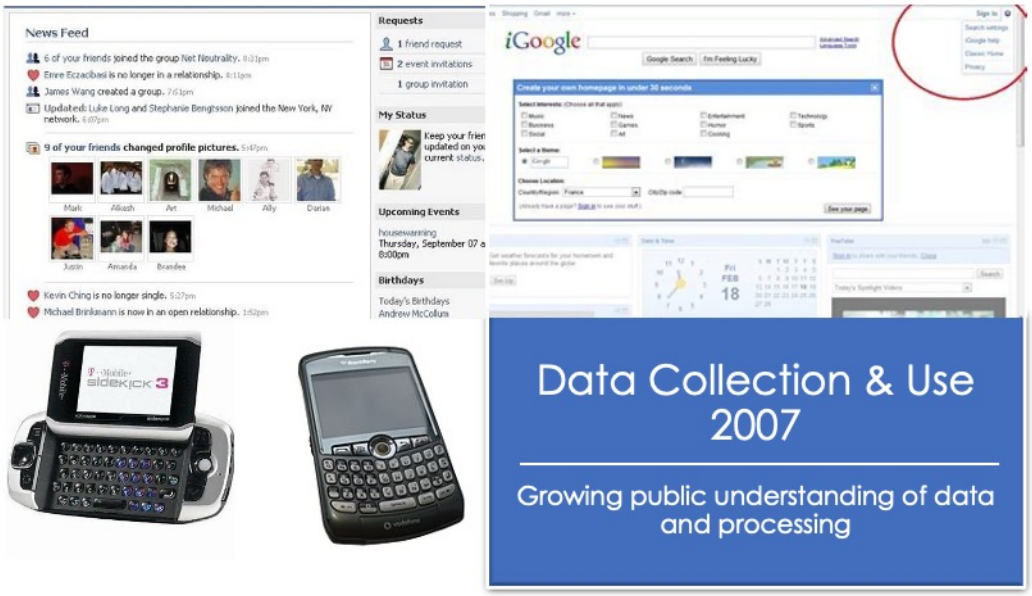
RESPONSIBLE DATA USE IN THE AGE OF MACHINE LEARNING

2/25/2022 UW CSE 163

EMILY MCREYNOLDS

@EMILYNEXUS

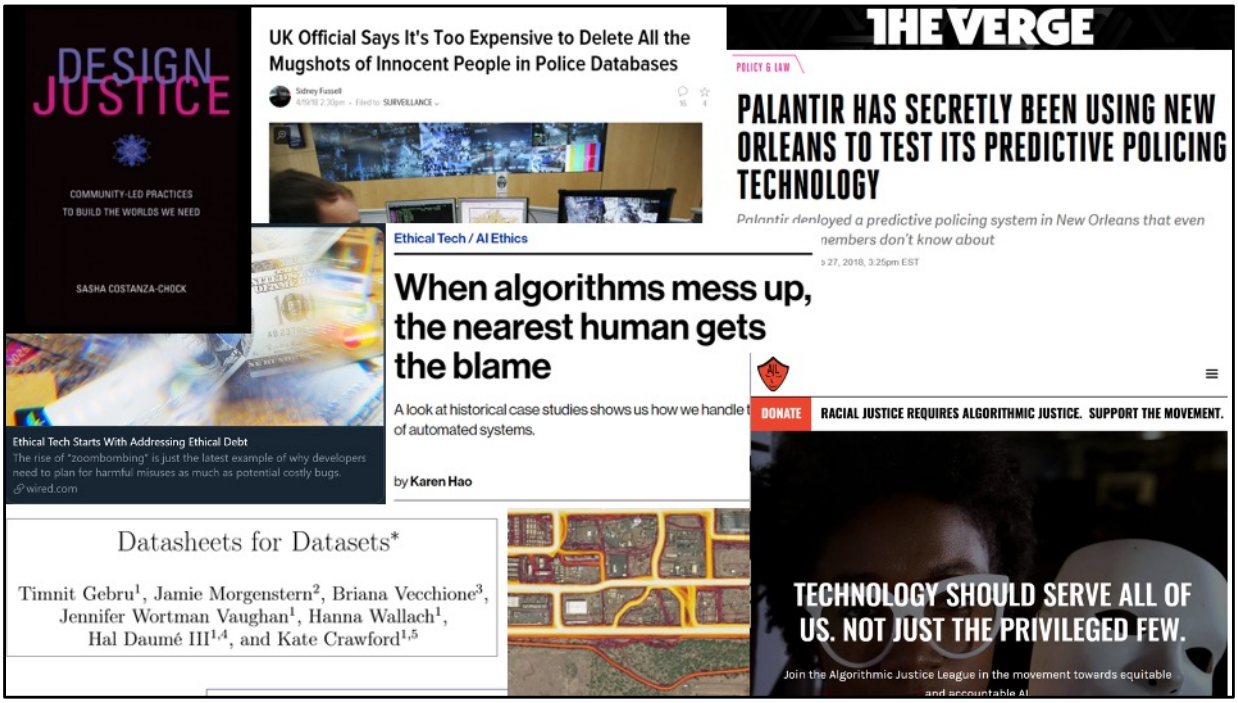


Data Collection & Use 2007

Growing public understanding of data and processing

Data Collection & Processing in 2022

Smartphones and ambient data collection



Data justice
Data ethics

Examples:
Strava's heat map revealed military bases around the world and exposed soldiers to real danger – https://www.nytimes.com/2018/01/29/world/middleeast/strava-heat-map.html
Datasheets for Datasets - https://arxiv.org/abs/1803.09010
Ethical Debt https://www.wired.com/story/opinion-ethical-tech-starts-with-addressing-ethical-debt/
Palantir - https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd
https://gizmodo.com/uk-official-says-its-too-expensive-to-delete-all-the-mu-1825388626
MIT https://www.technologyreview.com/2019/05/28/65748/ai-algorithms-liability-human-blame/
Algorithmic Justice League https://www.ajl.org/
Design Justice https://design-justice.pubpub.org/

https://datajusticelab.org/

TINY IMAGES – BIG PROBLEM

June 29th, 2020

It has been brought to our attention [1] that the Tiny Images dataset contains some derogatory terms as categories and offensive images. This was a consequence of the automated data collection procedure that relied on nouns from WordNet. We are greatly concerned by this and apologize to those who may have been affected.

The dataset is too large (80 million images) and the images are so small (32 x 32 pixels) that it can be difficult for people to visually recognize its content. Therefore, manual inspection, even if feasible, will not guarantee that offensive images can be completely removed.

We therefore have decided to formally withdraw the dataset. It has been taken offline and it will not be put back online. We ask the community to refrain from using it in future and also delete any existing copies of the dataset that may have been downloaded.

How it was constructed: The dataset was created in 2006 and contains 53,464 different nouns, directly copied from Wordnet. Those terms were then used to automatically download images of the corresponding noun from Internet search engines at the time (using the available filters at the time) to collect the 80 million images (at tiny 32x32 resolution; the original high-res versions were never stored).

Why it is important to withdraw the dataset: biases, offensive and prejudicial images, and derogatory terminology alienates an important part of our community -- precisely those that we are making efforts to include. It also contributes to harmful biases in AI systems trained on such data. Additionally, the presence of such prejudicial images hurts efforts to foster a culture of inclusivity in the computer vision community. This is extremely unfortunate and runs counter to the values that we strive to uphold.

Yours Sincerely,

Antonio Torralba, Rob Fergus, Bill Freeman.

[1] Large image datasets: A pyrrhic win for computer vision?, anonymous authors, OpenReview Preprint, 2020.

LARGE DATASETS: A PYRRHIC WIN FOR COMPUTER VISION?

Vinay Uday Prabhu*
UnifyID AI Labs
Redwood City
vinay@unify.id

Abeba Birhane*
School of Computer Science, UCD, Ireland
Lero - The Irish Software Research Centre
abeba.birhane@ucdconnect.ie

July 27, 2020

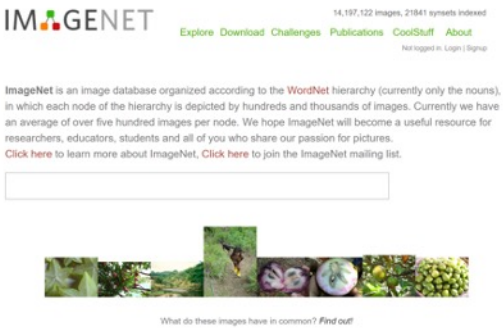
ABSTRACT

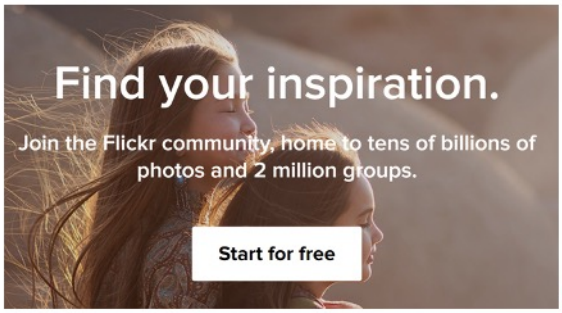
In this paper we investigate problematic practices and consequences of large scale vision datasets. We examine broad issues such as the question of consent and justice as well as specific concerns such as the inclusion of verifiably pornographic images in datasets. Taking the ImageNet-ILSVRC-2012 dataset as an example, we perform a cross-sectional model-based quantitative census covering factors such as age, gender, NSFW content scoring, class-wise accuracy, human-cardinality-analysis, and the semanticity of the image class information in order to statistically investigate the extent and subtleties of ethical transgressions. We then use the census to help hand-curate a look-up-table of images in the ImageNet-ILSVRC-2012 dataset that fall into the categories of verifiably pornographic: shot in a non-consensual setting (up-skirt), beach voyeuristic, and exposed private parts. We survey the landscape of harm and threats both society broadly and individuals face due to uncritical and ill-considered dataset curation practices. We then propose possible courses of correction and critique the pros and cons of these. We have duly open-sourced all of the code and the census meta-datasets generated in this endeavor for the computer vision community to build on. By unveiling the severity of the threats, our hope is to motivate the constitution of mandatory Institutional Review Boards (IRB) for large scale dataset curation processes.

Long used benchmark datasets can turn into challenging datasets.

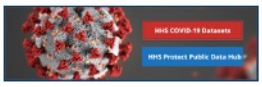
80 Million Tiny Images paper - https://cs.nyu.edu/~fergus/papers/tiny.pdf
Letter taking down dataset - https://groups.csail.mit.edu/vision/TinyImages/
Paper: “Large datasets: A Pyrrhic win for computer vision?”
https://arxiv.org/abs/2006.16923

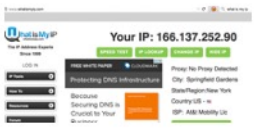
CHANGING CONSIDERATIONS





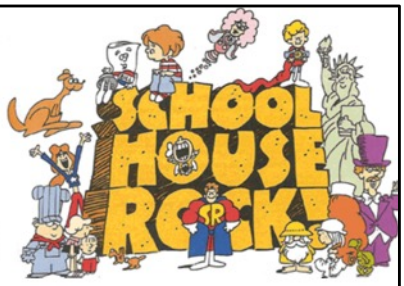
Spectrum of data



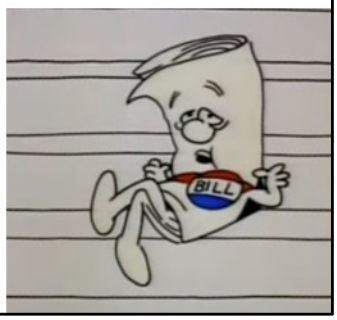


Not people data | De-identified but sensitive | Identifiable data

DATA RELATED LEGISLATION



- Federal Education Rights Privacy Act (FERPA)
- Health Information Portability & Accountability Act (HIPAA)
- Gramm-Leach-Bliley Act (GLBA) – Article V
- Children's Online Privacy Protection Act (COPPA)
- Federal Trade Commission Act (Article 5)



ALWAYS ASK YOURSELF THE HARD QUESTIONS

- Should we build it?
- Who are the stakeholders?
- Who will this help and how?
- Who might this hurt and how?
- Never forget: "Essentially, all models are wrong, but some are useful." George Box (sort of)

ETHICS IN DATA SCIENCE

CASE STUDY 0: PREDICTING CRIMINALITY

CASE STUDY 1: TRACKING FOR SAFETY

CASE STUDY 2: COMPAS

CASE STUDY 3: POTHOLES IN BALTIMORE

CASE STUDY 4: FACIAL RECOGNITION

PHRENOLOGY

- What is phrenology?
- Why do you think it comes up when papers/articles claim to predict criminology, sexuality, and other traits through pictures?


AMERICAN PHRENOLOGICAL JOURNAL.

KNOW THYSELF.

HOME TRUTHS FOR HOME CONSUMPTION.

1848.

VOL. X. MARCH. NO. 3.

O. S. FOWLER, EDITOR.

PHRENOLOGY, PHYSIOLOGY, PHYSIOGNOMY, MAGNETISM.

NEW YORK:

FOWLERS & WELLS,

PHRENOLOGICAL CABINET, 131 NASSAU-STREET.

Terms $1 a Year, invariably in advance. Twenty a Number.

FACIAL TRACKING

- Can this be done responsibly?
- Relevant research – Prof. Barrett
- Microsoft Research: Guidelines for Assessing and Minimizing Risks of Emotion Recognition Applications

https://cos.northeastern.edu/people/lisa-barrett
https://journals.sagepub.com/eprint/SAUES8UM69EN8TSMUGF9/full
https://news.northeastern.edu/2021/08/20/you-cant-determine-emotion-from-someones-facial-expression-and-neither-can-ai/
https://www.microsoft.com/en-us/research/publication/guidelines-for-assessing-and-minimizing-risks-of-emotion-recognition-applications/
https://www.microsoft.com/en-us/research/uploads/prod/2021/07/camera_ready_share.pdf

ETHICS IN DATA SCIENCE

CASE STUDY 0: PREDICTING CRIMINALITY

CASE STUDY 1: TRACKING FOR SAFETY

CASE STUDY 2: COMPAS

CASE STUDY 3: POTHOLES IN BALTIMORE

CASE STUDY 4: FACIAL RECOGNITION


9:41 AM
100%
Mail
Calendar
Photos
Camera
Maps
Clock
Weather
News
Home
Notes
Stocks
Reminders
Health
Wallet
Settings
07:00
NOKIA
10:10
motorola
11:35

Google

Privacy-Preserving Contact Tracing

Across the world, governments, and health authorities are working together to find solutions to the COVID-19 pandemic, to protect people and get society back up and running. Software developers are contributing by crafting technical tools to help combat the virus and save lives. In this spirit of collaboration, Google and Apple are announcing a joint effort to enable the use of Bluetooth technology to help governments and health agencies reduce the spread of the virus, with user privacy and security central to the design.

As part of this partnership Google and Apple are releasing draft documentation for an Exposure Notification system in service of privacy-preserving contact tracing:

Exposure Notification - Bluetooth Specification

Exposure Notification - Cryptography Specification

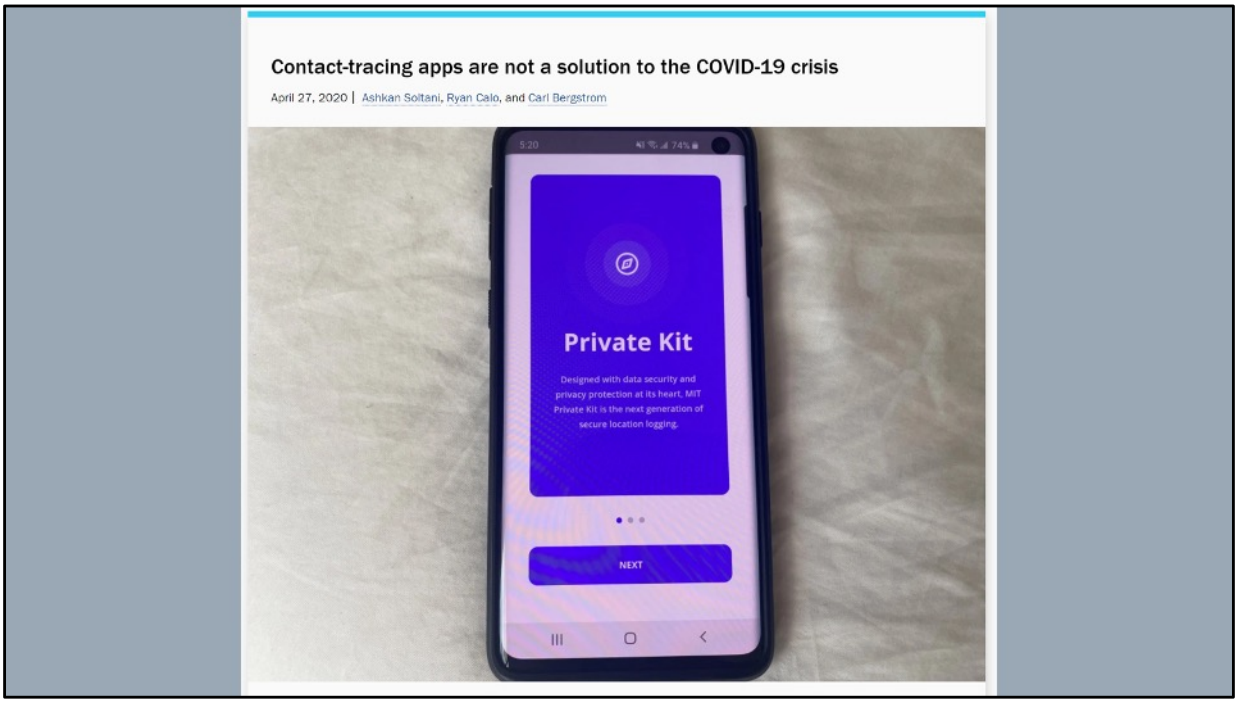
Exposure Notification - Framework API

Exposure Notification - Frequently Asked Questions

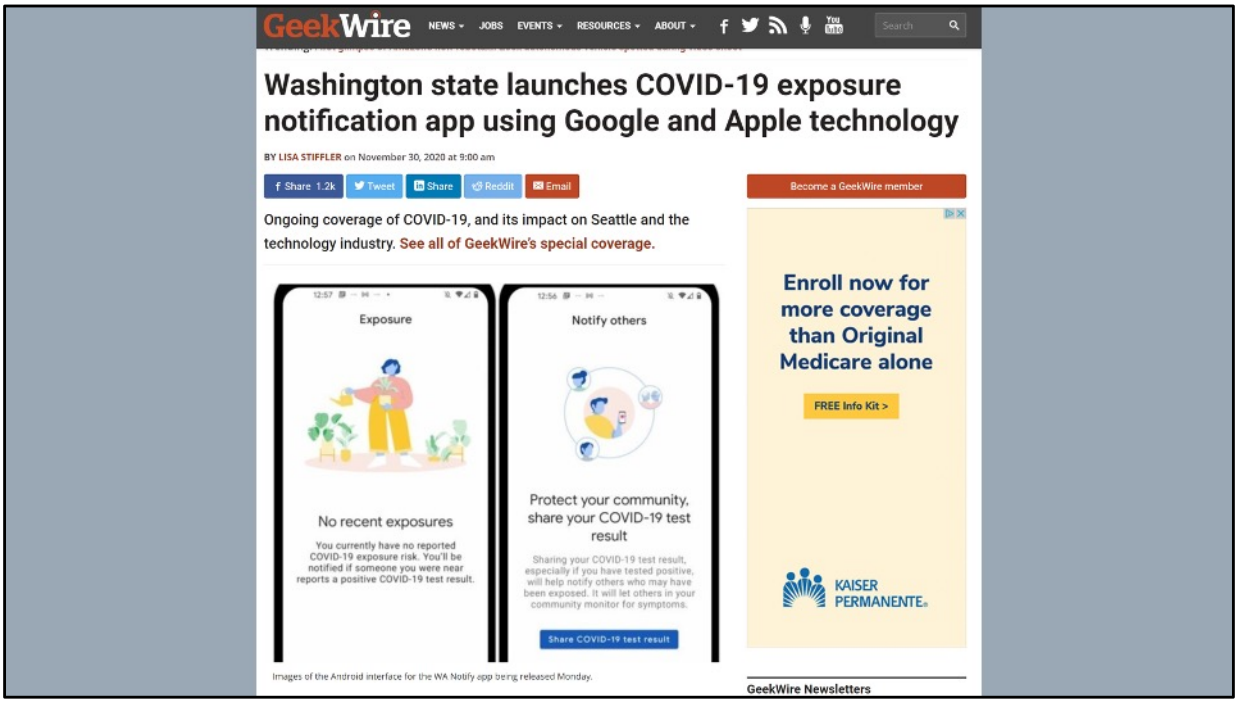
All of us at Apple and Google believe there has never been a more important moment to work together to solve one of the world's most pressing problems. Through close cooperation and collaboration with developers, governments, and public health providers, we hope to harness the power of technology to help countries around the world slow the spread of COVID-19 and accelerate the return of everyday life.

https://covid19.apple.com/contacttracing
https://www.wired.com/story/apple-google-contact-tracing-strengths-weaknesses/
https://www.technologyreview.com/2020/11/20/1012325/do-digital-contact-tracing-apps-work-heres-what-you-need-to-know/

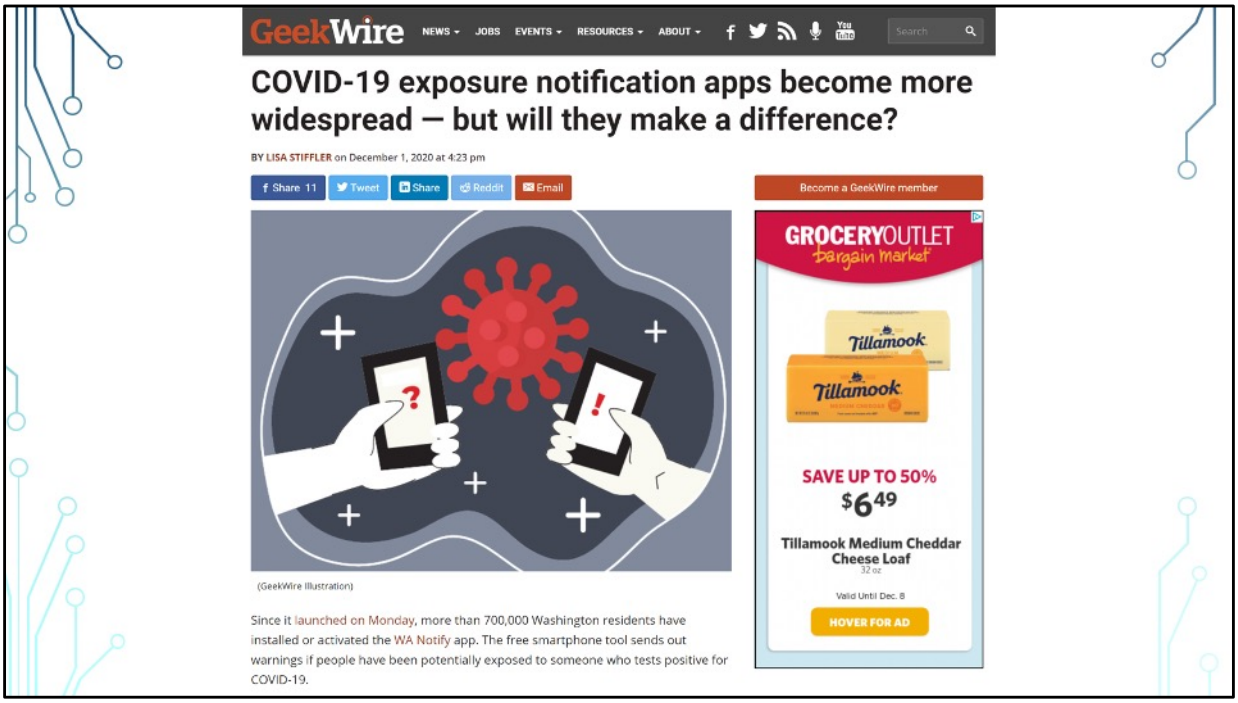
https://twitter.com/doctorow/status/1296248163653660672
https://threader.app/thread/1296248163653660672



https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/



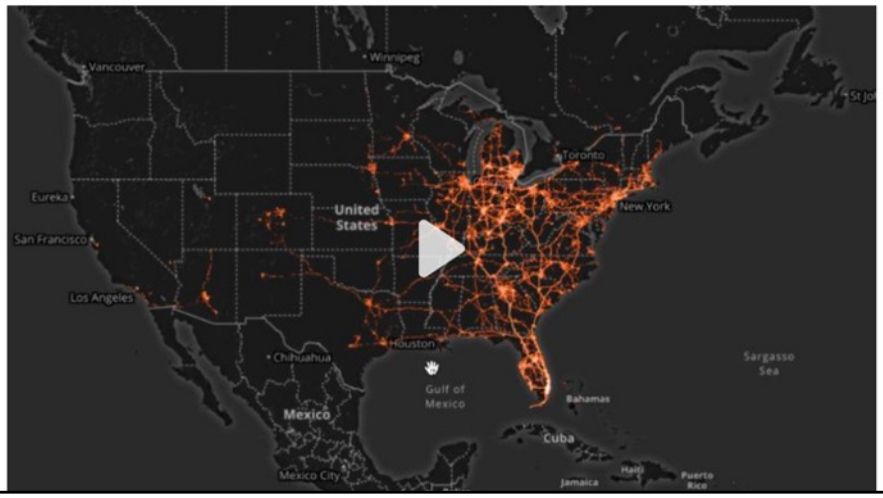
https://www.geekwire.com/2020/wa-notify-covid-exposure-app/



https://www.geekwire.com/2020/covid-19-exposure-notification-apps-become-widespread-will-make-difference/

How the cell phones of spring breakers who flouted coronavirus warnings were tracked

By Donie O'Sullivan
Video by Sofia Barrett
Updated 8:50 AM ET, Sat April 4, 2020



ETHICS IN DATA SCIENCE

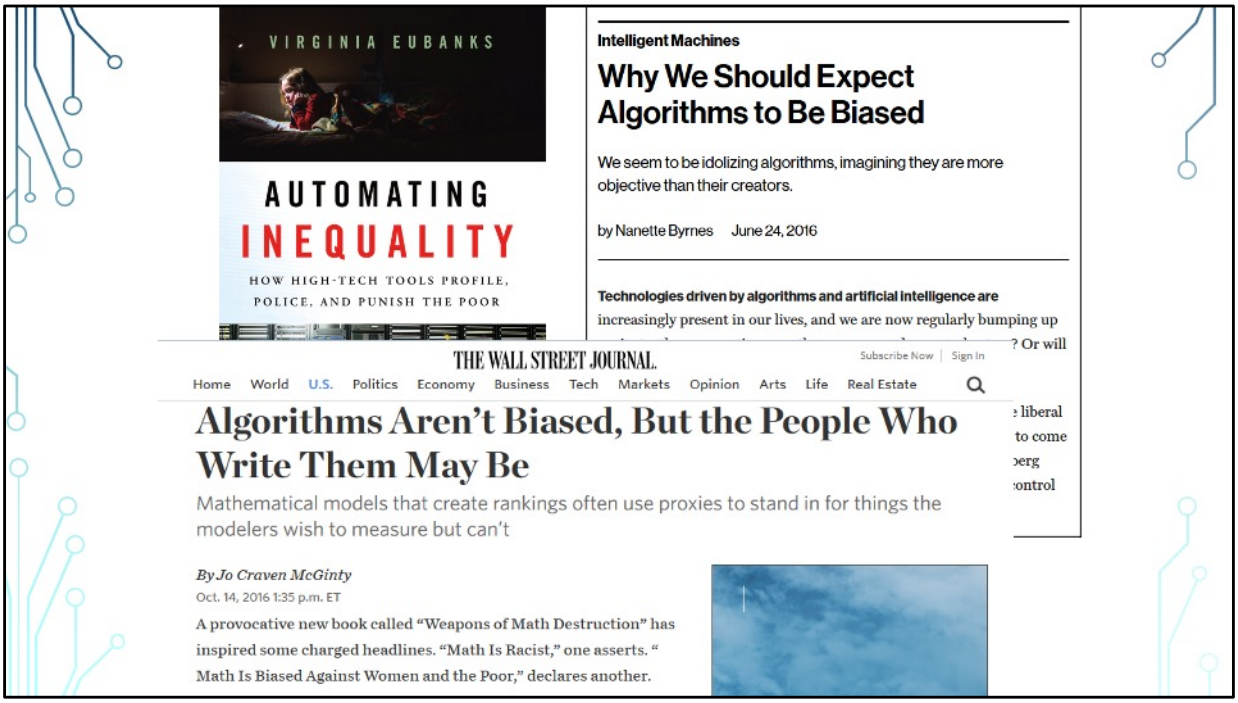
CASE STUDY 0: PREDICTING CRIMINALITY

CASE STUDY 1: TRACKING FOR SAFETY

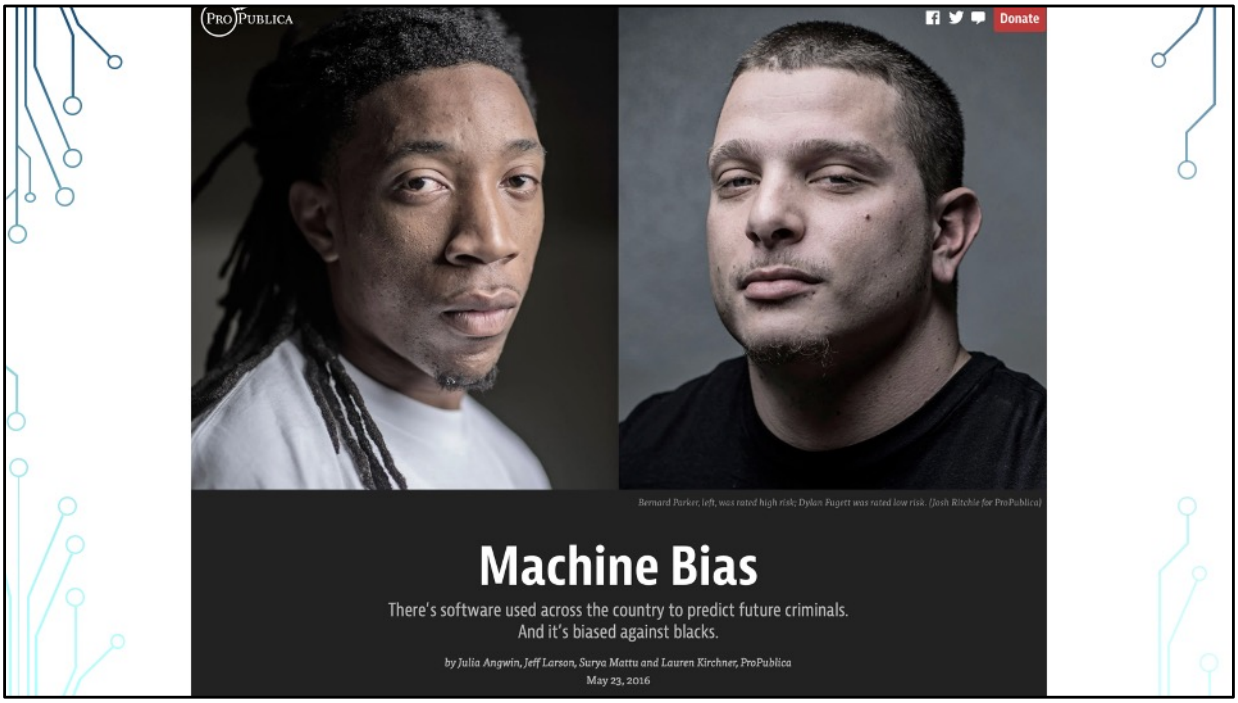
CASE STUDY 2: COMPAS

CASE STUDY 3: POTHOLES IN BALTIMORE

CASE STUDY 4: FACIAL RECOGNITION



https://www.technologyreview.com/2016/06/24/159118/why-we-should-expect-algorithms-to-be-biased/
https://www.wsj.com/articles/algorithms-arent-biased-but-the-people-who-write-them-may-be-1476466555
https://virginia-eubanks.com/books/



https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

NORTHEPOINTE'S COMPAS

Correctional Offender Management Profiling for Alternative Sanctions

- What do you think it was design to do?
- Does the data it's using make sense for COMPAS' purpose?

ETHICS IN DATA SCIENCE

CASE STUDY 0: PREDICTING CRIMINALITY

CASE STUDY 1: TRACKING FOR SAFETY

CASE STUDY 2: COMPAS

CASE STUDY 3: POTHOLES IN BALTIMORE

CASE STUDY 4: FACIAL RECOGNITION

ACCESS TO TECHNOLOGY

- Do potholes as an ethics issue seem innocuous?
- What could Baltimore have done differently?

ETHICS IN DATA SCIENCE

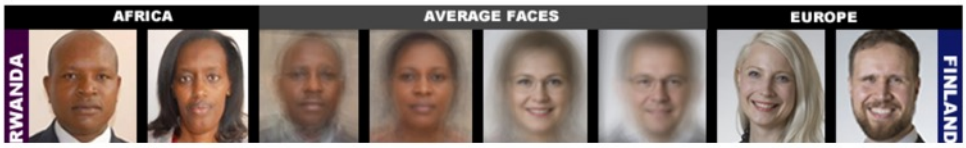
CASE STUDY 0: PREDICTING CRIMINALITY

CASE STUDY 1: TRACKING FOR SAFETY

CASE STUDY 2: COMPAS

CASE STUDY 3: POTHOLES IN BALTIMORE

CASE STUDY 4: FACIAL RECOGNITION



Proceedings of Machine Learning Research 81:1–15, 2018 Conference on Fairness, Accountability, and Transparency

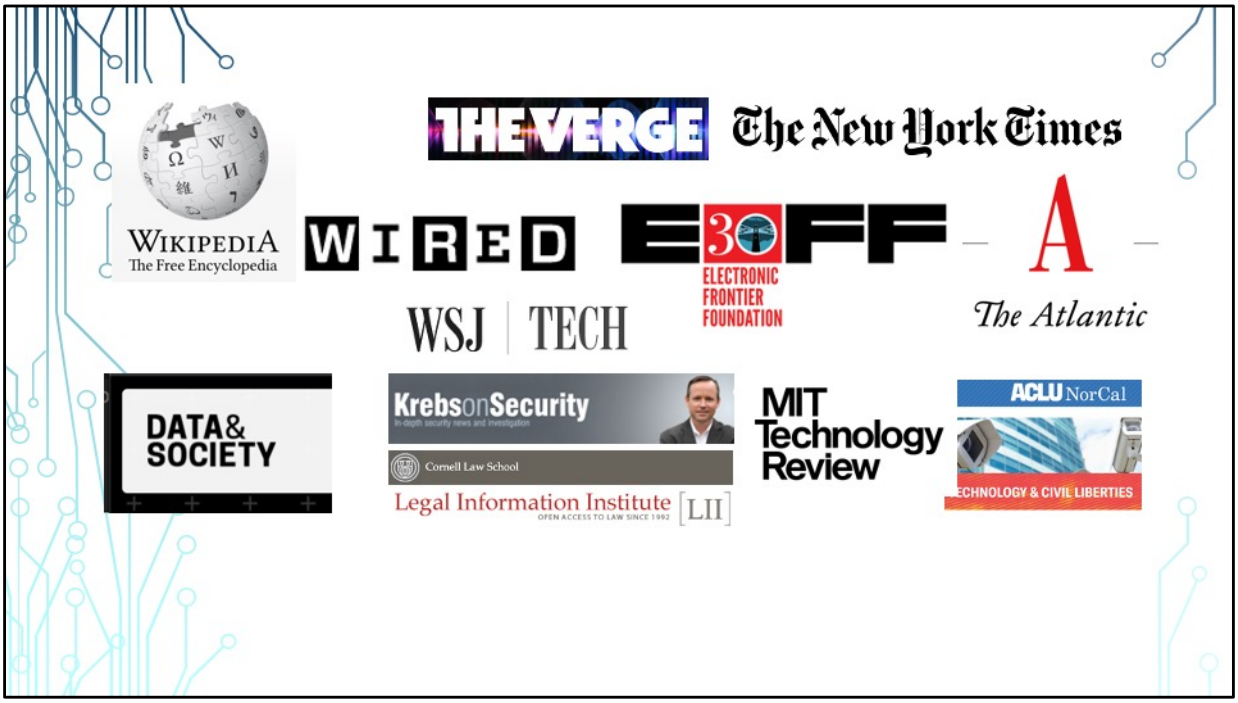
Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*

Joy Buolamwini JOYAB@MIT.EDU
MIT Media Lab 75 Amherst St. Cambridge, MA 02139

Timnit Gebru TIMNIT.GEBRU@MICROSOFT.COM
Microsoft Research 641 Avenue of the Americas, New York, NY 10011

Editors: Sorelle A. Friedler and Christo Wilson

https://partnershiponai.org/about-ml-2021/

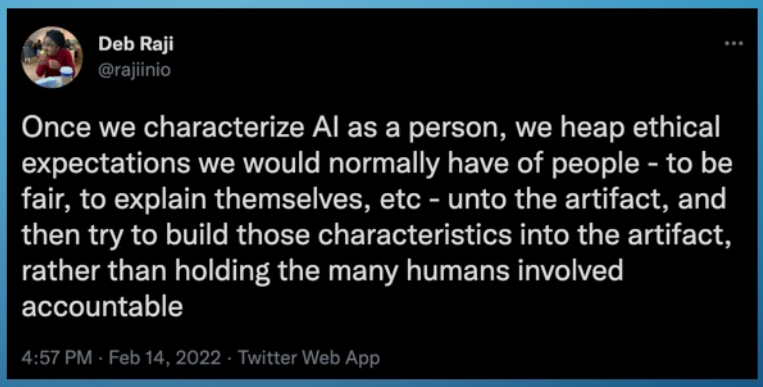


New York Times – Kash Hill
Https://www.eff.org
https://www.theatlantic.com
https://www.wsj.com
https://www.aclunc.org/issue/technology-civil-liberties
https://datasociety.net/
https://www.wired.com/
https://krebsonsecurity.com/
https://www.law.cornell.edu/uscode/text
https://www.technologyreview.com/

TAKEAWAYS


Matthew Hutson
@Silverjacket
Follow
Fun at #AIES2018. Speaker: "Here's our algorithm." Audience member: "Have you considered the ethical implications?" S: "I'm just an engineer." AM: "'Once the rockets are up, who cares where they come down?'" @RealAAAI @TheOfficialACM
CRIME SCENE DO NOT CROSS
Artificial intelligence could identify gang crimes—and ignite an ethical fir...
Law enforcement algorithm sparks heated debates
sciencemag.org
9:01 AM - 1 Mar 2018
8 Retweets 11 Likes
8
11

TAKEAWAYS



https://twitter.com/rajiinio/status/1493388989918629890?s=20&t=Ux2u3pRYsttVDmybPiD9vw

Responsibility in AI Systems & Experiences (RAISE) at the University of Washington presents:

Casey Fiesler

Data Is People: Unintended Consequences in AI and Data Science

Friday February 25, 2022, 9-10am PT
Join: https://washington.zoom.us/j/94636255672

Abstract: Hardly a day passes without a new technology ethics scandal in the news, and many of them touch on AI, machine learning, and data science—from privacy violations in data collection to biased data feeding algorithms to unexpected uses of research. This talk builds on two threads of my research: (1) empirical work towards best practices for research ethics in data science, including understanding possible harms of using public data without consent; and (2) ethical debt (as a parallel to technical debt) in technology design and research as the precursor to the types of unintended consequences that underly many ethical controversies. I will suggest strategies for being more forward-thinking and for remembering the people present in data.