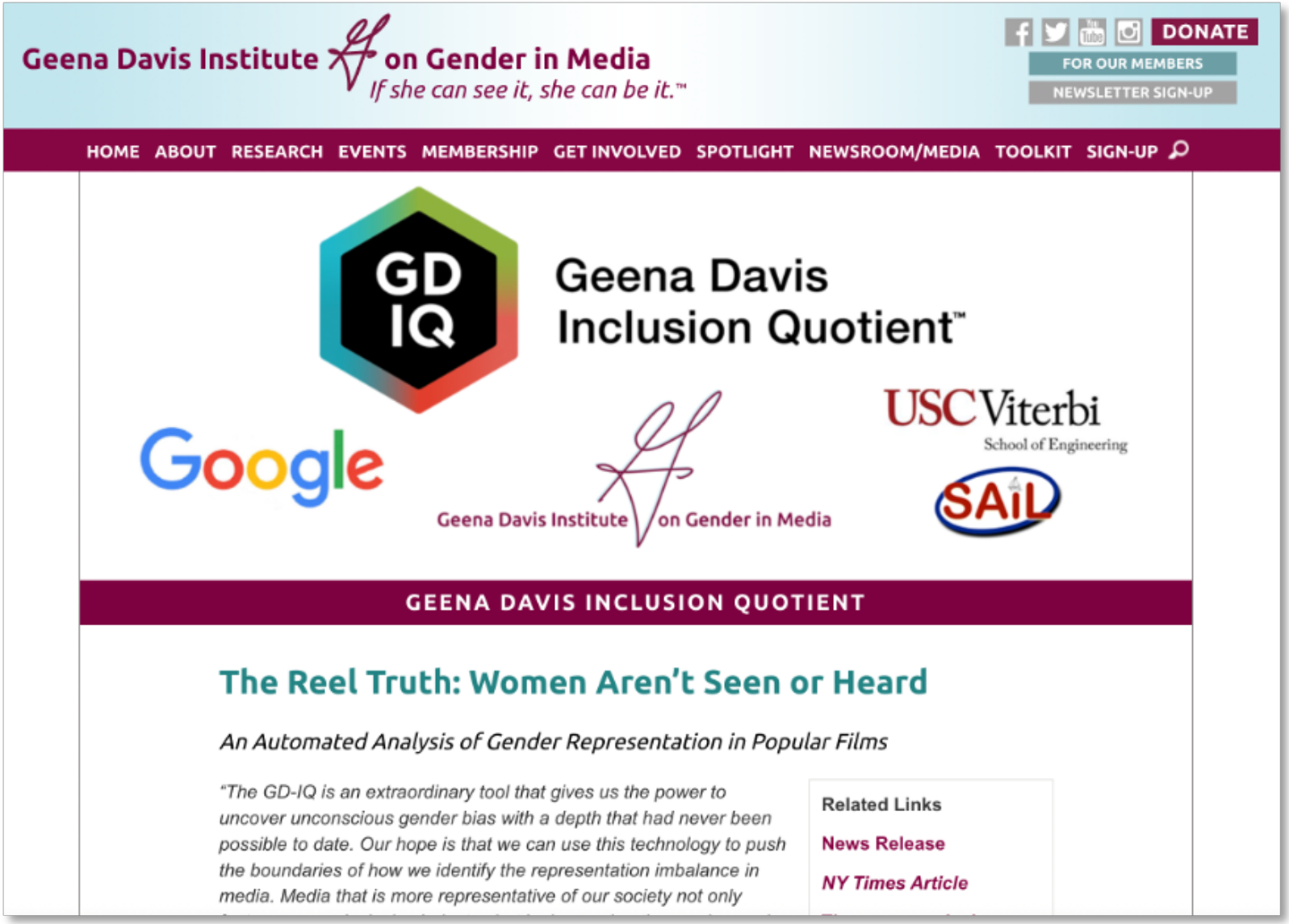
Teaching "Algorithms and Social Justice"

Larry Snyder and Suzanne Edwards

Lehigh University
Bethlehem, PA

Penn State–Berks, 11/10/23

The GD-IQ


Geena Davis Institute on Gender in Media
If she can see it, she can be it.™
DONATE
FOR OUR MEMBERS
NEWSLETTER SIGN-UP
HOME ABOUT RESEARCH EVENTS MEMBERSHIP GET INVOLVED SPOTLIGHT NEWSROOM/MEDIA TOOLKIT SIGN-UP
GD IQ
Geena Davis Inclusion Quotient™
Google
Geena Davis Institute on Gender in Media
USC Viterbi School of Engineering
SAIL
GEENA DAVIS INCLUSION QUOTIENT
The Reel Truth: Women Aren't Seen or Heard
An Automated Analysis of Gender Representation in Popular Films
"The GD-IQ is an extraordinary tool that gives us the power to uncover unconscious gender bias with a depth that had never been possible to date. Our hope is that we can use this technology to push the boundaries of how we identify the representation imbalance in media. Media that is more representative of our society not only
Related Links
News Release
NY Times Article

ISE/WGSS 296: Algorithms and Social Justice

- Cross-listed between Industrial and Systems Engineering (ISE) and Women, Gender, and Sexuality Studies (WGSS)
  - Larry: ISE
  - Suzanne: WGSS and English
- Taught Fall 2022 (11 students), Fall 2023 (18 students)
- Roughly equal split between engineering students and arts and sciences students
- No prerequisites

Pedagogical Approach

- Humanities and engineering approaches from the ground up
- We will read feminist theory one day and write Python code the next
- Everyone is pushed outside their comfort zone
- Model how to be confused, how to disagree, how to embrace the "pleasures of the difficult"
- Peer-to-peer learning

Avoiding "Mix-and-Stir"

The disciplines [...] come in separate bottles, with separate production histories before they are mixed. It would be better to meld their production processes together. Engineering and history, computing and feminist politics, platform capitalism and gig workers' unions should be in engagement and contestation from the beginning in order to have any real effect on systems before they settle into exploitative structures.

—Kavita Philips, "How to Stop Worrying about Clean Signals and Start Loving the Noise," *Your Computer Is On Fire*

Course Outline

- Introduction to ASJ
- Criminal justice
- Interlude: Visionary futures
- Algorithmic epistemologies
- Data labor

Crash Course on Interdisciplinarity

A Humanities and Social Science Perspective

- **IE/CS/OR/... :**
  - We are building neutral, unbiased models.
  - You provide the data, we provide the results.
  - If there's garbage out, it means there was garbage in.

A Humanities and Social Science Perspective

- **IE/CS/OR/... :**
  - We are building neutral, unbiased models.
  - You provide the data, we provide the results.
  - If there's garbage out, it means there was garbage in.
- **Humanists and social scientists:**
  - Your models produce real harms for real people.
  - They might not create bias, but they amplify and propagate it.
  - Nothing is neutral.

Constructed Categories

- Most humanists and social scientists consider categories such as race and gender as **socially constructed** categories, not biological ones.
- These categories (and the inequalities that attach to them) are produced by social interactions.
- The categories do not simply reflect the world "as it is."

Constructed Categories

- Most humanists and social scientists consider categories such as race and gender as **socially constructed** categories, not biological ones.
- These categories (and the inequalities that attach to them) are produced by social interactions.
- The categories do not simply reflect the world "as it is."
- We (IE/CS/OR/...) like to think our tools are neutral, independent of social forces.
- But because the tools are part of social systems, they are biased, human, messy, just like anything else.
- In fact, as tools that operate with immense social force, they contribute to **constructing social categories** and magnify inequalities.

Constructed Categories

- ML and other algorithms contribute to constructing social categories.
- But this process is:
  - mostly invisible or ignored
  - virtually unregulated
  - mostly controlled by large Silicon Valley corporations, which are:
  - mostly controlled by a small, homogeneous group of people.
- This is a problem.
- (In addition to the problem of magnifying the harms that are attached to these categories.)

Introduction to ASJ

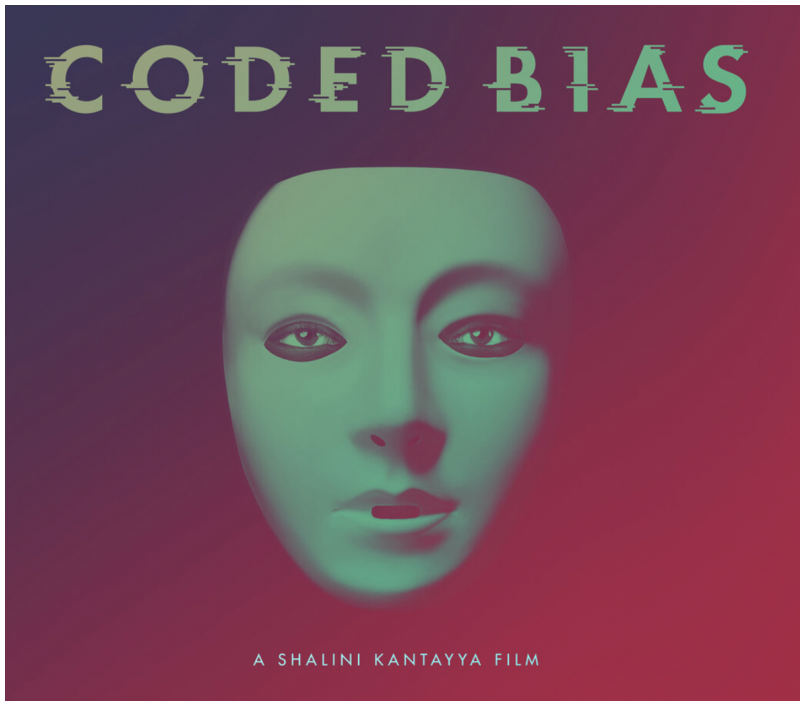
Why Algorithms and Social Justice Now?

CHAPTER 2

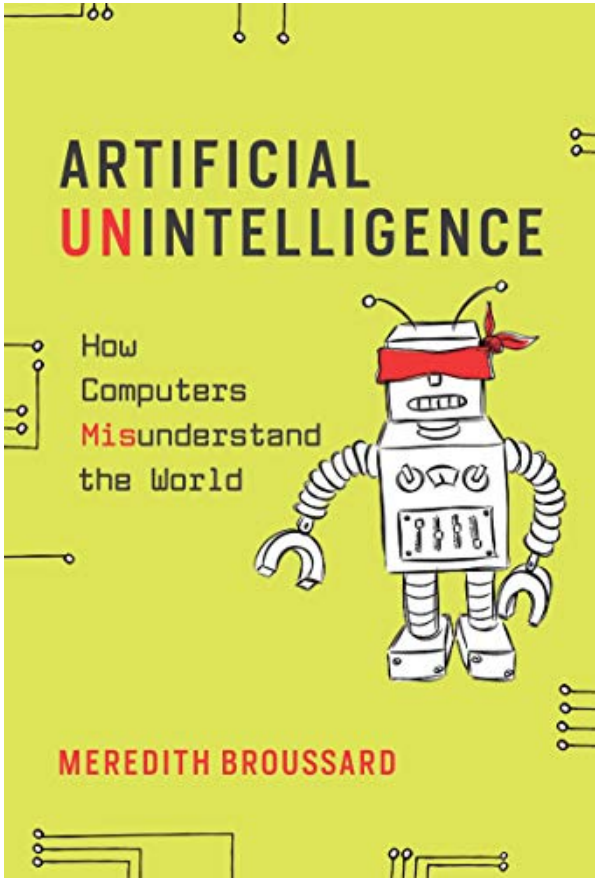
Five Faces of Oppression

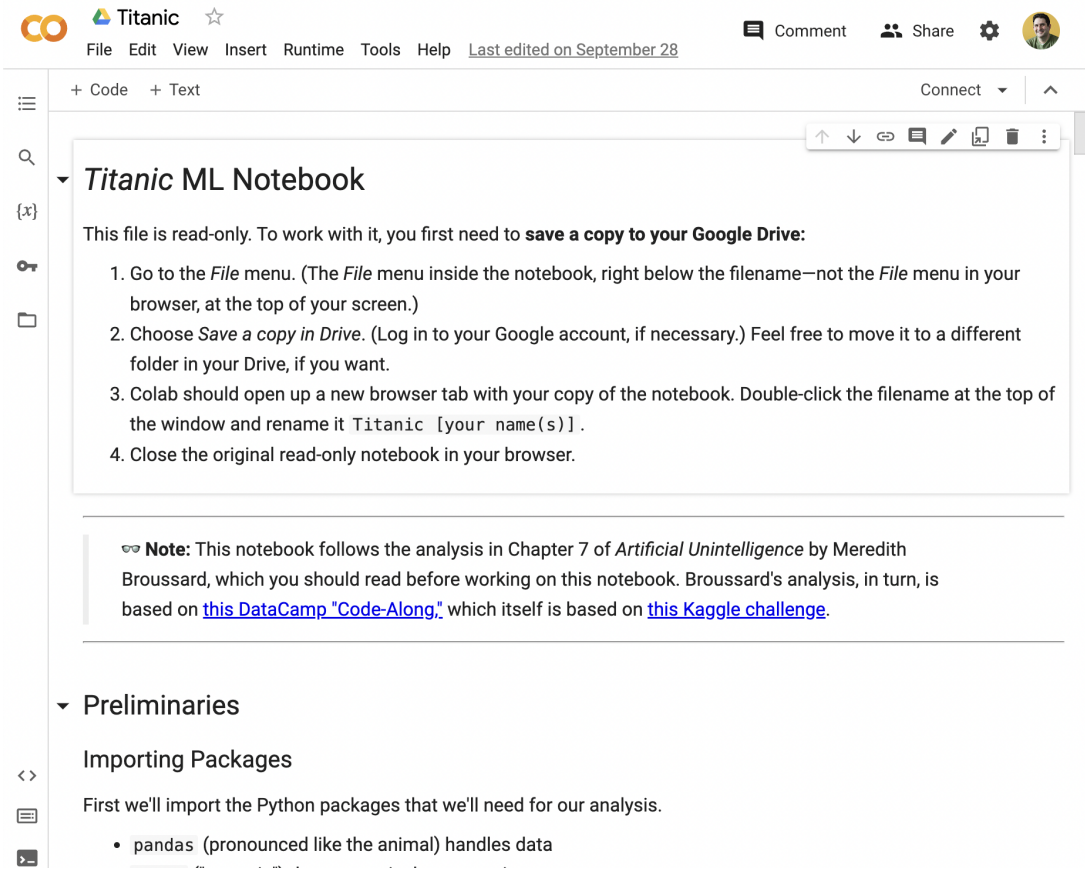
> Someone who does not see a pane of glass does not know that he does not see it. Someone who, being placed differently, does see it, does not know the other does not see it.
>
> When our will finds expression outside ourselves in actions performed by others, we do not waste our time and our power of attention in examining whether they have consented to this. This is true for all of us. Our attention, given entirely to the success of the undertaking, is not claimed by them as long as they are docile. . . .
>
> Rape is a terrible caricature of love from which consent is absent. After rape, oppression is the second horror of human existence. It is a terrible caricature of obedience.
>
> —Simone Weil

I HAVE proposed an enabling conception of justice. Justice should refer not only to distribution, but also to the institutional conditions necessary for the development and exercise of individual capacities and collective communication and cooperation. Under this conception of justice, injustice refers primarily to two forms of disabling constraints, oppression and domination. While these constraints include distributive patterns, they also involve matters which cannot easily be assimilated to the logic of distribution: decisionmaking procedures, division of labor, and culture.



What is AI?


ARTIFICIAL
UNINTELLIGENCE
How
Computers
Misunderstand
the World
MEREDITH BROUSSARD


Titanic
File Edit View Insert Runtime Tools Help Last edited on September 28
Comment
Share
+ Code + Text
Connect
Titanic ML Notebook
This file is read-only. To work with it, you first need to save a copy to your Google Drive:
1. Go to the File menu. (The File menu inside the notebook, right below the filename—not the File menu in your browser, at the top of your screen.)
2. Choose Save a copy in Drive. (Log in to your Google account, if necessary.) Feel free to move it to a different folder in your Drive, if you want.
3. Colab should open up a new browser tab with your copy of the notebook. Double-click the filename at the top of the window and rename it Titanic [your name(s)].
4. Close the original read-only notebook in your browser.
Note: This notebook follows the analysis in Chapter 7 of Artificial Unintelligence by Meredith Broussard, which you should read before working on this notebook. Broussard's analysis, in turn, is based on this DataCamp "Code-Along," which itself is based on this Kaggle challenge.
Preliminaries
Importing Packages
First we'll import the Python packages that we'll need for our analysis.
pandas (pronounced like the animal) handles data

Criminal Justice

Biometric Normativity and the Gaze

Physiognomy's New Clothes

Blaise Aguera y Arcas · Follow
38 min read · May 6, 2017

Listen Share ••• More

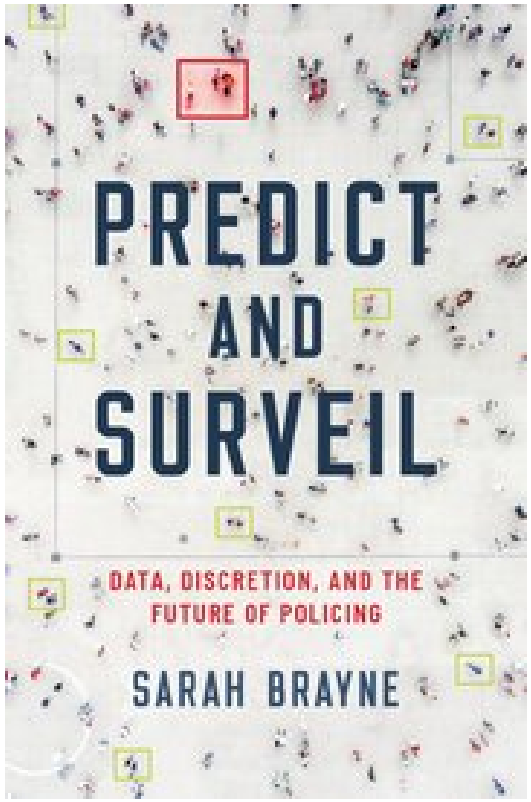
by Blaise Agüera y Arcas, Margaret Mitchell and Alexander Todorov

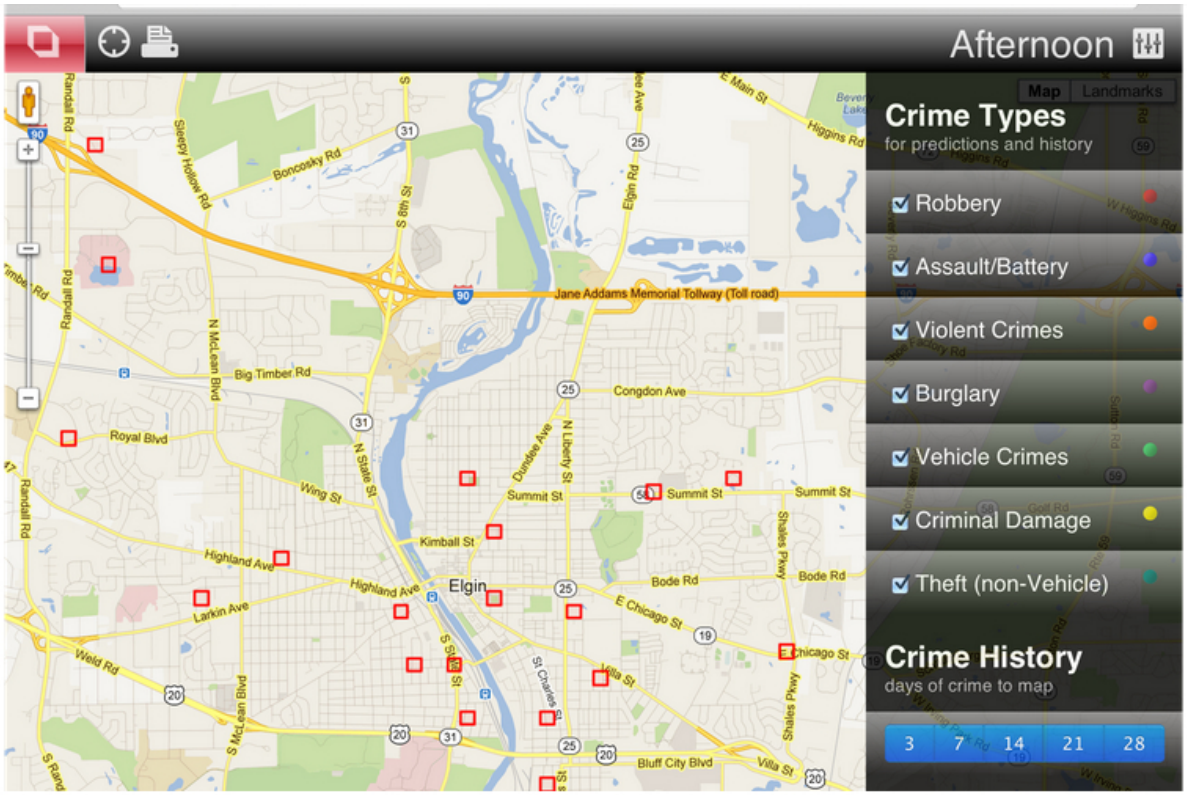
The Male Gazed

Surveillance, Power, and Gender

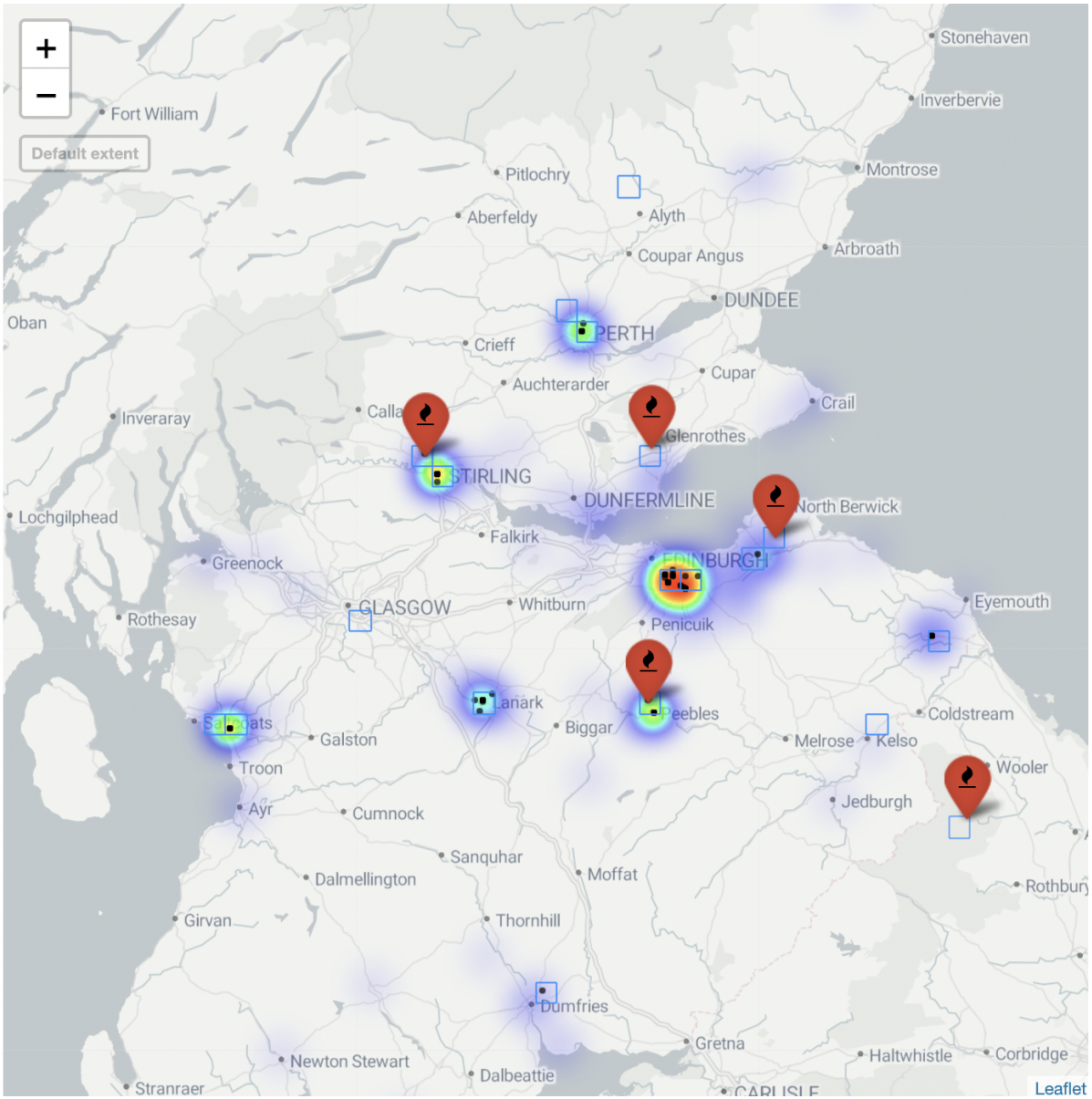
—*by* Kate Losse *on* **January 13th, 2014**

Predictive Policing and Feedback Loops

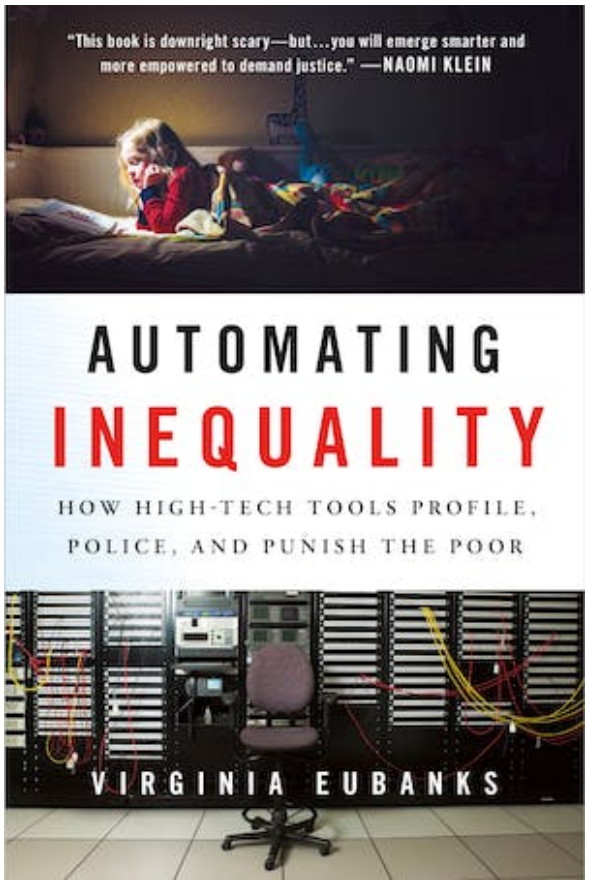




Scottish Witches App

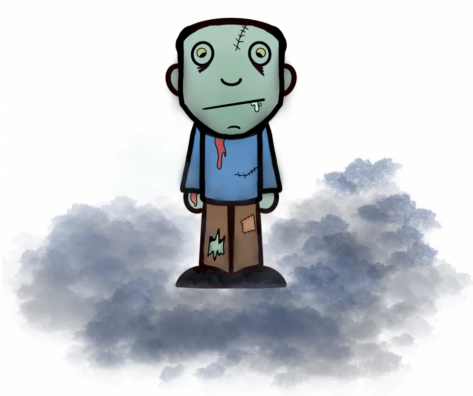


Predictive Algorithms

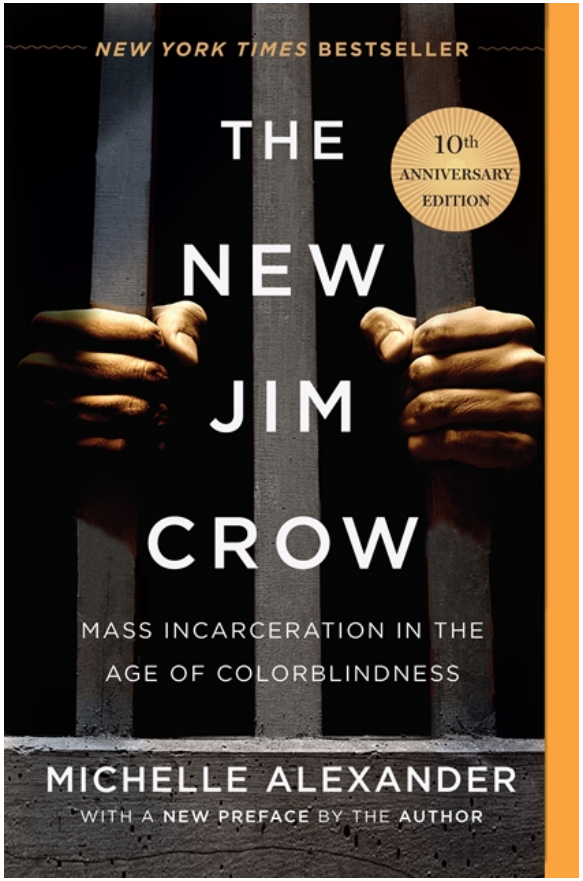


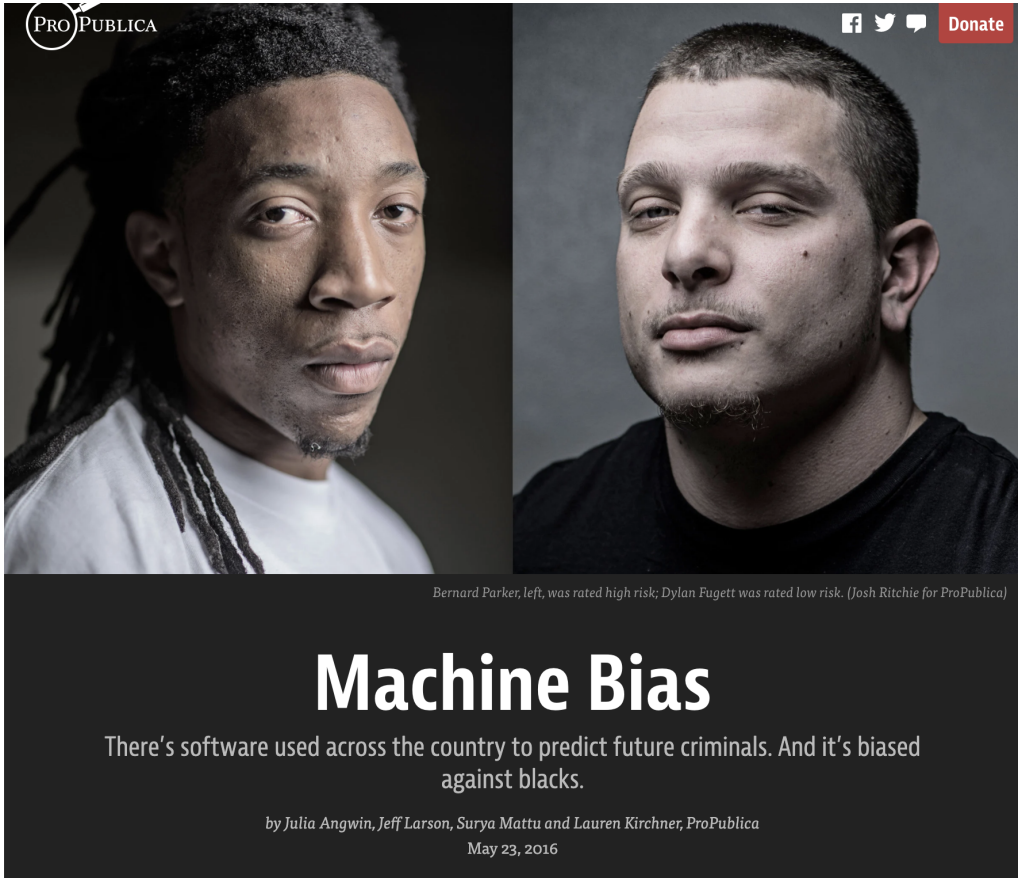
The "Zombie Predictive Model"

- False positives / false negatives
- The confusion matrix
- Threshold models



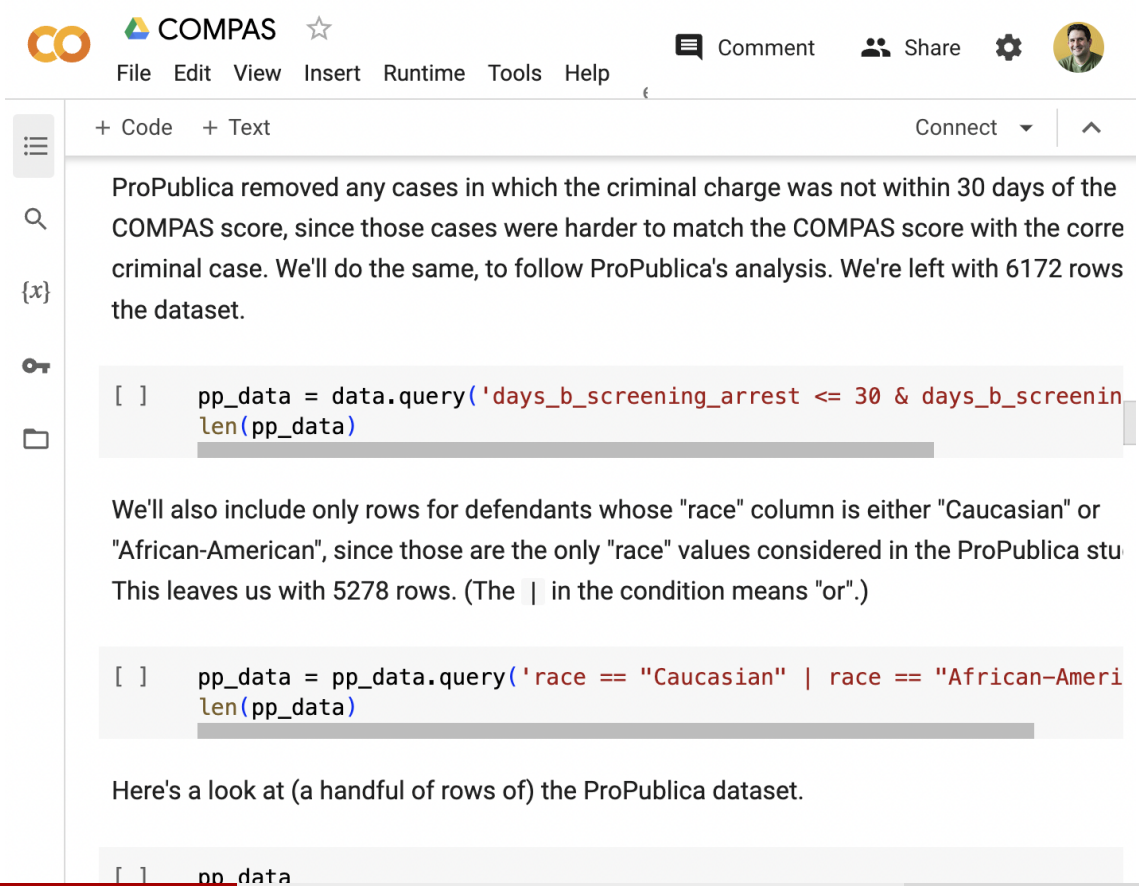
Recidivism-Risk Models





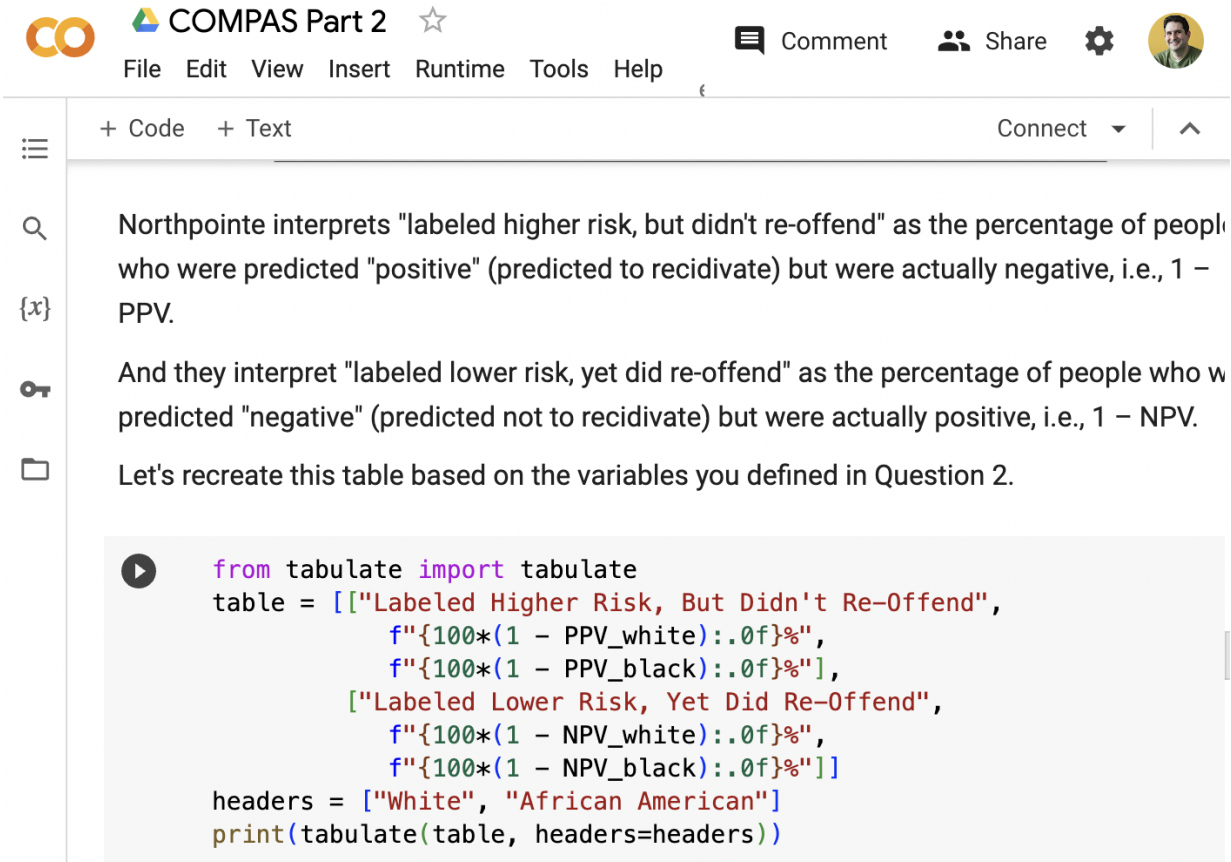
The COMPAS Conundrum

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%



The COMPAS Conundrum

	White	African American
Labeled Higher Risk, But Didn't Re-Offend	41%	37%
Labeled Lower Risk, Yet Did Re-Offend	29%	35%



The COMPAS Conundrum

Theorem

Equalized odds and predictive parity cannot both hold for the same predictor, unless the predictor is perfect *or the two groups have the* same base rate*.*

(Chouldechova 2016, Kleinberg, et al. 2016, Miconi 2017)

An Interpretable Model

nature machine intelligence

Search | Log in

Explore content | About the journal | Publish with us

nature > nature machine intelligence > perspectives > article

Perspective | Published: 13 May 2019

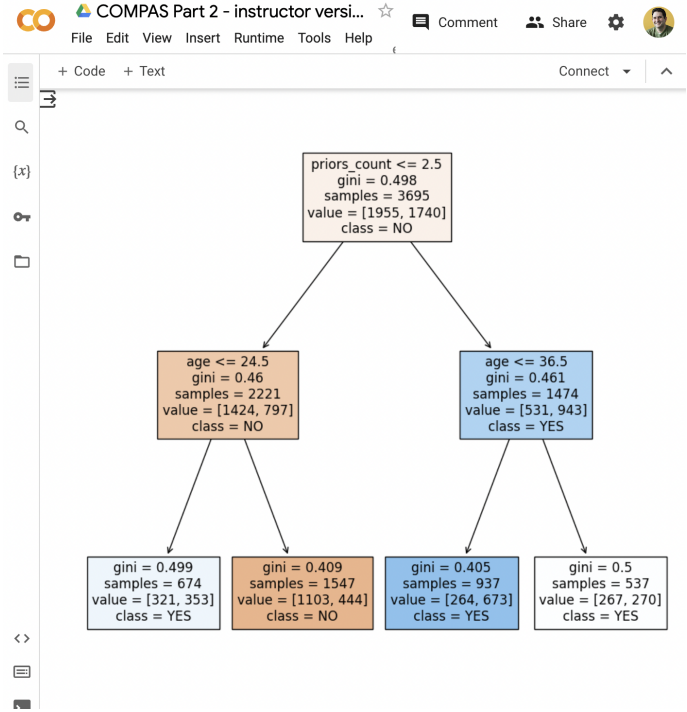
Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead

Cynthia Rudin

Nature Machine Intelligence **1**, 206–215 (2019) | Cite this article

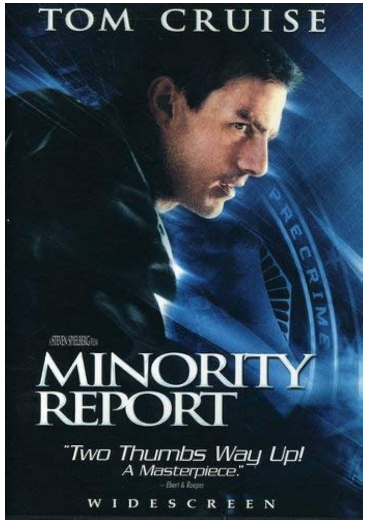
69k Accesses | **2541** Citations | **484** Altmetric | Metrics

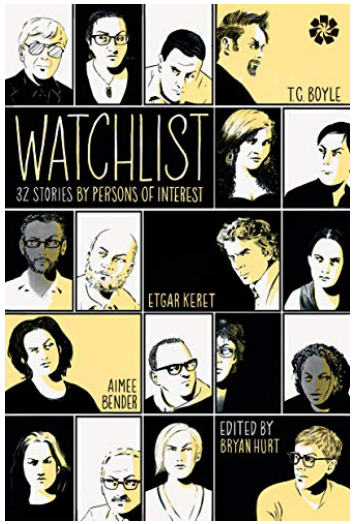
A preprint version of the article is available at arXiv.

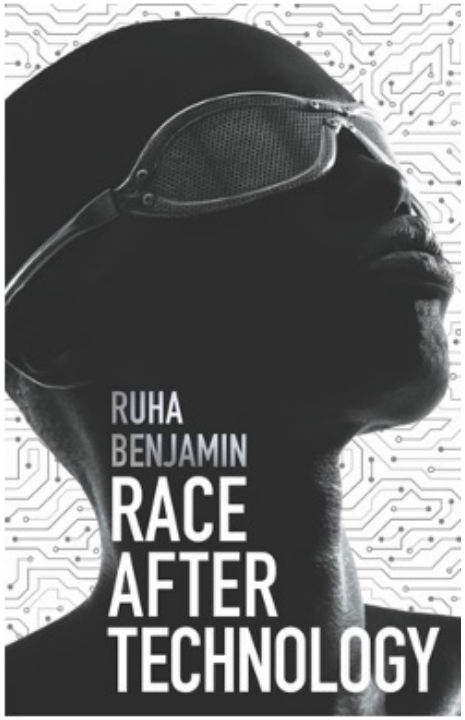


Interlude: Visionary Futures

Visionary Futures

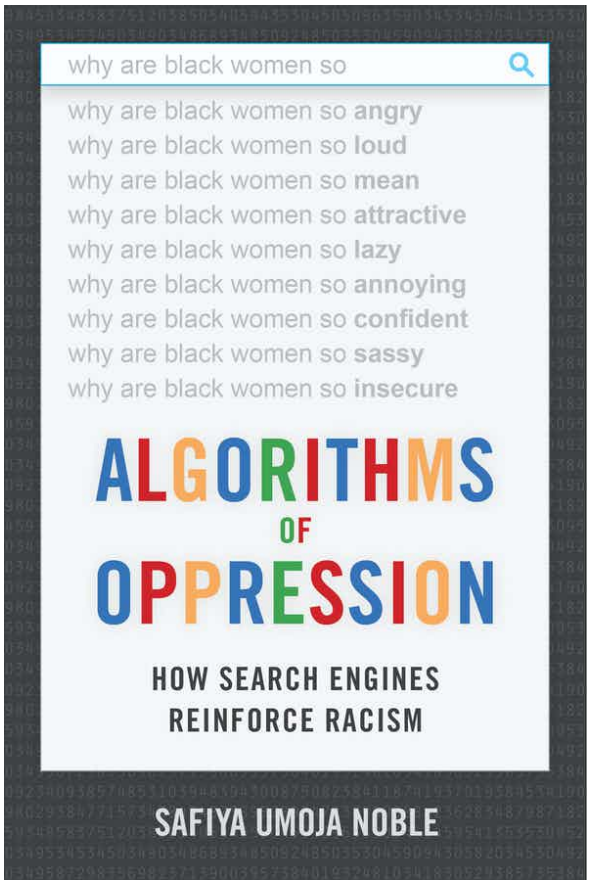






Algorithmic Epistemologies

Searching



Situating Search

Chirag Shah
chirags@uw.edu
University of Washington
Seattle, Washington, USA

Emily M. Bender
ebender@uw.edu
University of Washington
Seattle, Washington, USA

ABSTRACT

Search systems, like many other applications of machine learning, have become increasingly complex and opaque. The notions of relevance, usefulness, and trustworthiness with respect to information were already overloaded and often difficult to articulate, study, or implement. Newly surfaced proposals that aim to use large language models to generate relevant information for a user's needs pose even greater threat to transparency, provenance, and user interactions in a search system. In this perspective paper we revisit the problem of search in the larger context of information seeking and argue that removing or reducing interactions in an effort to retrieve presumably more relevant information can be detrimental to many fundamental aspects of search, including information verification, information literacy, and serendipity. In addition to providing suggestions for counteracting some of the potential problems posed by such models, we present a vision for search systems that are intelligent and effective, while also providing greater transparency and accountability.

CCS CONCEPTS

• **Information systems** → **Users and interactive retrieval**; **Language models**.

KEYWORDS

Search models; Language models; Information Seeking Strategies

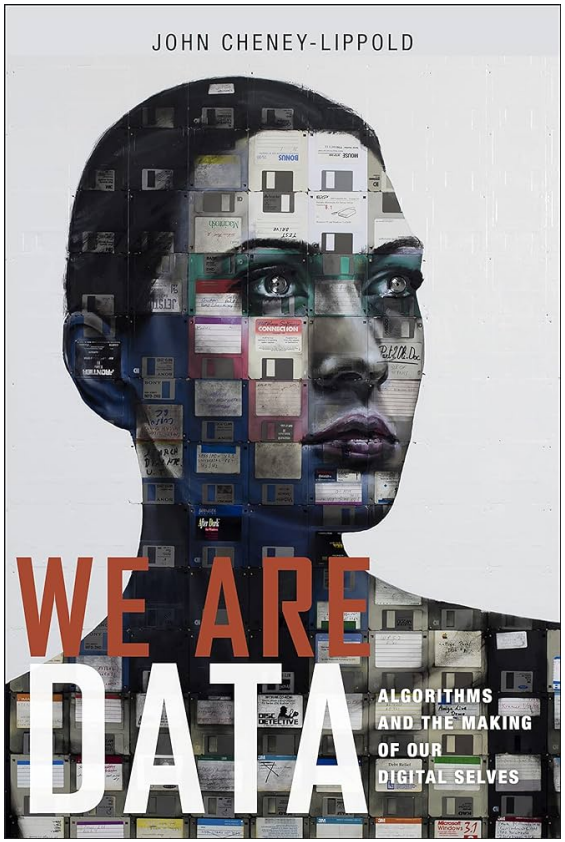
ACM Reference Format:
Chirag Shah and Emily M. Bender. 2022. Situating Search. In *Proceedings of the 2022 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '22), March 14–18, 2022, Regensburg, Germany.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3498366.3505816

user with the potentially useful information as quickly and effectively as possible. Examples include passage retrieval [58], question-answering systems [39], and dialogue or conversational systems [56]. We observe a trend towards valuing speed and convenience and ask: Is getting the user to a piece of relevant information as fast as possible the only or the most important goal of a search system? We argue that it should not be; that a search system needs to support more than matching or generating an answer; that an information processing system should provide more ways to interact with and make sense out of information than simply retrieving it based on programmed in notions of relevance and usefulness. More importantly, we argue that searching is a socially and contextually situated activity with diverse set of goals and needs for support that must not be boiled down to a combination of text matching and text generating algorithms.

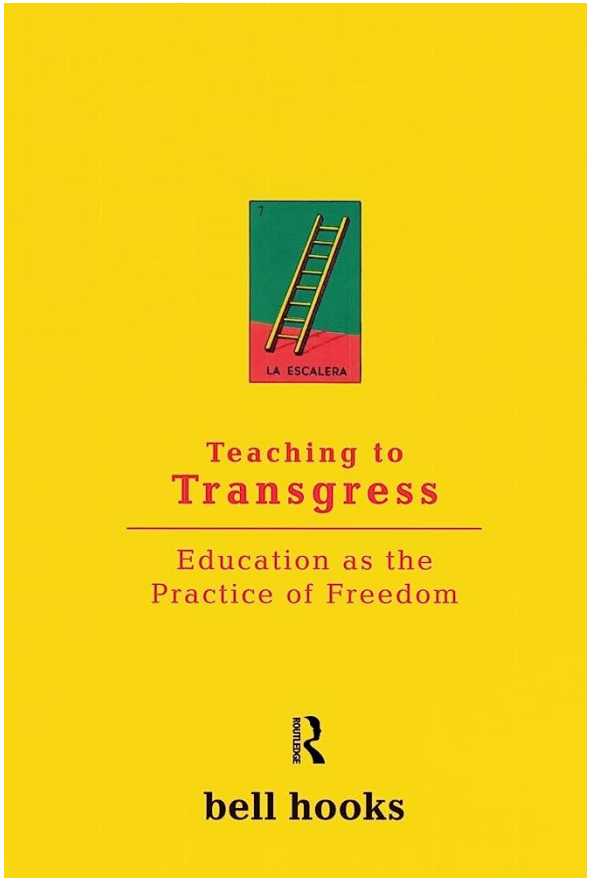
In this perspective paper we examine a couple of new proposals, which we refer to collectively as *the Google proposals* since they stem from Google, which involve closed-off systems that could generate relevant text in response to a user's queries, aiming to leverage large amounts of data and language models (LMs). We argue that such approaches miss the big picture of why people seek information and how that process contains value beyond simply retrieving relevant information. Beyond the critique of a few proposals, this paper offers a broad perspective of how search systems and society have evolved with each other and where we should go from here.

We begin by describing and examining the essential ideas of proposals by Metzler et al. [49] and Google in the next section (§2). We summarize why we believe these proposals are flawed in technical and conceptual terms. To better understand these flaws and to envision better systems, we need to examine search as an information seeking activity embedded in specific social and technological contexts. Therefore, we take a step back in Section 3 to

Digital Selves



US News Rankings



U.S.News & WORLD REPORT EDUCATION

Home / Education / Colleges / Best National University Ranki...

Best National University Rankings

BEST COLLEGES U.S.News & WORLD REPORT RANKINGS

Schools in the National Universities category, such as the University of Texas at Austin and the University of Vermont, offer a full range of undergraduate majors, plus master's and doctoral programs. These colleges also are committed to producing groundbreaking research. Read the methodology »

To unlock full rankings, SAT/ACT scores and more, sign up for the U.S. News College Compass!

Update: The 2024 rankings have been updated since the September 18, 2023 publication date. The updated rankings are noted below. For more information, please read this correction.

Some Takeaways

Every Model is Biased

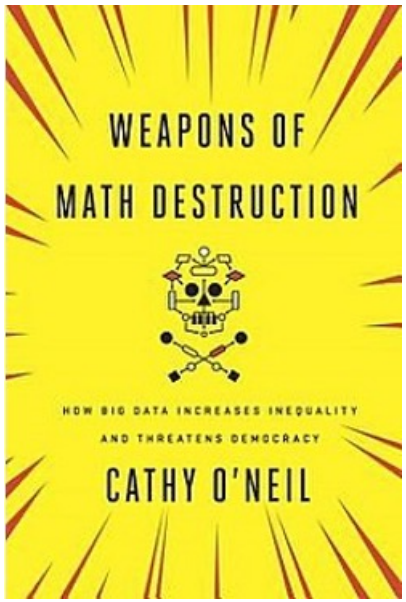
Models are opinions embedded in mathematics.

(O'Neil, *Weapons of Math Destruction*, p. 21)

- Every model is built, trained, and deployed by humans
- The models with the greatest impact are often built by extremely powerful private corporations, with little or no government oversight

The Errors Aren't iid

> *And the victims? Well, an internal data scientist might say, no statistical system can be* perfect. *Those folks are collateral damage. [...] Think of the astounding scale, and ignore the imperfections.*



pp. 12–13

But the errors are not iid. They magnify privilege and inequality.

It's Not Just about the Training Data

Rachel Thomas @math_rachel · Dec 7, 2020

Replying to @math_rachel

All data has context. All data has bias. We need to understand how & why it was collected, and in which contexts it makes sense to use. 2/

Rachel Thomas @math_rachel · Dec 7, 2020

Bias is just one ethical issue. Questions of power are crucial. If bias in facial recognition software is addressed, but USA police continue to use it to surveil protesters, we haven't solved the fundamental issue 3/

It's Not Just about the Models, Either

- As optimizers, we want to "fix" the problem by changing the formulation.
- Equity objectives, fairness constraints, etc.
- How would this translate to facial recognition / predictive policing / search engines?

The Big AI Players are Moving *Backwards* on Ethics

The New York Times

Google Researcher Says She Was Fired Over Paper Highlighting Bias in A.I.

Timnit Gebru, one of the few Black women in her field, had voiced exasperation over the company's response to efforts to increase minority hiring.

Timnit Gebru, a respected researcher at Google, questioned biases built into artificial intelligence systems. Cody O'Loughlin for The New York Times

The Big AI Players are Moving *Backwards* on Ethics

WINNI WINTERMEYER

by **Karen Hao**
March 11, 2021

Joaquin Quiñonero Candela, a director of AI at Facebook, was apologizing to his audience.

The Big AI Players are Moving *Backwards* on Ethics

THE WALL STREET JOURNAL.

POLITICS | NATIONAL SECURITY

Tech Industry Seeks Bigger Role in Defense. Not Everyone Is on Board.

Ex-Google CEO Eric Schmidt says the U.S. is in danger of losing its military edge if it doesn't join forces with Silicon Valley

Key Pentagon officials and members of Congress have endorsed ideas that former Google CEO Eric Schmidt and other tech-industry leaders back.

PHOTO: KEVIN DIETSCH/GETTY IMAGES

By Ryan Tracy

Sept. 7, 2021 9:00 am ET

Outcome vs. Intent

Jessica Rose @jesslynnrose · Jan 7

The function of a system is its output.

If you have dog grooming machine that sometimes smashes puppies and you keep running it, you're in the dog smashing business.

If you work for a mass surveillance company that keeps enabling genocide and undermining democracy...

27 1K 2.2K

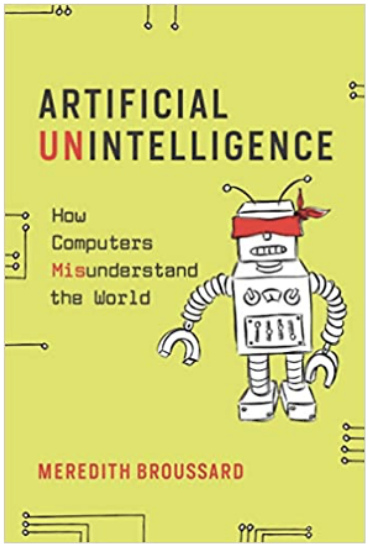
Start to Think More Critically

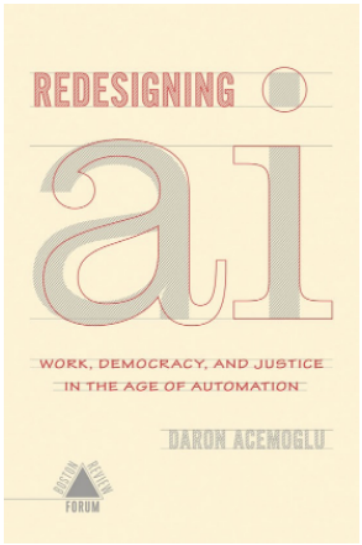
Rachel Thomas @math_rachel · Nov 3, 2019

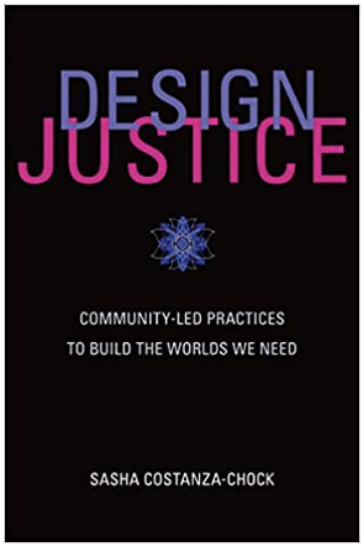
Some questions to ask when analyzing an algorithmic system:
Should we even be doing this?
What bias is in the data? (all data is biased, need to know how)
Error rates on different sub-groups? (e.g. GenderShades)
Is there an appeals process?
How diverse is the team that built it?

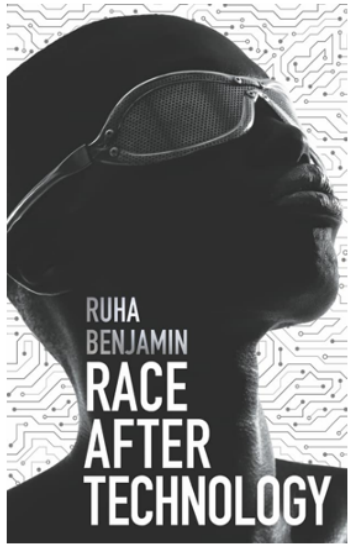
(Should we be asking similar questions about optimization/OR models?)

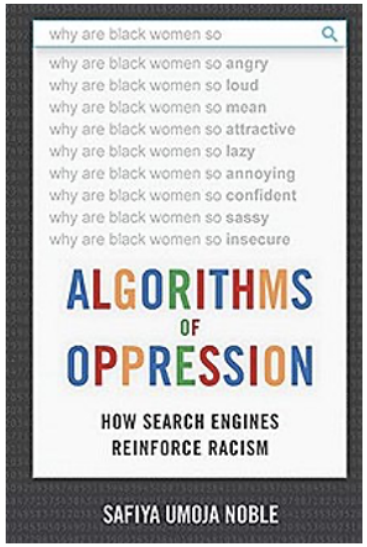
Read

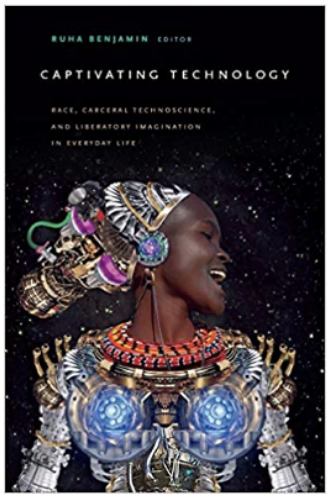

ARTIFICIAL
UNINTELLIGENCE
How
Computers
Misunderstand
the World
MEREDITH BROUSSARD

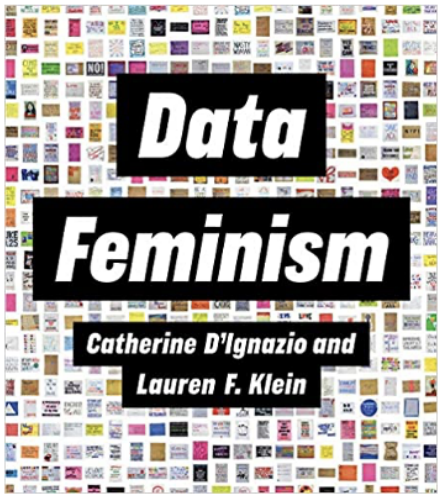

REDESIGNING
ai
WORK, DEMOCRACY, AND JUSTICE
IN THE AGE OF AUTOMATION
DARON ACEMOGLU

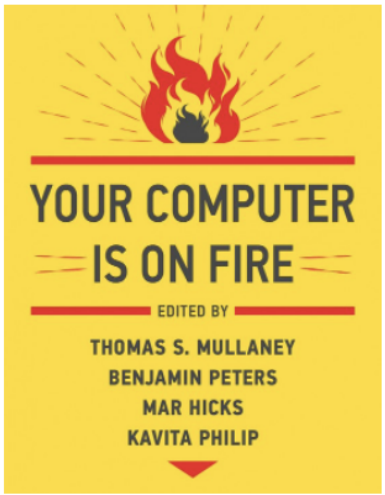

DESIGN
JUSTICE
COMMUNITY-LED PRACTICES
TO BUILD THE WORLDS WE NEED
SASHA COSTANZA-CHOCK

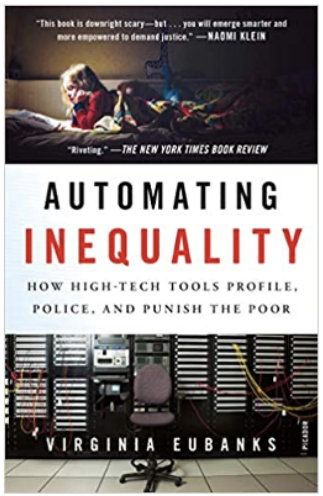

RUHA
BENJAMIN
RACE
AFTER
TECHNOLOGY

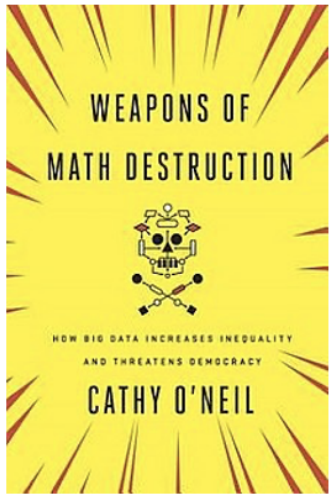

why are black women so
why are black women so angry
why are black women so loud
why are black women so mean
why are black women so attractive
why are black women so lazy
why are black women so annoying
why are black women so confident
why are black women so sassy
why are black women so insecure
ALGORITHMS
OF
OPPRESSION
HOW SEARCH ENGINES
REINFORCE RACISM
SAFIYA UMOJA NOBLE


RUHA BENJAMIN EDITOR
CAPTIVATING TECHNOLOGY
RACE, CARCERAL TECHNOSCIENCE,
AND LIBERATORY IMAGINATION
IN EVERYDAY LIFE


Data
Feminism
Catherine D'Ignazio and
Lauren F. Klein


YOUR COMPUTER
IS ON FIRE
EDITED BY
THOMAS S. MULLANEY
BENJAMIN PETERS
MAR HICKS
KAVITA PHILIP


"This book is downright scary—but . . . you will emerge smarter and more empowered to demand justice." —NAOMI KLEIN
"Riveting." —THE NEW YORK TIMES BOOK REVIEW
AUTOMATING
INEQUALITY
HOW HIGH-TECH TOOLS PROFILE,
POLICE, AND PUNISH THE POOR
VIRGINIA EUBANKS


WEAPONS OF
MATH DESTRUCTION
HOW BIG DATA INCREASES INEQUALITY
AND THREATENS DEMOCRACY
CATHY O'NEIL

Listen


A.I. Nation
WHYY
PRINCETON
UNIVERSITY


THE RADICAL AI PODCAST


IN
MACHINES
WE
TRUST
MIT Technology Review

Follow

- Sarah T. Roberts @ubiquity75
- Ayodele @DataSciBae
- Andrea Hicks @Andrea_HicksPhD
- Yim Register @yimregister
- Meredith Whittaker @mer__edith
- Jessica Rose @jesslynnrose
- Rachel Thomas @math_rachel
- Mar Hicks @histoftech
- Cathy O'Neil @mathbabedotorg
- Timnit Gebru @timnitGebru
- Meredith Broussard @merbroussard
- Shannon Vallor @ShannonVallor
- Casey Fiesler @cfiesler
- Alex Hanna @alexhanna
- Ruha Benjamin @Ruha9
- MMitchell @mmitchell_ai
- Karen Hao @_KarenHao
- Algorithmic Justice League @AJLUnited
- Emily M. Bender @emilymbender
- Tom Mullaney @tsmullaney
- Paul N. Edwards @AVastMachine
- Safiya Umoja Noble @safiyanoble
- Tawana Petty @Combsthepoet
- Luke Stark @luke_stark
- Joy Buolamwini @jovialjoy
- Jiahao Chen @acidflask
- Sasha Costanza-Chock @schock
- Electronic Frontier Foundation @EFF

...and many others

Questions?

larry.snyder@lehigh.edu
sme6@lehigh.edu

github.com/LarrySnyder/ASJ

Thank You!

larry.snyder@lehigh.edu
sme6@lehigh.edu

github.com/LarrySnyder/ASJ