

Ethics, Bias and AI in Medicine: More than Just Data

Matthew DeCamp, MD PhD

Associate Professor

Division of General Internal Medicine & Center for Bioethics and Humanities



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Disclosures

I have no relevant conflicts of interest to disclose.

My research has been funded by the NIH, Agency for Healthcare Research and Quality, R2HC programme, the Greenwall Foundation, and the American College of Physicians.



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Objectives

By the end of this presentation, participants will be able to:

1. Describe why bias is such a pervasive problem for artificial intelligence
2. Explain how “bias” is not just a data problem



<https://commons.wikimedia.org/wiki/File:Aspen-PopulusTremuloides-2001-09-27.jpg>

A Provocative Starter

CU

University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Where Bioethics and Informatics Intersect



A prisoner in a compression chamber loses consciousness (and later dies) during an experiment to determine at what altitude aircraft crews could survive without oxygen. Dachau, Germany, December 1942.



Victim of a medical experiment immersed in freezing water at the Dachau concentration camp. Dachau, Germany, between August 1942 and May 1943.



Defendants seated under guard in the dock behind the defense counsel during the Doctors Trial, which was held in Nuremberg, Germany, from December 9, 1946, to August 20, 1947.

Respect for Choice – Risks & Benefits – Justice as Fairness

<https://www.ushmm.org/information/exhibitions/online-exhibitions/special-focus/doctors-trial>



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Where Bioethics and Informatics Intersect

How IBM Technology Jump Started the Holocaust



Edwin Black

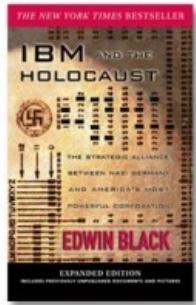
6/16/11 11:07pm • Filed to: IBM ▾



109



Save



IBM AND THE HOLOCAUST

THE STRATEGIC ALLIANCE
BETWEEN NAZI GERMANY
AND AMERICA'S MOST
POWERFUL CORPORATION

Kinder	Se	Werksta	Merkmale	Familie	Körper	Sitz-	Ges	Rassenanteil
0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9

RASSENAMT - SS
1

At the vanguard of Hitler's intellectual shock troops were the statisticians. Naturally, statistical offices and census departments were Dehomag's [IBM's German subsidiary] number one clients. In their journals, Nazi statistical experts boasted of what they expected their evolving science to deliver. All of their high expectations depended on the continuing innovation of IBM punch cards and tabulator technology. Only Dehomag could design and execute systems to identify, sort, and quantify the population to separate Jews from Aryans.



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Were statisticians (“informaticists”) complicit in the Nazi evil?

Would thinking about ethics have made any difference?

Was this just an example of an ethically neutral “tool” being used for wrongdoing?



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

The Potential of AI in Health Care



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities



Some AI Basics

Artificial Intelligence refers broadly to machines that can perform tasks typically thought to require human judgment/intelligence.

ARTIFICIAL INTELLIGENCE		
NARROW	GENERAL	SUPER
Performs a specific task (e.g., diabetic retinopathy diagnosis)	Capable of performing any or a range of tasks (e.g., science fiction robots)	Capable of exceeding human capacity (e.g., moral AI)

*In bioethics, “super AI” is intriguing because it brings up the potential that should **exceed** (not just equal) the capacity of physicians...or even **exceed our moral abilities** (by avoiding human-created biases).*

The “Explainability” Problem



[This Photo](#) is licensed under [CC BY-SA](#)

One of the major challenges for AI can be when the inner workings are not just unknown – but unknowable.
*You can imagine this *could* complicate the ability to obtain fully informed “consent.”*

Will AI in Health Care Live Up to the Hype?



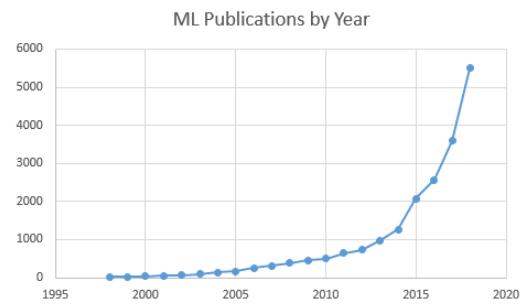
- Up to \$6 billion (10-fold increase) by 2021

(Frost & Sullivan. Transforming healthcare through artificial intelligence systems. 2016)

Table 2 FDA AI approvals are accelerating

From: High-performance medicine: the convergence of human and artificial intelligence

Company	FDA Approval	Indication
Apple	September 2018	Atrial fibrillation detection
Aidoc	August 2018	CT brain bleed diagnosis
iCAD	August 2018	Breast density via mammography
Zebra Medical	July 2018	Coronary calcium scoring
Bay Labs	June 2018	Echocardiogram EF determination
Neural Analytics	May 2018	Device for paramedic stroke diagnosis
IDx	April 2018	Diabetic retinopathy diagnosis
Icometrics	April 2018	MRI brain interpretation
Imagen	March 2018	X-ray wrist fracture diagnosis
Viz.ai	February 2018	CT stroke diagnosis
Arterys	February 2018	Liver and lung cancer (MRI, CT) diagnosis
MaxQ-AI	January 2018	CT brain bleed diagnosis
Alivecor	November 2017	Atrial fibrillation detection via Apple Watch
Arterys	January 2017	MRI heart interpretation



- Dozens of FDA “approvals” in the past 2 years

- Exponential increase in publications

Use Cases Across the Health Care Spectrum

- Research & Development (e.g., screening drug targets)
- Diagnosis (e.g., pathology, radiology)
- Treatment Decision-making (e.g., cancer treatments)
- Prognostication (e.g., about mortality)
- Resource Allocation (e.g., coverage decisions)
- Surveillance (e.g., public health & COVID-19)

Topol EJ. *Nature Medicine* 2019;25:44-56.

Uses Across Translational Science

Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients ⁸

Peter Heydon¹, Catherine Egan^{1,2}, Louis Bolter³, Ryan Chambers³, John Anderson³, Steve Aldington⁴, Irene M Stratton⁴, Peter Henry Scanlon⁴, Laura Webster⁵, Samantha Mann⁵, Alain du Chemin⁵, Christopher G Owen⁶, Adnan Tufail^{1,2}, Alicia Regina Rudnicka⁶

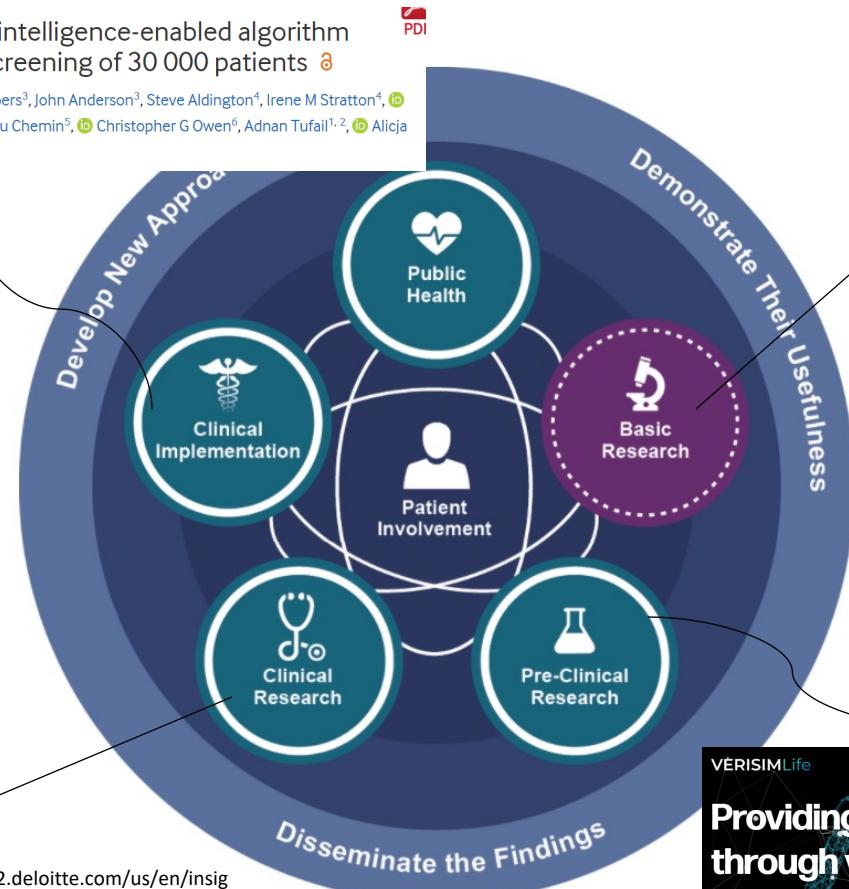
TRIAL CONDUCT

Assess site performance (e.g. enrolment and dropout rates) with real-time monitoring.

Analyse digital biomarkers on disease progression, and other quality-of-life indicators.

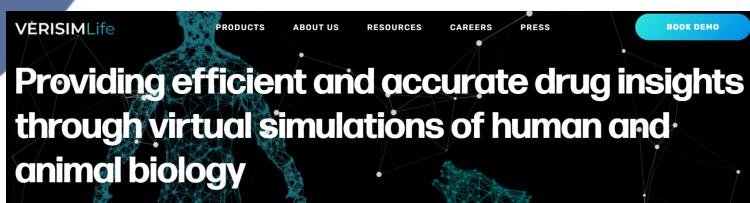
Automate sharing of data across multiple systems.

<https://www2.deloitte.com/us/en/insights/industry/life-sciences/artificial-intelligence-in-clinical-trials.html>



Hundreds of start-up companies using AI in drug discovery

<https://blog.benchsci.com/startups-using-artificial-intelligence-in-drug-discovery>



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

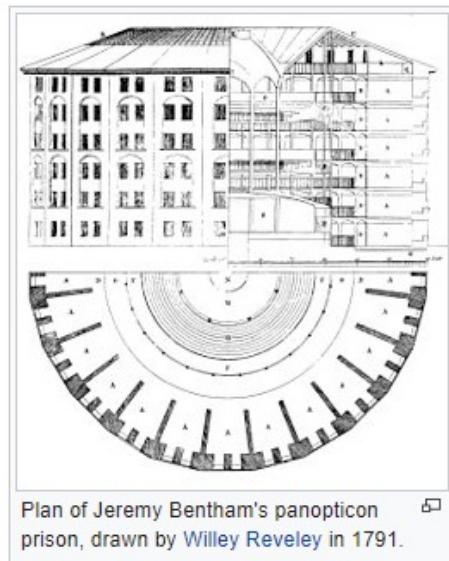
AI and Surveillance

Towards Vision-Based Smart Hospitals: A System for Tracking and Monitoring Hand Hygiene Compliance

Albert Haque, Michelle Guo, Alexandre Alahi, Serena Yeung, Zelun Luo, Alisha Rege, Jeffrey Jopling, Lance Downing, William Beninati, Amit Singh, Terry Platckok, Arnold Milstein, Li Fei-Fei

(Submitted on 1 Aug 2017 ([v1](#)), last revised 24 Apr 2018 (this version, v3))

One in twenty-five patients admitted to a hospital will suffer from a hospital acquired infection. If we can intelligently track healthcare staff, patients, and visitors, we can better understand the sources of such infections. We envision a smart hospital capable of increasing operational efficiency and improving patient care with less spending. In this paper, we propose a non-intrusive vision-based system for tracking people's activity in hospitals. We evaluate our method for the problem of measuring hand hygiene compliance. Empirically, our method outperforms existing solutions such as proximity-based techniques and covert in-person observational studies. We present intuitive, qualitative results that analyze human movement patterns and conduct spatial analytics which convey our method's interpretability. This work is a step towards a computer-vision based smart hospital and demonstrates promising results for reducing hospital acquired infections.



POLITICO AI: Decoded: How AI is helping fight a pandemic — Europe's coronavirus app — Insights from Valencia

By JANOSCH DELCKER | 4/8/20, 11:00 AM CET

The constant surveillance panopticon (*thanks to Glenn Cohen for the comparison*)



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Many, Many Ethical Issues

ETHICS CONCEPT	EXAMPLE ISSUES
Respect for Choice ("autonomy")	Privacy Informing patients Clinician choice
Beneficence ("well-being," risk/benefit)	Risks and benefits Managing AI-based errors
Justice (fairly distributing benefits)	Bias Stigma

Today I am going to focus on bias – though I will also hint at how issues of bias affect other ethics issues.

Ethics Issues: Not just Theoretical

EXCLUSIVE

STAT+

IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show

By CASEY ROSS @caseymross and IKE SWETLITZ / JULY 25, 2018

Internal IBM documents show that its Watson supercomputer recommended 'unsafe and incorrect' cancer treatment advice and that company customers identified "multiple examples of unsafe and erroneous recommendations" as IBM was promoting the product to hospitals around the world.

Major Security Flaws Found in South Korea Quarantine App

The defects, which have been fixed, exposed private details of people in quarantine. The country has been hailed as a pioneer in digital public health.



Google and the University of Chicago Are Sued Over Data Sharing

The New York Times

By Daisuke Wakabayashi

June 26, 2019



SAN FRANCISCO — When the University of Chicago Medical Center announced a partnership to share patient data with Google in 2017, the alliance was promoted as a way to unlock information

electronic health records and improve predictive medicine.

But, the University of Chicago, the medical center and Google, sued in a potential class-action lawsuit accusing the search giant without stripping identifiable date stamps or

The Pervasive Problem of Bias in AI



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Prominent Examples of Bias in AI

MIT Technology Review

Topics Magazine Newsletters Events

Sign in

Subscribe

Tech policy / AI Ethics

AI is sending people to jail — and getting it right



ALGORITHM
WATCH

Newsletter · Events ·

ABOUT / RESEARCH / STORIES /

Search...

Using historical data to train risk assessment tools, machines are copying the mistakes of humans

by Karen Hao

Story

Google apologizes after its Vision AI produced racist results



Examples of Bias in Health Care AI

RESEARCH ARTICLE

Dissecting racial bias in an algorithm used to manage the health of populations

 Ziad Obermeyer^{1,2,*}, Brian Powers³, Christine Vogeli⁴,  Sendhil Mullainathan^{5,*†}

* See all authors and affiliations

Science 25 Oct 2019;
Vol. 366, Issue 6464, pp. 447-453
DOI: 10.1126/science.aax2342

DISCOVER LATEST OBSESSIONS

QUARTZ

FEATURED

 Google's AI for mammograms doesn't account for racial differences

A commercial algorithm disadvantages black patients for directing care resources.

Annals of Internal Medicine®

LATEST ISSUES IN THE CLINIC JOURNAL CLUB MULTIMEDIA CME / MOC AUTHORS / SUBMIT

Special Article | 18 December 2018

Ensuring Fairness in Machine Learning to Advance Health Equity

Alvin Rajkomar, MD , Michaela Hardt, PhD, Michael D. Howell, MD, MPH, Greg Corrado, PhD, ... 

Concerns over diagnostic performance that differs by race or ethnicity.

Case management for early hospital discharge (would have) preferentially given resources to wealthy white individual.

Center for Bioethics and Humanities

Why is Bias in AI such a Pervasive Problem?

The usual answer goes something like this:



And in fact, data are a large part of the problem:

- Radical inequities in clinical trial participation
- Historical patterns of racism and inequities in access that cast a shadow over current algorithms
- Algorithms that start fair can “learn” these patterns - even when the variable in question is removed (adaptability)

But today I want to suggest that the problem may be even worse.

Before We Get There, What is “Bias”?

BIAS could be defined as any disproportionate, systematic weighting of an idea or thing.

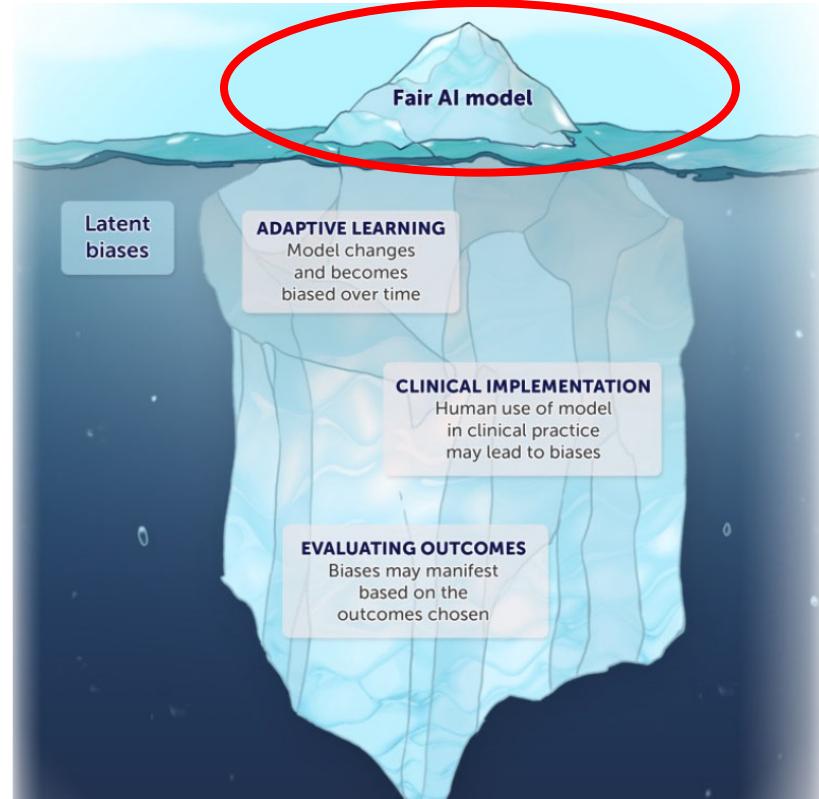
And we know that bias can be:

- Intended (explicit) or unintended (implicit)
- The product of cognitive errors (“anchoring bias”) or a cognitive intent (“reporting bias,” i.e., misconduct)

If we keep this more general idea of bias in mind, we can see that biases have implications for far more than just the data.

Biases in Data: The Tip of the Iceberg

The implication of understanding bias in this was is that even if we solve the “data problem” and create a fair model, *latent biases* lie beneath the surface – biases waiting to happen.



DeCamp M, Lindvall C. Latent bias and the implementation of artificial intelligence in medicine [published online ahead of print, 2020 Jun 18]. *J Am Med Inform Assoc.* 2020;ocaa094. doi:10.1093/jamia/ocaa094

Three More Examples of Bias in AI: *Bias in selection of the scientific problem or task*

Systematic favoring of certain problems or tasks that are “worthy” of solving is itself a bias.



[This Photo](#) is licensed under [CC BY-SA](#)

What reasons do we have to believe this might be an issue?

RESEARCH ARTICLE | SCIENTIFIC COMMUNITY

Topic choice contributes to the lower rate of NIH awards to African-American/black scientists

<https://www.theverge.com/2019/4/3/18293410/ai-artificial-intelligence-ethics-boards-charters-problem-big-tech>

Travis A. Hoppe^{1,2}, Aviva Litovitz^{1,2}, Kristine A. Willis^{3,*}, R

+ See all authors and affiliations

Science Advances 09 Oct 2019;
Vol. 5, no. 10, eaaw7238
DOI: 10.1126/sciadv.aaw7238

University of Colorado
Anschutz Medical Camp

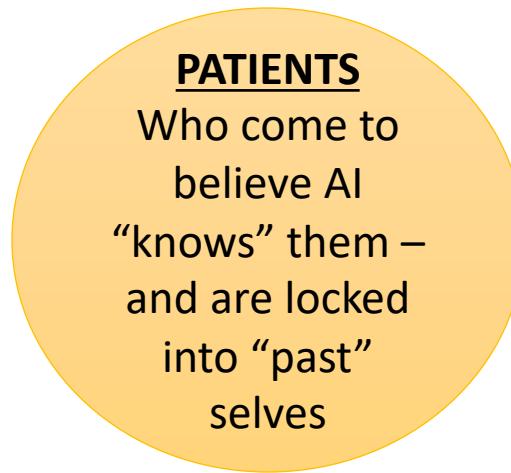
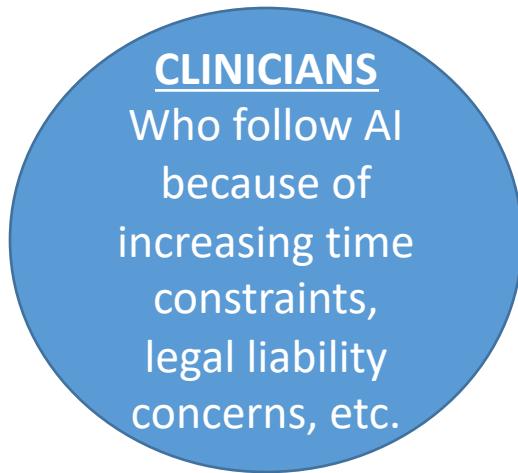
THE PROBLEM WITH AI ETHICS

Is Big Tech's embrace of AI ethics boards actually helping anyone?

By James Vincent | Apr 3, 2019, 11:47am EDT

Three More Examples of Bias in AI: *Bias in “Choices”*

Automation bias is the phenomenon where we tend to over-rely on technology or follow it unquestioningly.



Biases in “choices” can be particularly problematic in settings of historical distrust, racism, and so on.

Three More Examples of Bias in AI: *Bias in Application*

When certain AI applications are only used or accessible in particular settings, it can be known as “privilege bias.” Of course, our hope is that AI *reduces* privilege bias by leapfrogging.



However, we need accountability measures for ensuring AI's benefits and risks are accessible equitably.

Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring Fairness in Machine Learning to Advance Health Equity. *Ann Intern Med.* 2018;169(12):866-872

Implications of a More Comprehensive Account of “Bias”

- Issues of bias are primarily - but not only - issues of justice; they also affect *choice* and *beneficence* (well-being).
- It will not be enough to create “fair” systems because these systems operate in an unfair world.
- It will not be enough to keep systems as “decision support” only (i.e., non-autonomous AI).

So what should we do?



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

What Can We Do?



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Building “Trust” in AI is Not the Answer.

Building trust in AI



[What's next for AI](#) [Featured articles](#) ▾ [Featured interviews](#) ▾

AI is no longer the future—it's now here in our living rooms and cars and, often, our pockets. As the technology has continued to expand its role in our lives, an important question has emerged: What level of trust can—and should—we place in these AI systems?

To explore this question, we spoke to 30 AI scientists and leading thinkers. They told us that building trust in AI will require a significant effort to instill in it a sense of morality, operate in full transparency and provide education about the opportunities it will create for business and consumers. And, they say, this effort must be collaborative across scientific disciplines, industries and government.



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Why We Cannot “Trust” AI

The richest concept of trust requires the following:

1. *Vulnerability* – placing something of value at risk to another’s discretion.
2. An *expectation* that what was entrusted will be fulfilled
3. An implicit or explicit promise of the trustee
4. Mutual understanding of both parties
5. Discretion on the part of the trustee regarding how to fulfill the trust
6. Confidence in the motive of the trustee

Not only is AI by this definition incapable of trust – but also, I do not want to “trust” AI, I want to be assured it is not biased.



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

What Can We Do?

- (1) **We must address issues of “data” causes of fairness in the algorithms...but it’s not a matter of “tweaking” or “correcting” them.**

Addressing biases in AI requires us to address longer standing issues in diversity among scientists, clinical trials recruitment, health system distrust, structural racism, implicit biases, etc.

- (2) **We must address issues of bias *ex ante* (not wait for a *Science* publication to reveal them)...and we might even want to consider issues of bias as safety issues.**

Recent FDA guidance does not mention bias!

- (3) **As AI is implemented, we must use translational D&I science to support dissemination to not just “more” people but to ensure the “right” people.**



Yousefi Nooraie, R., Kwan, B., et al. (2020). Advancing health equity through CTSA programs: Opportunities for interaction between health equity, dissemination and implementation, and translational science. *Journal of Clinical and Translational Science*, 4(3), 168-175

An example of managing bias *ex ante*

MIT AI tool can predict breast cancer up to 5 years early, works equally well for white and black patients

Darrell Etherington @etherington / 4 weeks ago



Comment

MIT's Computer Science and Artificial Intelligence Lab has developed a new deep learning-based AI prediction model that can anticipate the development of breast cancer up to five years in advance. Researchers working on the product also recognized that other similar projects have often had inherent bias because they were based overwhelmingly on white patient populations, and specifically designed their own model so that it is informed by "more equitable" data that ensures it's "equally accurate for white and black women."



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Back to the Bigger Picture

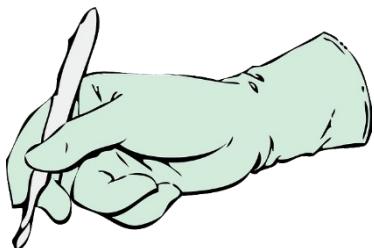


University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Mere tools?

You might still think, “Wait a minute, these are just how AI is used, not problems with technology itself.”



The essence of technology is by no means anything technological.

— Martin Heidegger —

AZ QUOTES

On the one hand: “things we use” to achieve our end goals.

On the other hand: tools fundamentally change the way we see the world – even changing our end goals.

Karches K. Theoretical Medicine and Bioethics. 2018; 39(2):91-110.



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Is AI simply a “tool”?

<http://chronicle.com/weekly/v50/i21/21b02601.htm>

How Computers Change the Way We Think

By SHERRY TURKLE

The tools we use to think change the ways in which we think. The invention of written language brought about a radical shift in how we process, organize, store, and transmit representations of the world. Although writing remains our primary information technology, today when we think about the impact of technology on our habits of mind, we think primarily of the computer.

From the issue dated January 30, 2004

The Chronicle of Higher Education

<http://www.chronicle.com>

Section: The Chronicle Review

Volume 50, Issue 21, Page B26

Scholarship in sociology points out that the tools we use can change the way we think and act toward each other. If this is true, it has significant implications for ethics.

It's not just that we need ethics to inform our use of tools – if tools are not neutral, we must continually be aware of how our core values and technology affect each other.



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

How might AI change the way we think?

- Will it result in “reductionism” about patients (e.g., seeing them as packets of data)?
- Could it depersonalize care (or enable greater empathy)?
- Might it change the way we design health care spaces for the sake of surveillance (rather than healing)?
- Others?

Should some uses be off limits?

Original Investigation | Health Informatics



July 12, 2019

Development and Validation of a Deep Learning Algorithm for Mortality Prediction in Selecting Patients With Dementia for Earlier Palliative Care Interventions

Liqin Wang, PhD^{1,2}; Long Sha, MS³; Joshua R. Lakin, MD^{1,4,5}; [et al](#)

[» Author Affiliations](#) | Article Information

JAMA Netw Open. 2019;2(7):e196972. doi:10.1001/jamanetworkopen.2019.6972

Key Points

| [Español](#) | [中文 \(Chinese\)](#)

Question How does a deep learning algorithm using patient demographic information and longitudinal clinical notes to predict mortality risk perform as a proxy indicator for identifying patients with dementia who need palliative care?



University of Colorado
Anschutz Medical Campus

Center for Bioethics and Humanities

Concluding Thoughts

- AI has great potential for improving health, but it is critical that issues of bias and equity be at the center of AI – not an afterthought.
- We will not fix AI biases by fixing the algorithm – it is yet another opportunity and motivation to address by structural and social issues of power, class, and race.
- As AI moves forward, we must continue to advocate for decisions and decision processes that both account for issues of equity and include marginalized communities and people.