Data Ethics in Biomedicine

Mid-range considerations

Stephanie Holmquist
slh12@Columbia.edu
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Setting 1

- The AI narrative in medicine is part of a bigger one
- Narrative of becoming the next kind of human being in the context of data driven cultures, market economies, self-conceptions (seems quite certain)
- Surveillance societies, whether by fiat from the center or by the logic of the market paired with perceived individual needs (Alexa, order me dinner.)

Setting 2

- Datasets are in and of themselves inert repositories of mute digital or physical materials. Traditional analytic and statistical tools, plus increasingly computationally sophisticated algorithms transform datasets into sources for choice and recommendation in clinical and public health medicine.
- A cluster of related ethical questions emerge from AI-based advances in dataset mining. These include the need for “explainable AI”, accounting for machine and human bias in dataset pattern recognition, and a reconsideration of ownership and privacy of personal bio-medical information.

Setting 3

Because we are constantly leaving data trails, without knowing who is harvesting and storing from them, then... (add several steps)…

Data…

Logic of a move to a more public health/research ethics style framework alongside the traditional bioethical individual orientation

- Privacy (?)
- Anti-discrimination
- Procedural and substantive fairness

- Framework of a right to science
- Right to recognition—”moral and material rights”

UDHR 27(1) (1948)

We have reasons to be mistrustful

NHS and Alphabet

- Highs of achievement
- NHS/Moorfields Hospital + Deep Mind
- One 3D scan→ 50 diagnosable eye diseases and conditions

De Fauw, Jeffrey, et al. "Clinically Applicable Deep Learning for Diagnosis and Referral in Retinal Disease." *Nature Medicine 24.9 (2018): 1342-50. Print.*

Lows of mistrust

- Reveals demographic characteristics
- Accessory findings—AD?
- Uncontrolled monetization
- Opaque decision-making
- Unaccountable policy reversal

Unsuccessful ML applications

Facial recognition algorithms made by Microsoft, IBM and Face++ were more likely to misidentify the gender of black women than white men.

Gender was misidentified in **up to 1 percent of lighter-skinned males** in a
385 photos.

Gender was misidentified in **35 percent of darker-skinned females** in a set of 271 photos.

Buolamwini, Joy and Gebru, Timnit Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification
Proceedings of Machine Learning Research 81:1–15, 2018; Conference on Fairness, Accountability, and Transparency

https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

Successful ML application

The researchers had access to data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), a major multi-site study focused on clinical trials to improve prevention and treatment of this disease. The ADNI dataset included more than 2,100 FDG-PET brain images from 1,002 patients. Researchers trained the deep learning algorithm on 90 percent of the dataset and then tested it on the remaining 10 percent of the dataset. Through deep learning, the algorithm was able to teach itself metabolic patterns that corresponded to Alzheimer's disease.
Finally, the researchers tested the algorithm on an independent set of 40 imaging exams from 40 patients that it had never studied. The algorithm achieved 100 percent sensitivity at detecting the disease an average of more than six years prior to the final diagnosis.
"We were very pleased with the algorithm's performance," Sohn said. "It was able to predict every single case that advanced to Alzheimer's disease."

Academic (UCSF)

Ding, Yiming, et al. "A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18f-Fdg Pet of the Brain." *Radiology* (2018): 180958. Print. 10.1148/radiol.2018180958

Successful ML application

The researchers had access to data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI), a major multi-site study focused on clinical trials to improve prevention and treatment of this disease. The ADNI dataset included more than 2,100 FDG-PET brain images from 1,002 patients. Researchers trained the deep learning algorithm on 90 percent of the dataset and then tested it on the remaining 10 percent of the dataset. Through deep learning, the algorithm was able to teach itself metabolic patterns that corresponded to Alzheimer’s disease.

Finally, the researchers tested the algorithm on an independent set of 40 imaging exams from 40 patients that it had never studied. The algorithm achieved 100 percent sensitivity at detecting the disease an average of more than six years prior to the final diagnosis.

“We were very pleased with the algorithm’s performance,” Sohn said. “It was able to predict every single case that advanced to Alzheimer’s disease.”

TRUST WORTHY?

Academic (UCSF)

Ding, Yiming, et al. "A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18f-Fdg Pet of the Brain." *Radiology* (2018): 180958. Print. 10.1148/radiol.2018180958

Reasonable questions

- Are the algorithm and other computational processes explainable?
  (Auditable, correlation versus causation)

- Does the data set have integrity?

- Algorithm + data set = biased output

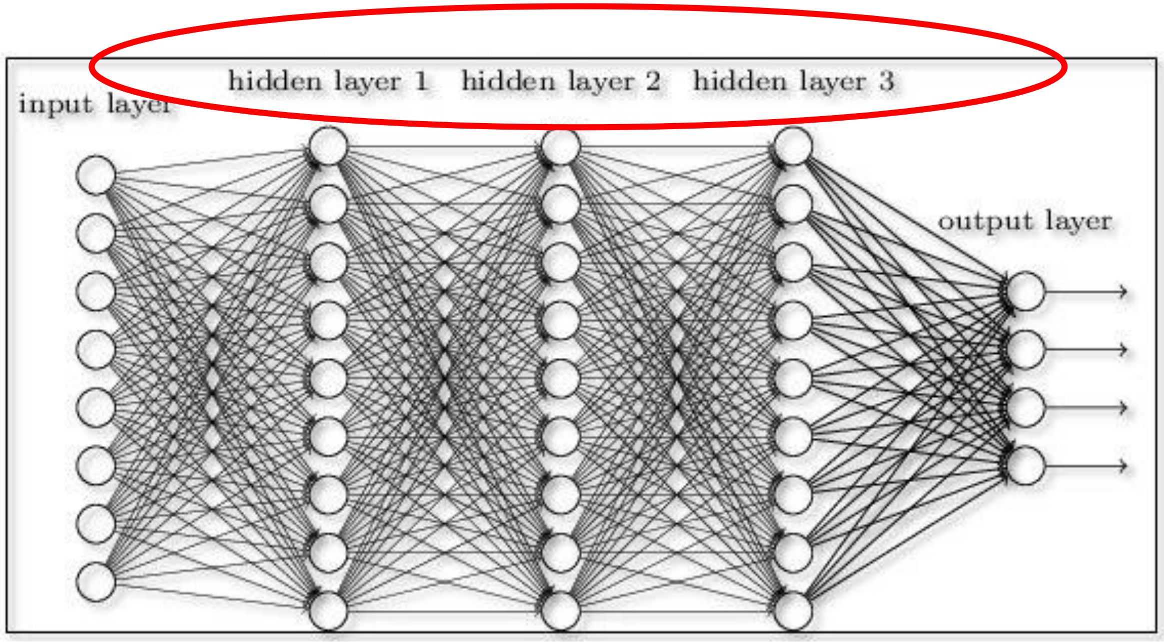
Over-testing and over-medicalization

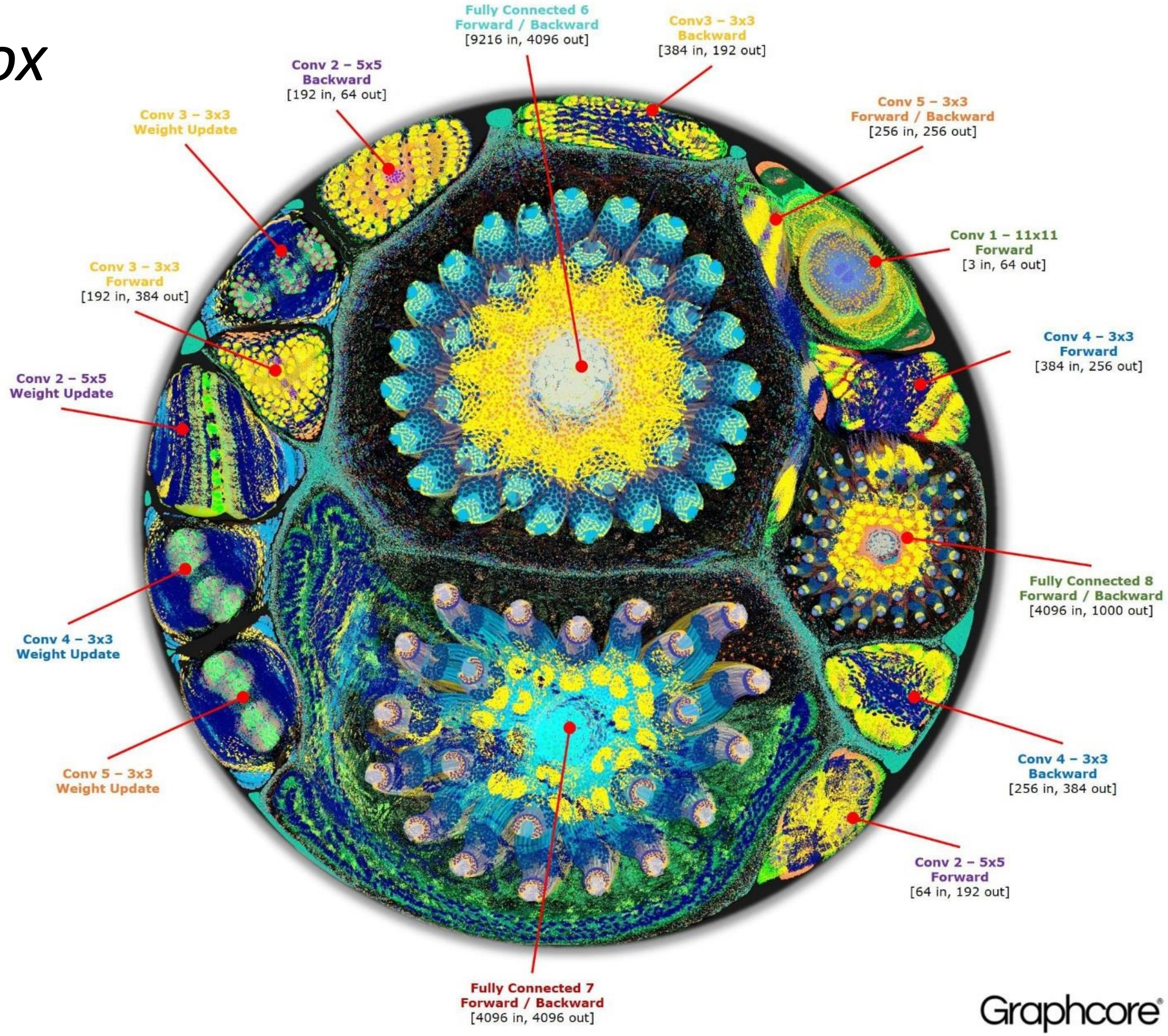
Complexity and trust—the black box

```
data what;
set weather;
if  city ='Halifax' and weather='rain'
then whattotake="Sigh! I knew it, I'll take an umbrella";
else if city ='San Antonio' and weather='rain'
then whattotake="No way! ok I'll take an umbrella";
flag='SAS user group conferences 3-8 november 2016';
run;

NOTE: There were 8 observations read from the data set WEATHER.
NOTE: The data set WORK.WHAT has 8 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds
```

Complexity and trust—the black box
Non-linearity





Explainable AI

Does the GDPR provide a right to and explanation for AI/ML?

Wachter, Sandra and Mittelstadt, Brent and Floridi, Luciano,

Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation

(December 28, 2016). International Data Privacy Law, 2017. Available at SSRN: https://ssrn.com/abstract=2903469 or http://dx.doi.org/10.2139/ssrn.2903469

Selbst, Andrew D. and Powles, Julia, Meaningful Information and the Right to Explanation

Articles 13-15 provide rights to “meaningful information about the logic involved” in automated decisions. This is a right to explanation, whether one uses the phrase or not.

(November 27, 2017). International Data Privacy Law, vol. 7(4), 233-242 (2017). Available at SSRN: https://ssrn.com/abstract=3039125

FDA 510(k) Clearance
The AI as a device

Explainable AI

We value explanation for practical reasons—evaluative tool for safety, efficacy, robustness, fair economic assessment

We value explanation for deontic reasons—showing respect for patients

Some valuable things should be distributed fairly, in some non-procedural sense—?
Some valuable things should be distributed with procedural fairness

In medicine current best use cases are decision making-support modules with allocational implications

Transformed into the patient or not, triaged as the priority or not, etc.

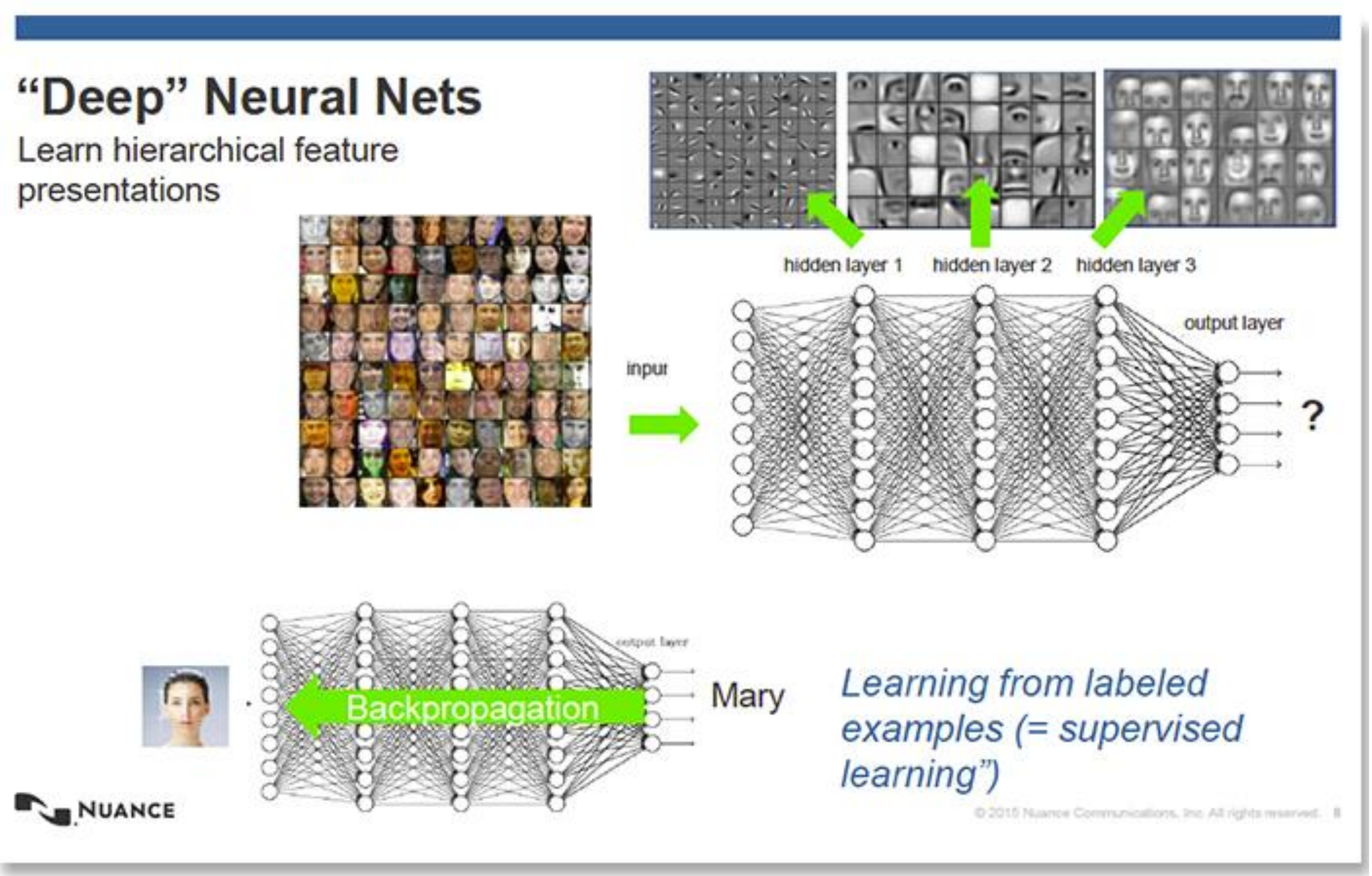
Explainable AI

- Centralized locus of error
  - to AIs (data scientists) from many HCPs making medical errors
  - however AIs are better learners—number of HCPs who are chronic error makers and not improving their practices
  - IRL algorithmic learning
- Some irreducible uncertainty about safety from black boxness—is this different because one AI influences and entire cohort of physicians and patients?
  - Absence of end points for safety and effectiveness
  - Algorithms that design algorithms
  - Undetectable, widespread error

Explainable AI—explanatory strategies

- Visualization
- Visualization with weighting for most determinative pixels
- Context and user relevant explanations supports inference to a "salient purpose"
  (Rune Nyrup)

Instances when contextual and functional explanations are not good enough?
Uber accident?



The structure and training of deep neural networks. Image: Nuance

Explainable AI—explanatory strategies

Moorfield’s retinal imaging AI
Allows a sophisticated assessor some insight into a diagnostic and decision support system

Network 1
Familiar representation—but highly processed
- map of the different types of eye tissue and the features of disease it sees, such as haemorrhages, lesions, irregular fluid or other symptoms of eye disease.
- Insight into the system’s rational

Network 1
- “classification network” analyses the map
- diagnoses and a referral recommendation.
- recommendation as a percentage, allowing clinicians to assess the system’s confidence in its analysis.

Explainable AI

Fully causal explanation

- Non-AI examples
  - US Medicaid home care
- AI as expert systems—old-style credit risk assessment/COMPAS

Less supervised/unsupervised/Deep learning

- Black box is not designed, not coached, non-cognitive logic of pattern recognition and correlation
- Error to as for “explanation” as step-wise or network propagated series of events that reliably produce XXX output from YYY inputs
- Pragmatic explanation

Explainable AI

Kinds of explanation

Ex-post explanations as

- General system functions
- For a given output/decision

Trust as a substitute for explanation—familiar

Practices and institutions of trust

Is explanation necessary for “intelligent trust”; “proxy evidence of trustworthiness” for patients, HCP, HC systems

Alzheimer’s Disease Neuroimaging Initiative—the account of the inputs, procedures—training and testing, actors

Give us reason to adopt an attitude of trust, just so long as...

Reasonable questions

- Are the algorithm and other computational processes explainable? (Auditable, correlation versus causation)

- Does the data set have integrity?

- Is algorithm + data set = biased output

(Over-testing and over-medicalization)

Data set integrity

- Translation from data capture, which may be well structured for one purpose, but not for big data analytics—esp. decision support, predictive analytics, drug development
- (Unstructured text, interoperability, cooperation across national borders.)
- 3Vs-volume, velocity, variety
- More v’s…
- V=veracity

An ethically salient practical goal for BD is leveraging to provide “big evidence”. (Patrick Ryan)

- Biomedical projects may collect as much as six terrabytes about a one patient
- Flatiron’s longitudinal tracking of lung cancer patients
- Nature article

Data set integrity/RWE

Real-World Evidence — What Is It and What Can It Tell Us?

Rachel E. Sherman, M.D., M.P.H., Steven A. Anderson, Ph.D., M.P.P., Gerald J. Dal Pan, M.D., M.H.S., Gerry W. Gray, Ph.D., Thomas Gross, M.D., M.P.H., Nina L. Hunter, Ph.D., Lisa LaVange, Ph.D., Danica Marinac-Dabic, M.D., Ph.D., Peter W. Marks, M.D., Ph.D., Melissa A. Robb, B.S.N., M.S., Jeffrey Shuren, M.D., J.D., Robert Temple, M.D., Janet Woodcock, M.D., Lilly Q. Yue, Ph.D., and Robert M. Califf, M.D.

The term “real-world evidence” is widely used by those who develop medical products or who study, deliver, or pay for health care, but its specific meaning is elusive. We believe it refers to information on health care that is derived from multiple sources outside typical clinical research settings, including electronic health records (EHRs), claims and billing data, product and dis- shortage of researchers with adequate methodologic savvy could result in poorly conceived study and analytic designs that generate incorrect or unreliable conclusions. Accordingly, if we are to realize the full promise of such evidence, we must be clear about what it is and how it can be used most effectively, and we must have appropriate expectations about what it can tell us.

Sherman, Rachel E; Anderson, Steven A; Dal Pan Gerald J; Gray, Gerry W; Gross, Thomas; et al. **The New England Journal of Medicine; Boston** Vol. 375, Iss. 23, (Dec 8, 2016): 2293-2297.
DOI:10.1056/NEJMsb1609216

Data set Integrity//Strategy for shifting the social compact

Trope is: "Use my data"—promoted by pharma and tech entrants into the big data health space, including…

- Pfizer/Flatiron
- Alphabet/Deep Mind
- GSK+23&me+Amazon-Alexa
- Tencent

- Large scale benefits from "total population enrollment"—specific conditions to total surveillance
- Move from the 3% to the 97% of the affected population
- Move from the snapshot, RCT-end-point-oriented, to the longitudinal, increasingly dense data narrative for v. large cohorts
- Move from post-marketing surveillance to population scale tracking of drug and device and efficacy over time

Data set integrity

Biomedical projects may collect as much as six terabytes about one patient

Multiply useful

- To the patient
- Healthcare system
  - Population characteristics and trends
  - Allocation
- RWE applications
  - Treatment protocols—in real time, who is benefiting, how are new interventions behaving (clinical phenotype and genotype)
  - Drug development
  - Post marketing multivigilance

Expensive—initial classification by trained actual people

Ethical pressure derives from intensity of potential benefit

Data set Integrity/Novel, novel sources

- Social media scraping
- Health-related connected devices
- Connected device interactions
- Real world activity
  - Public space surveillance
  - General daily life patterns—travel, shopping, transport
  - Voice analysis (insurance)

“Creative” in and outside of medicine/Deep Learning software is cheap or opensource

Bias

- We understand traditional sources of bias in HC datasets fairly well
- Sample sizes that are too small—race, uncommon disease subtypes
- AI will introduce new forms of novel bias
- Our best resources in biomedicine for bias awareness and de-biasing strategies is from hiring selection

There will be algorithmic de-biasing in medicine

Given the physician-AI partnership model, should patients be entitled to an only AI diagnostic and treatment assessment or a comparative one?

YES

Privacy, data control

Many models

- Ownership—highly controlled
- Divided ownership—land versus mineral rights
- National asset
- Control
  - Data generator—us
  - Entity that adds value
  - Entity that adds patient benefit now or later
- Process oriented control
- Information and awareness and not good tools
- Social guard rails

Privacy/Self as data set

- Do we increasingly see our selves as complex data sets
- We prefer depersonalized interactions—preference for texting over voice phoning
- Rob—comfort may be the care—behaviorally related care
- Derrick—Can the technical aspect of healthcare be clearly separated from the affective-relational aspect?
- Patients may opt out of relational healthcare
  - Practical reasons—ATM→”XXXpay”
  - Their affective style
  - Move to distributed healthcare—from the clinic to anywhere

Again

Because we are constantly leaving data trails, without knowing who is harvesting and storing from them, then... (add several steps)…

Data…

Logic of a move to a more public health/research ethics style framework alongside the traditional bioethical individual orientation

- Privacy (?)
- Anti-discrimination
- Procedural and substantive fairness

- Framework of a right to science
- Right to recognition—”moral and material rights”

UDHR 27(1) (1948)