

# Excavating Awareness and Power for Trustworthy Data Science

---

Katie Shilton

Work done with the PERVADE Project

October 18, 2023

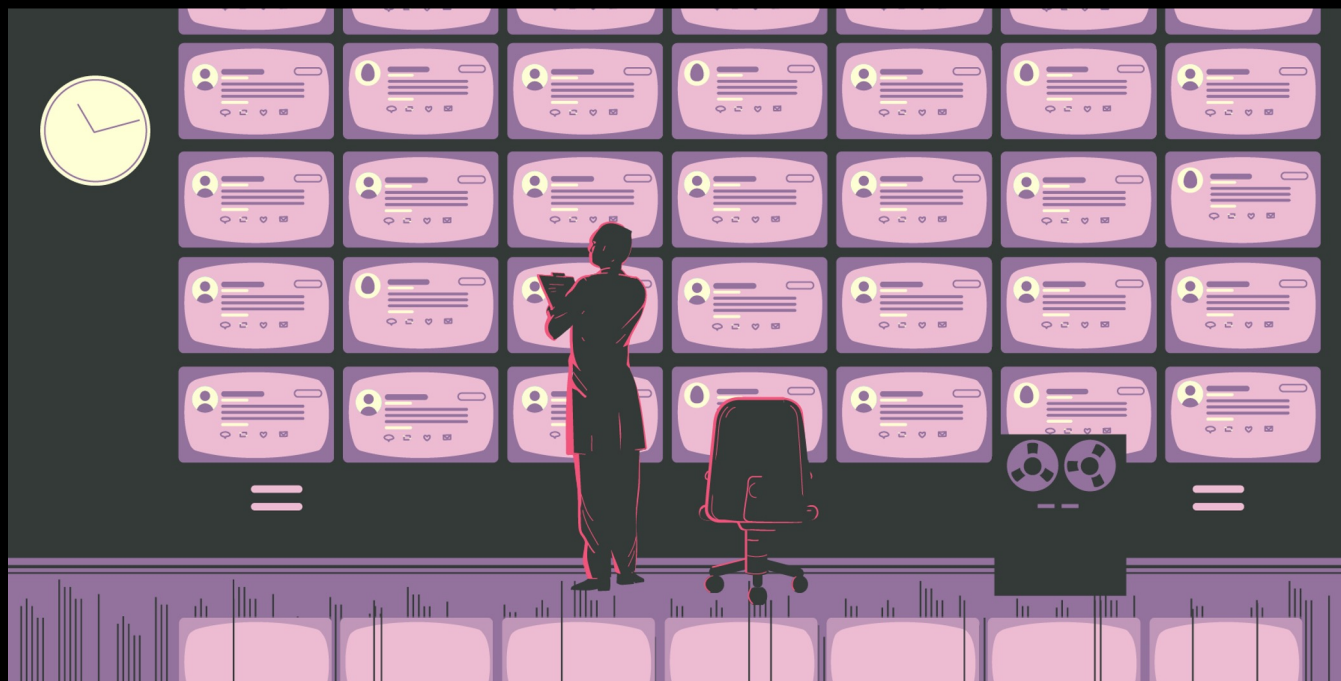


# Outline

1. Data use and research backlash
2. Why Belmont, IRBs & REBs aren't necessarily enough
3. Interventions for trustworthy data use



# A crisis of trust



Fiesler, C., & Proferes, N. (2018). "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society*, 4(1). (Image credit: Darren Garrett for How We Get To Next).

Hemphill, Libby, Angela Schöpke-Gonzalez, and Anmol Panda. 2022. "Comparative Sensitivity of Social Media Data and Their Acceptable Use in Research." *Scientific Data* 9 (1): 643.

# Data use, AI and the “techlash”



# New systems, new challenges

Hopkins Bloomberg  
**PUBLIC HEALTH**

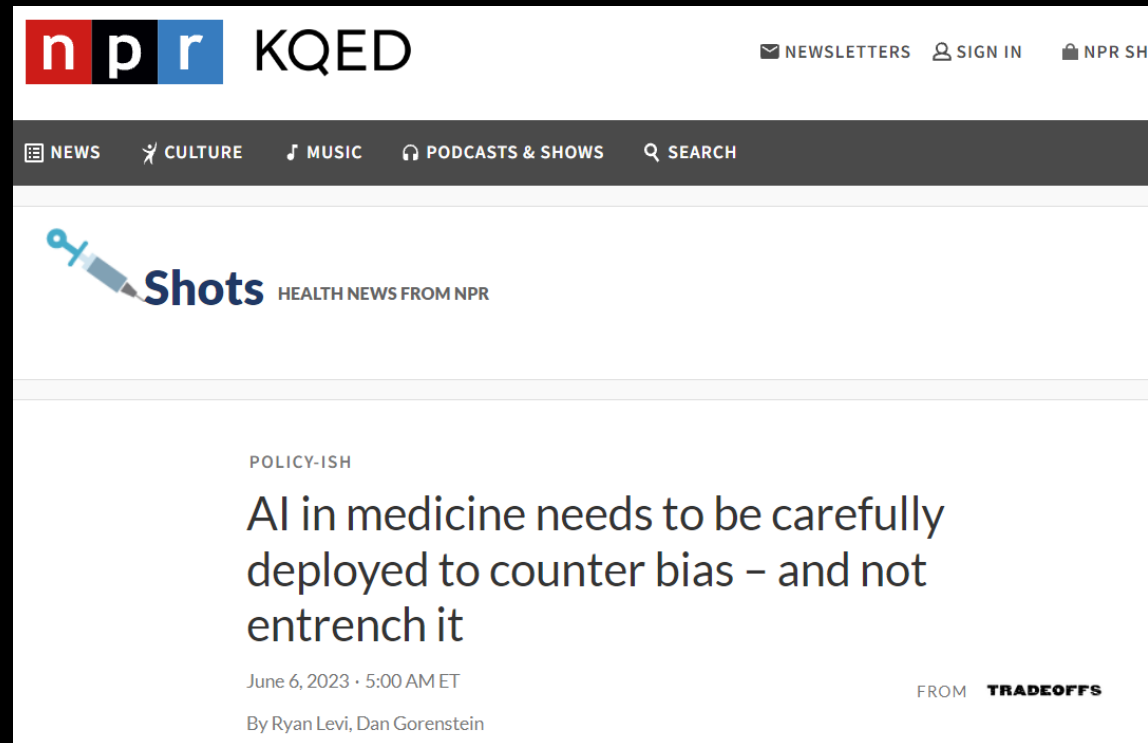
FEATURES

## How Biased Data and Algorithms Can Harm Health

Public health researchers are working to uncover and correct unfairness in AI.

By Carrie Arnold • Photo Illustrations by Patrick Kirchner

<https://magazine.jhsph.edu/2022/how-biased-data-and-algorithms-can-harm-health>



The screenshot shows the NPR KQED website. At the top, there is a navigation bar with the NPR logo (n p r) and the text "KQED". To the right of the logo are links for "NEWSLETTERS", "SIGN IN", and "NPR SH". Below the navigation bar is a secondary menu with links for "NEWS", "CULTURE", "MUSIC", "PODCASTS & SHOWS", and "SEARCH". The main content area features a "Shots" logo with a syringe icon and the text "HEALTH NEWS FROM NPR". Below this, there is a section titled "POLICY-ISH" with the article headline "AI in medicine needs to be carefully deployed to counter bias – and not entrench it". The article is dated "June 6, 2023 · 5:00 AM ET" and is by "Ryan Levi, Dan Gorenstein". At the bottom right of the article, it says "FROM TRADEOFFS".

# Empirical questions in data ethics

PERVADE is an interdisciplinary collaboration between seven researchers at six institutions to answer empirical questions in data ethics.



Funding for this project was provided by NSF Cyber-Human Systems (CHS) Award #1704315

# Empirical + Theoretical Investigations

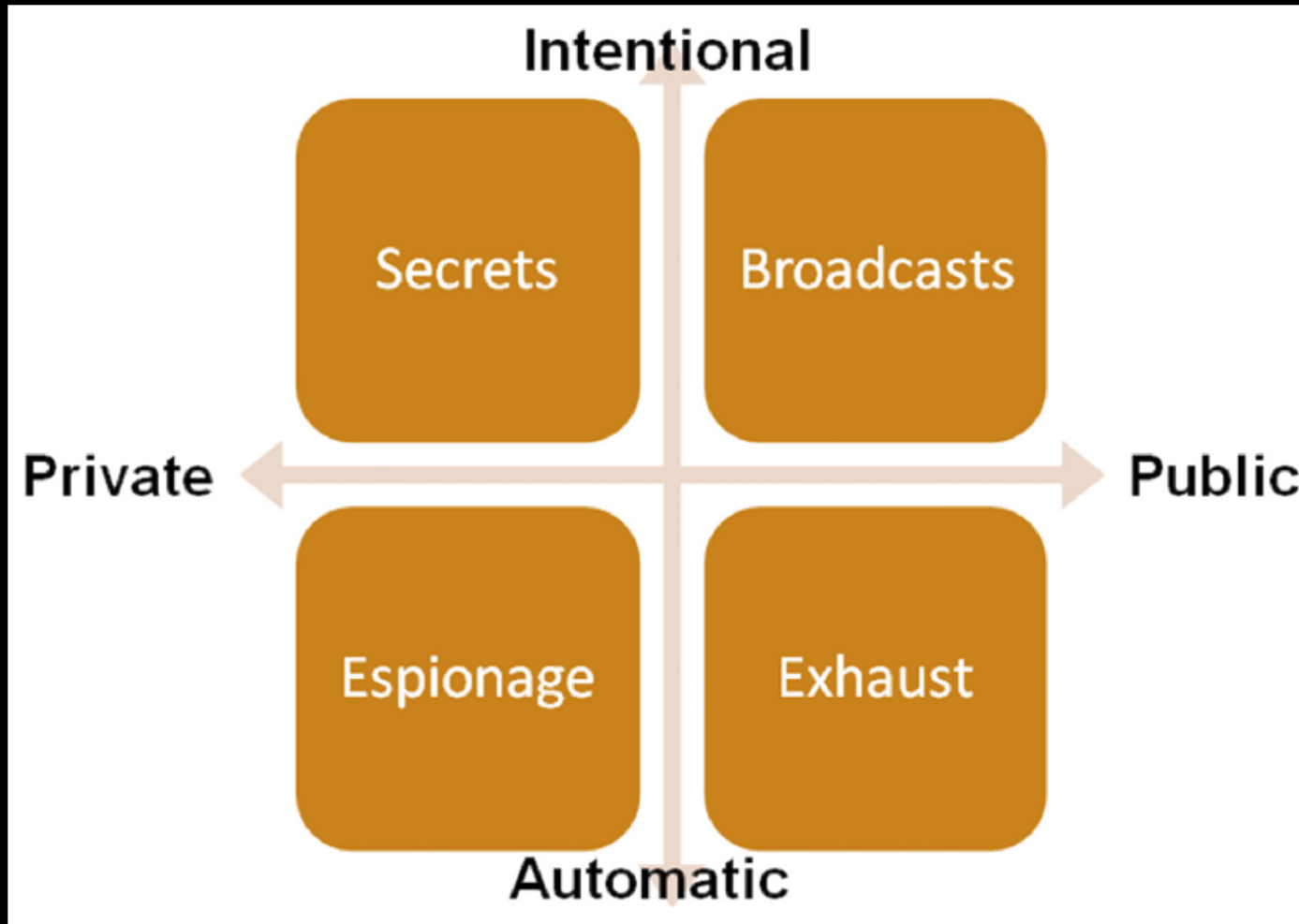
## Empirical studies: Participant awareness and expectations are paramount

- Gilbert, S., Shilton, K., & Vitak, J. (2023). When research is the context: Cross-platform user expectations for social media data reuse. *Big Data & Society*, 10(1), 32 pages. <https://doi.org/10.1177/20539517231164108>
- Gilbert, S., Vitak, J., & Shilton, K. (2021). Measuring Americans' Comfort With Research Uses of Their Social Media Data. *Social Media + Society*, 7(3), 1–13. <https://doi.org/10.1177/20563051211033824>
- Hallinan, B., Brubaker, J., & Fiesler, C. (2019). Unexpected expectations: Public reaction to the Facebook emotional contagion study. *New Media & Society*. <https://doi.org/10.1177/1461444819876944>
- Fiesler, C., & Proferes, N. (2018). "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society*, 4(1), 2056305118763366. <https://doi.org/10.1177/2056305118763366>

## Theoretical work: Power matters

- Klassen, S., & Fiesler, C. (2022). "This Isn't Your Data, Friend": Black Twitter as a Case Study on Research Ethics for Public Data. *Social Media + Society*, 8(4), 20563051221144317. <https://doi.org/10.1177/20563051221144317>
- Shilton, K., Moss, E., Gilbert, S. A., Bietz, M. J., Fiesler, C., Metcalf, J., Vitak, J., & Zimmer, M. (2021). Excavating awareness and power in data science: A manifesto for trustworthy pervasive data research. *Big Data & Society*, 8(2), 1–12. <https://doi.org/10.1177/20539517211040759>
- Dencik, L., Jansen, F., & Metcalfe, P. (2018, August 30). A conceptual framework for approaching social justice in an age of datafication. *DATAJUSTICE Project*. <https://datajusticeproject.net/2018/08/30/a-conceptual-framework-for-approaching-social-justice-in-an-age-of-datafication/>

# Participant Awareness



Shilton, K., Moss, E., Gilbert, S. A., Bietz, M. J., Fiesler, C., Metcalf, J., Vitak, J., & Zimmer, M. (2021). Excavating awareness and power in data science: A manifesto for trustworthy pervasive data research. *Big Data & Society*. <https://doi.org/10.1177/20539517211040759>.



# Reflecting on Power

What kind of consent was given for data?

Is personal information being revealed out of context?

Can individuals or vulnerable groups be identified from your data, analysis, or product?

Can anonymized datasets be de-anonymized?

Does your data come from a particularly vulnerable population?

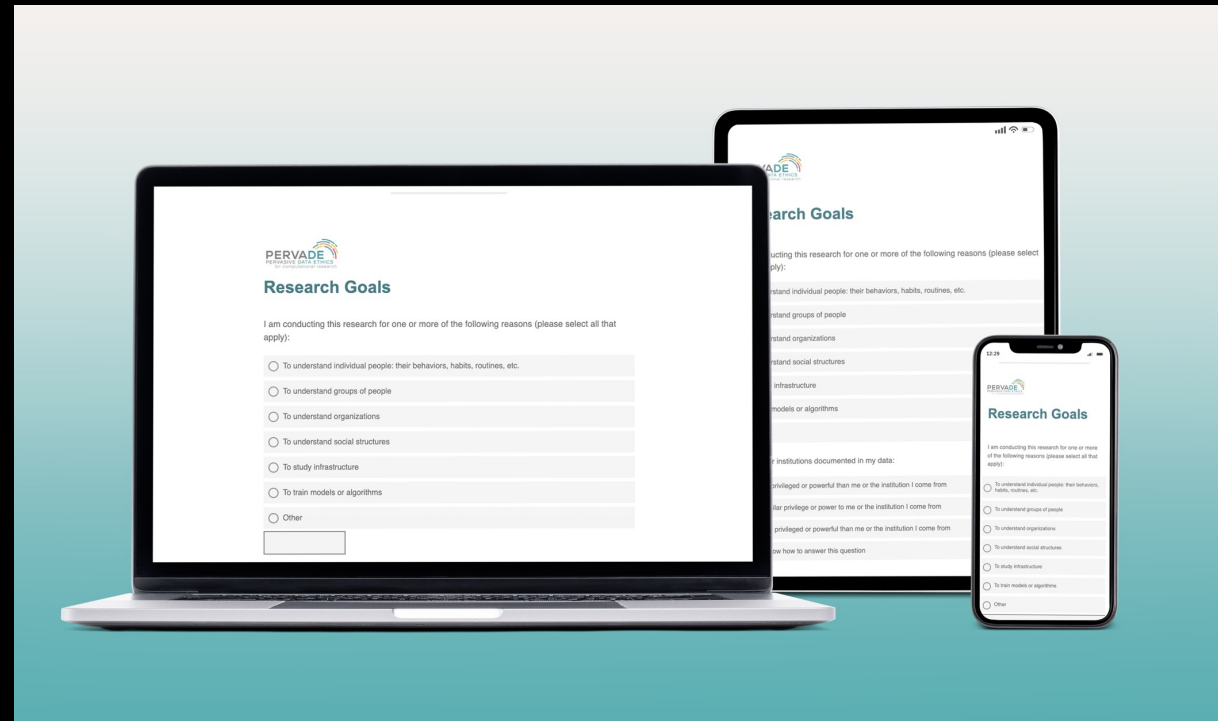
Is it more likely to produce errors when it is applied to a population that has historically been harmed by such errors?

Does it help a less vulnerable population at the expense of a more-vulnerable population?

Does the research design include members of the population group it is targeted at helping??

# Design Intervention: The PERVADE Decision Support Tool

How can we help data users reflect on *awareness* and *power*?



# Project Design for Trust

- Project goals
  - Whose interests are represented?
  - Are there clear IRB or disciplinary guidelines?
- Power and data extraction
  - Studying up and down
  - Representation
- Data sources and contextual expectations
  - Data quadrants
  - Acquisition methods (scraping, APIs, hacks)
  - Norms of the collection context



## Collection

The data that I'm working with for this project (please select all that apply):

- was collected with the informed consent of participants
- was collected without explicit informed consent of participants (e.g. is a reuse of publicly-available data)
- was produced knowingly and intentionally (e.g. Tweets)
- was collected without subjects' knowledge (e.g. location traces)
- is public-facing (e.g. Tweets)
- has an expectation of privacy (e.g. text messages)
- has unclear privacy expectations
- was collected under a terms of service agreement that permits research

### "was produced knowingly and intentionally (e.g. Tweets)"


We call such data broadcast: it was created in public, on purpose. Though this data is generally understood as public, people may still be quite surprised about this use of their data. We recommend you try to increase data subject awareness of your work. Consider using informed consent, or if not possible, other tools of awareness such as sharing results with participants ([link to awareness best practices/resources](#)). For a discussion of challenges using broadcast data, we recommend the video of the PERVADE Conversation with [Dr. Stevie Chancellor](#).

### "is public-facing (e.g. Tweets)"

We call this data exhaust: it was collected in public, but people didn't know they were creating it. We recommend you try to increase data subject awareness of your work, share data and results directly with data subjects, and/or use informed consent. We also recommend the video of the PERVADE Conversation with [Dr. Mark Dredze](#) for a discussion of using exhaust data.

# Research Design for Trust

- Storage and processing
  - Security
  - Retention
- Analysis
  - Binaries and categories
  - Linking datasets
  - Distressing inferences
  - Harms for participants and beyond
- Sharing
  - Deidentification challenges
  - Balancing open science and data protection



## Data Storage and Processing

I am storing my data (please select all that apply):

- In an institutional repository
- On local machines owned by my institution
- In cloud storage provided by my institution
- In cloud storage that I purchased

*"On local machines owned by my institution"*

This is a good start, but you might consider taking advantage of data stewardship offered by an institutional repository. For recommendations on data stewardship, a good place to start is with ICPSR's best practices: [here](#)

When my project is done, we will (please select all that apply):

- Archive the data in an institutional repository
- Share the data via an open repository (GitHub, a website)
- Destroy the data

*"Archive the data in an institutional repository"*

Good choice! For any data you decide not to keep, or that must be destroyed because of sensitivity, a good resource to start with is CDT's guide to data destruction: [click here](#)

[Previous](#) [Next](#)

# Project Prescriptions

## Recommended Resources

These are the data ethics resources that have been recommended to you based on your responses:

- Using pervasive data to train models or algorithms can raise ethical questions about data quality, appropriateness and bias. For discussion of factors to consider, we recommend the video of the PERVADE Conversation with Dr. Amandalynne Paullada. [click here](#)
- That's great. Users of even public platforms still express a preference for giving informed consent before participating in research. For a discussion of consent in data science research, we recommend the video of the PERVADE Conversation with [Dr. Camille Nebeker](#).
- "Social context describes the values and norms of a social setting, and may be the single most important predictor of how people react to research uses of their data. A good resource to start your exploration of social context is Michael Zimmer. 2018. How Contextual Integrity can help us with Research Ethics in Pervasive Data. Medium. Retrieved July 27, 2018 from [here](#).

**Have additional questions or maybe have a resource you would like to share with us?**

Contact the PERVADE team at [pervade@school@umd.edu](mailto:pervade@school@umd.edu)

[Visit the PERVADE Website](#)

[View this email in your browser](#)



## Thank you for using the PERVADE Data Ethics Tool!

Below you can find your results from your most recent completion of the data ethics evaluation tool.

## Your Diagnostic Score

**50**

**Hard Mode: 39 - 60**

To do this project ethically will take considerable time and effort. The project studies individuals; it may use data we would label as "espionage"; the analysis involves tricky binaries and potentially invasive conclusions; you haven't yet explored contextual norms for analysis and sharing; you have plans to combine datasets or share your data; and you're not yet sure of your discipline or profession's ethical norms. These answers indicate that you have a very difficult task in front of you. We recommend some homework to help you get started on this work. 1) Follow up with the resources and guidelines provided throughout the quiz. 2) Try some design fiction. Create a dystopian or worst case scenario for your research. What could go wrong? Who could use your data or results for evil? Then work backwards to try to mitigate some of those potential hazards. 3) Run a contextual integrity audit for your project. [See here to get started.](#)

## Recommended Resources

These are the data ethics resources that have been recommended to you based on your responses:

# Going Forward: Increasing Participation in AI



## Participatory Design

We will create new AI technology and advocate for its deployment and use in a way that aligns with the values and interests of diverse groups of people, particularly those that have been previously marginalized in the development process.

Lead: Katie Shilton, an associate professor in UMD's College of Information Studies who specializes in ethics and sociotechnical systems, particularly as it relates to big data.



## Methods and Metrics

We will develop novel methods, metrics and advanced machine learning algorithms that reflect the values and interests of relevant stakeholders, allowing them to understand the behavior and best uses of AI-infused systems.

Lead: Tom Goldstein, a UMD associate professor of computer science who leverages mathematical foundations and efficient hardware for high-performance systems.



## Evaluating Trust

We will effectively evaluate how people make sense of AI systems, and the degree to which their levels of reliability, fairness, transparency and accountability can lead to appropriate levels of trust.

Lead: David Broniatowski, a GW associate professor of engineering whose research focuses on decision-making under risk, group decision-making, system architecture, and behavioral epidemiology.



## Participatory Governance

We will explore how policymakers at all levels in the U.S. and abroad can foster trust in AI systems, as well as how policymakers can incentivize broader participation, accountability, and inclusiveness in the design and deployment of these systems.

Lead: Susan Ariel Aaronson, a research professor of international affairs at GW who is an expert in data-driven change and international data governance.

# Thank you!

[kshilton@umd.edu](mailto:kshilton@umd.edu)

TRAILS

<https://www.trails.umd.edu/>

PERVADE Project

<http://pervade.umd.edu>

Ethics & Values in Design (EViD) Lab

<http://evidlab.umd.edu/>



*This material is based on work supported by National Science Foundation under awards 1704369 and 2229885.*