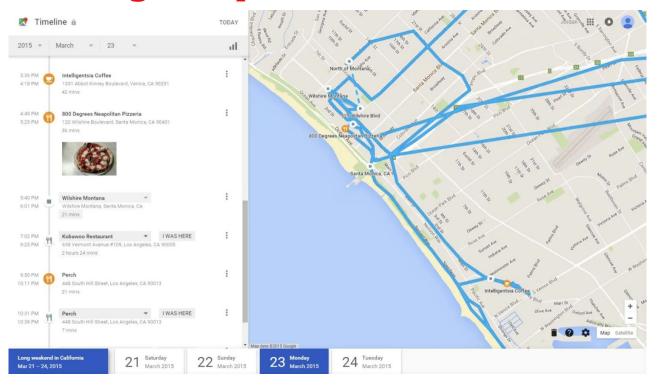
Google, Amazon, Apple, and Microsoft are all central players in "surveillance capitalism" and prey on our data.



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:



Check out a 2015 Venturebeat article about the "freaky" new Google Maps "Your Timeline" feature here.



Northeastern University

NULab for Texts, Maps, and Networks

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.
- <u>Blacklight</u> 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:



We may collect information about the images and audio that are a part of your User Content, such as identifying the objects and scenery that appear, the existence and location within an image of face and body features and attributes, the nature of the audio, and the text of the words spoken in your User Content. We may collect this information to enable special video effects, for content moderation, for demographic classification, for content and ad recommendations, and for other non-personally-identifying operations. We may collect biometric identifiers and biometric information as defined under US laws, such as faceprints and voiceprints, from your User Content. Where required by law, we will seek any required permissions from you prior to any such collection.



Downloading Your Data & Tightening your Privacy

Facebook: Settings > Your Facebook Information > Download your Information

Google: https://support.google.com/accounts/answer/3024190?hl=en

Instagram: Settings > Privacy and Security > Data download/Request Download

TikTok: Profile > 3-line icon (top right) > Settings & Privacy > Privacy > Download your data

Want to make your life more private? Follow this "DIY Guide to Feminist Cybersecurity" https://hackblossom.org/cybersecurity/

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 _____ Authors of All Time"

Now try these searches:

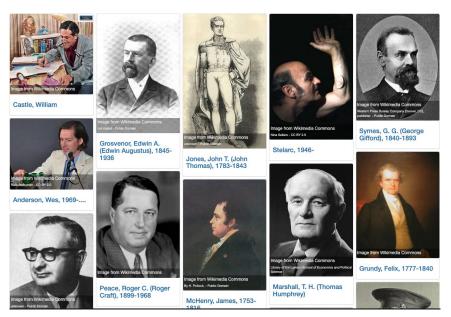
- Greatest women authors
- Greatest Black women authors
- Greatest Black authors
- Greatest white authors

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



Representation in digital archives

- <u>SNAC</u> (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?







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, and most people 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 <u>Vox article</u> for more information on algorithmic bias.



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: 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.
- But 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

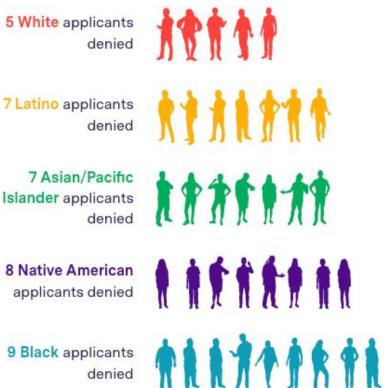


Source: Jenice Kim, The New York Times



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."





Alleviating Injustice

- 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?
- 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?

"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

"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

Als 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

- Als are fundamentally based on statistics. A text-generation
 Al 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.

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 "hallucitations."
 - 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.



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."

AI and Automated Plagiarism Checkers

- 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!

AI and Automated Plagiarism Checkers

- 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!



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.



Moving Forward How can we use data responsibly?

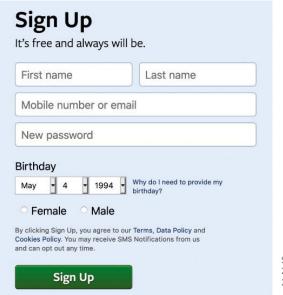
Questions Researchers Must Ask

- 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?



Collecting Data

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



Woman		Non-binary
(including trans	woman)	In another way
Man (including trans man)		Prefer not to say
(including trans	man)	
this the came	e gender you were	assigned at birth?
tills the sain		

account creation nage circa

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



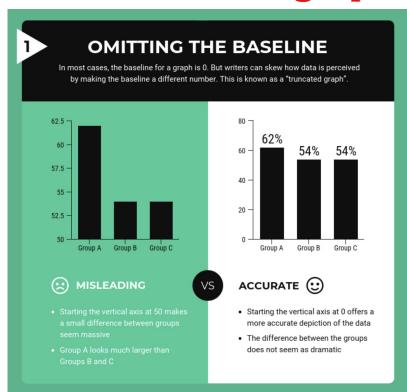
Northeastern University

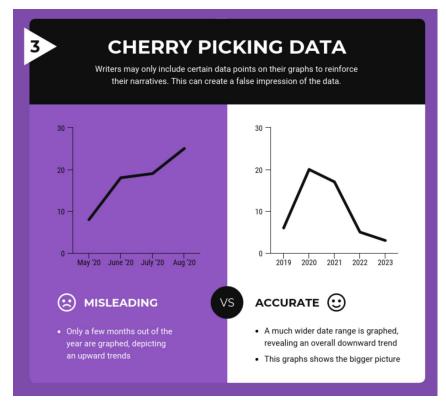
Responsibly Using Big (or any kind of) Data

- 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?



Be Mindful of Infographics and Data Visualizations







Northeastern University

NULab for Texts, Maps, and Networks

Finding and Using Non-Traditional Sources

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

- <u>Public Media</u> (written/broadcast journalism)
- <u>Crowdsourced projects</u> (including Wikipedia, aggregate reviews, etc.)
- <u>Multimedia sources</u> (including social media and blog posts)
 - Using Twitter for academic research
 - o 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.



Thank you!

—**Developed by:** 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: <u>nulab.info@gmail.com</u>

