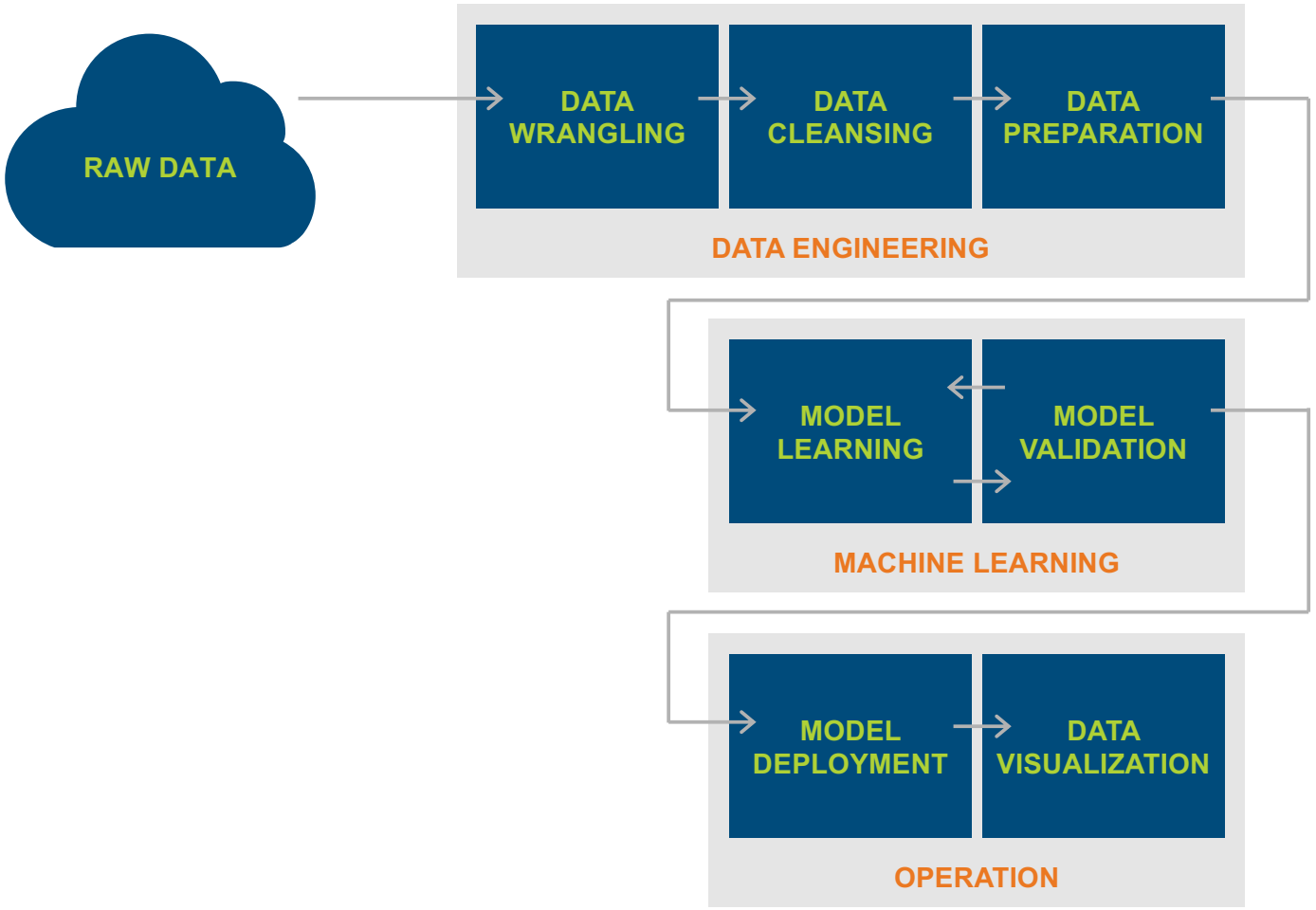
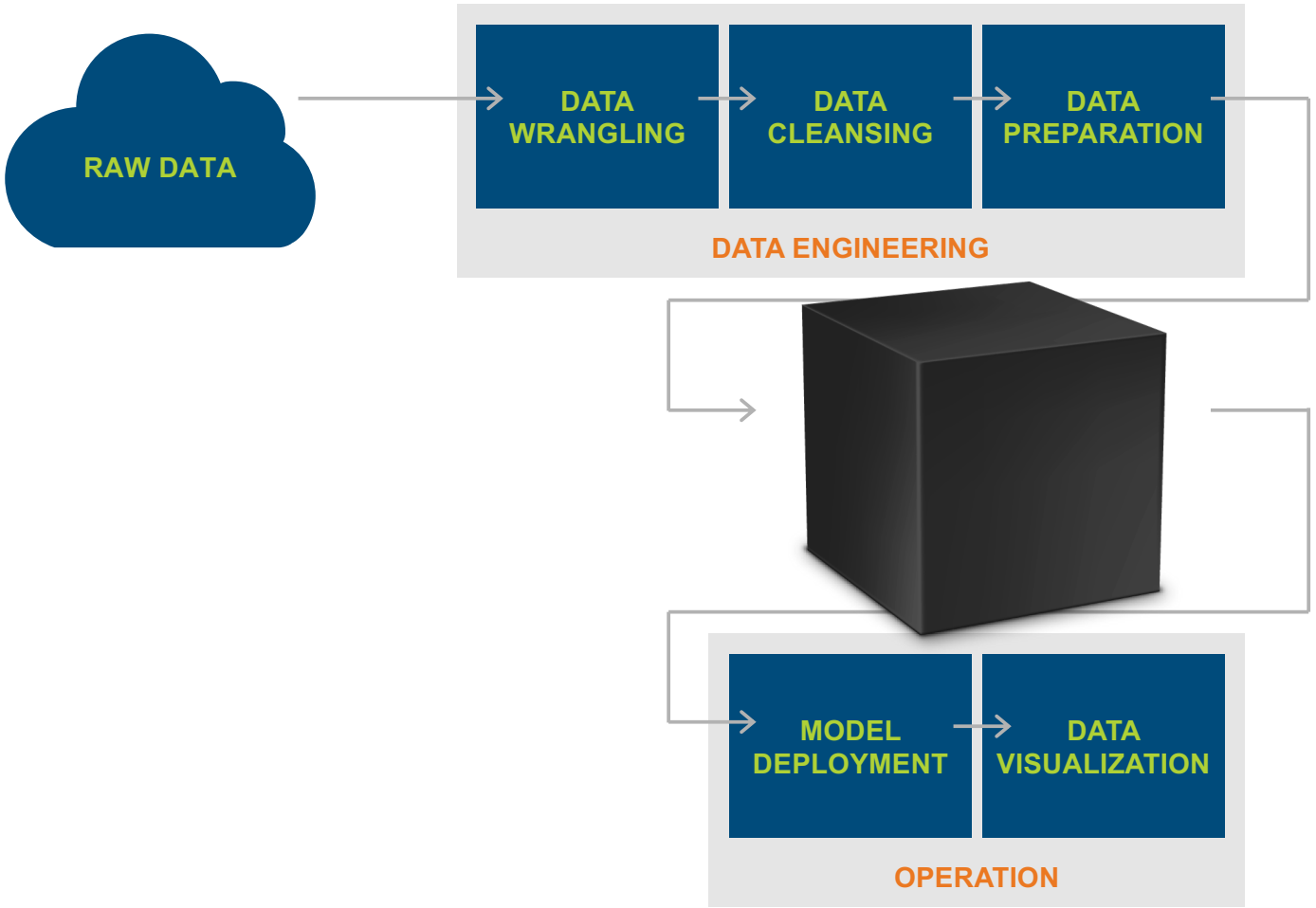
RESPONSIBLE DATA SCIENCE

IEEE Big Data 2019 Keynote | December 10, 2019

Lise Getoor | UC Santa Cruz


maintainable
reliable
open
data ownership
accountable
validated
data rights
justice
security
safe
robust
bias
fairness
transparency
privacy
equitable
usability
understandable
ethics
trust
interpretable


RAW DATA
DATA WRANGLING
DATA CLEANSING
DATA PREPARATION
DATA ENGINEERING
MODEL LEARNING
MODEL VALIDATION
MACHINE LEARNING
MODEL DEPLOYMENT
DATA VISUALIZATION
OPERATION

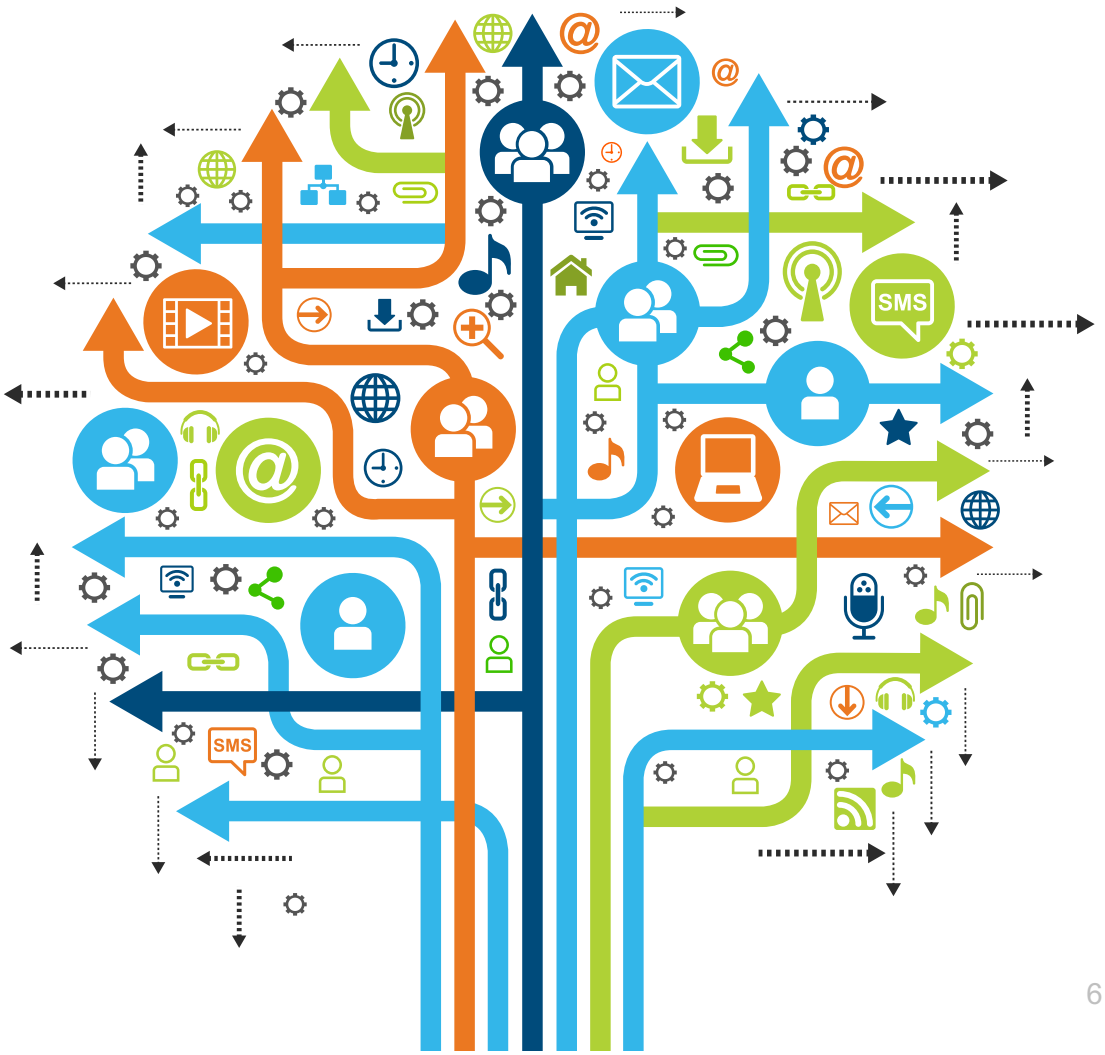

RAW DATA
DATA WRANGLING
DATA CLEANSING
DATA PREPARATION
DATA ENGINEERING
MODEL DEPLOYMENT
DATA VISUALIZATION
OPERATION

encourage some healthy skepticism!

and some curiosity…

L.Zielonko

SOCIO-TECHNICAL SYSTEMS



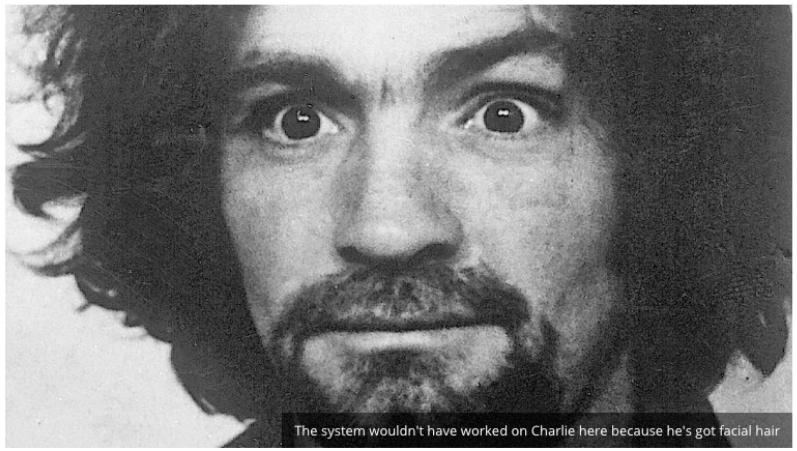

UNPACK
GOALS
EXCITE
INCITE

Data Science

IN THE NEWS

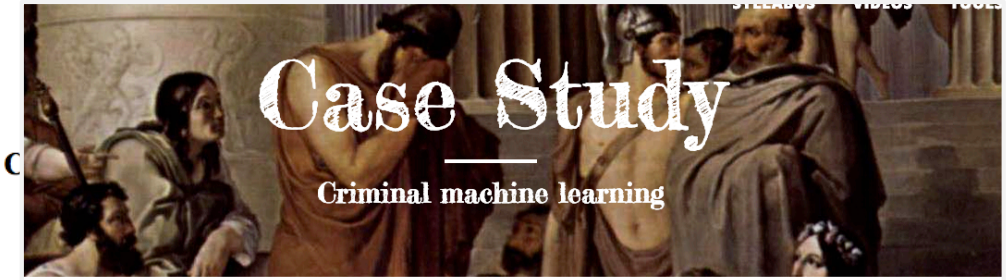
AI supposedly learns to identify criminals by their faces, takes us back to the 19th century

By Vlad Dudau · Nov 25, 2016 07:42 EST · HOT! 23



The system wouldn't have worked on Charlie here because he's got facial hair

Measuring the worth or character of a person by their physical traits is generally thought to be a sad staple of the quasi-scientific 19th and 20th centuries. But some of these ideas seem to be making a comeback, apparently dressed in full scientific garb, and accessorizing in trendy buzzphrases like machine-learning and artificial intelligence. That's the case with a new "scientific" paper, currently waiting to be published, which reports that researchers in China managed to teach a machine to identify criminals just by looking at their faces.



Criminal machine learning

For those who prefer video, this case study is described in the April 26th lecture of our Spring 2017 course.

In November of 2016, engineering researchers Xiaolin Wu and Xi Zhang posted an article entitled "Automated Inference on Criminality using Face Images" to a widely used online repository of research papers known as the arXiv. In their article, Wu and Zhang explore the use of machine learning to detect features of the human face that are associated with "criminality"—and they claim to have developed algorithms that can use a simple headshot to distinguish criminals from non-criminals with high accuracy. If this strikes you as frighteningly close to Philip K. Dick's notion of pre-crime, the film Minority Report, and other dystopian science fiction, you're not alone. The media thought so, too. A number of technology-focused press outlets [1,2,3] picked up on the story and explored the ethical implications.

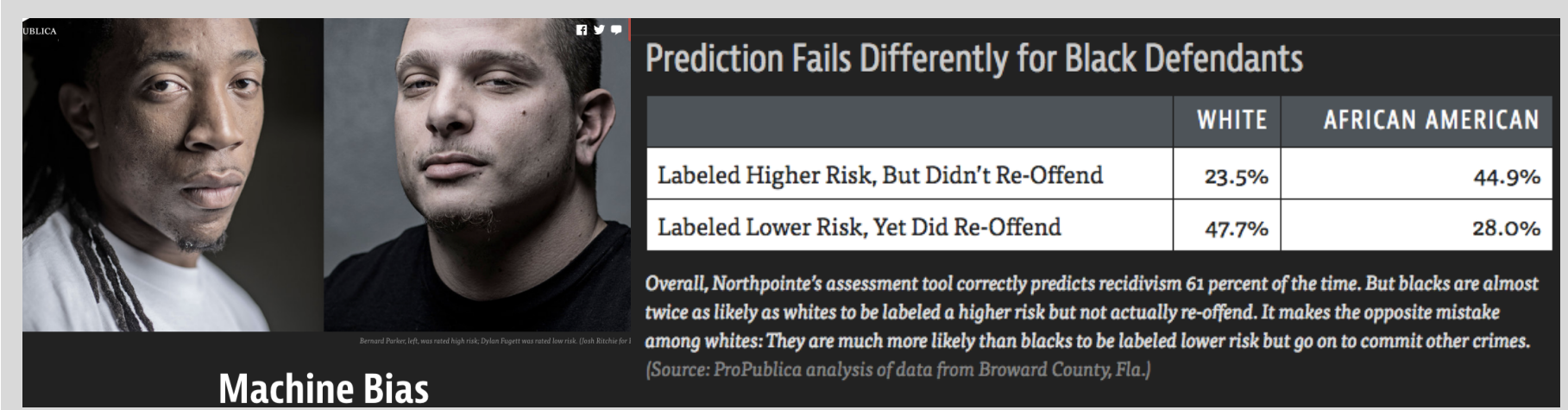
If one could really detect criminality from the structure of a person's face, we

RECIDIVISM RISK

defendant's likelihood of committing a crime

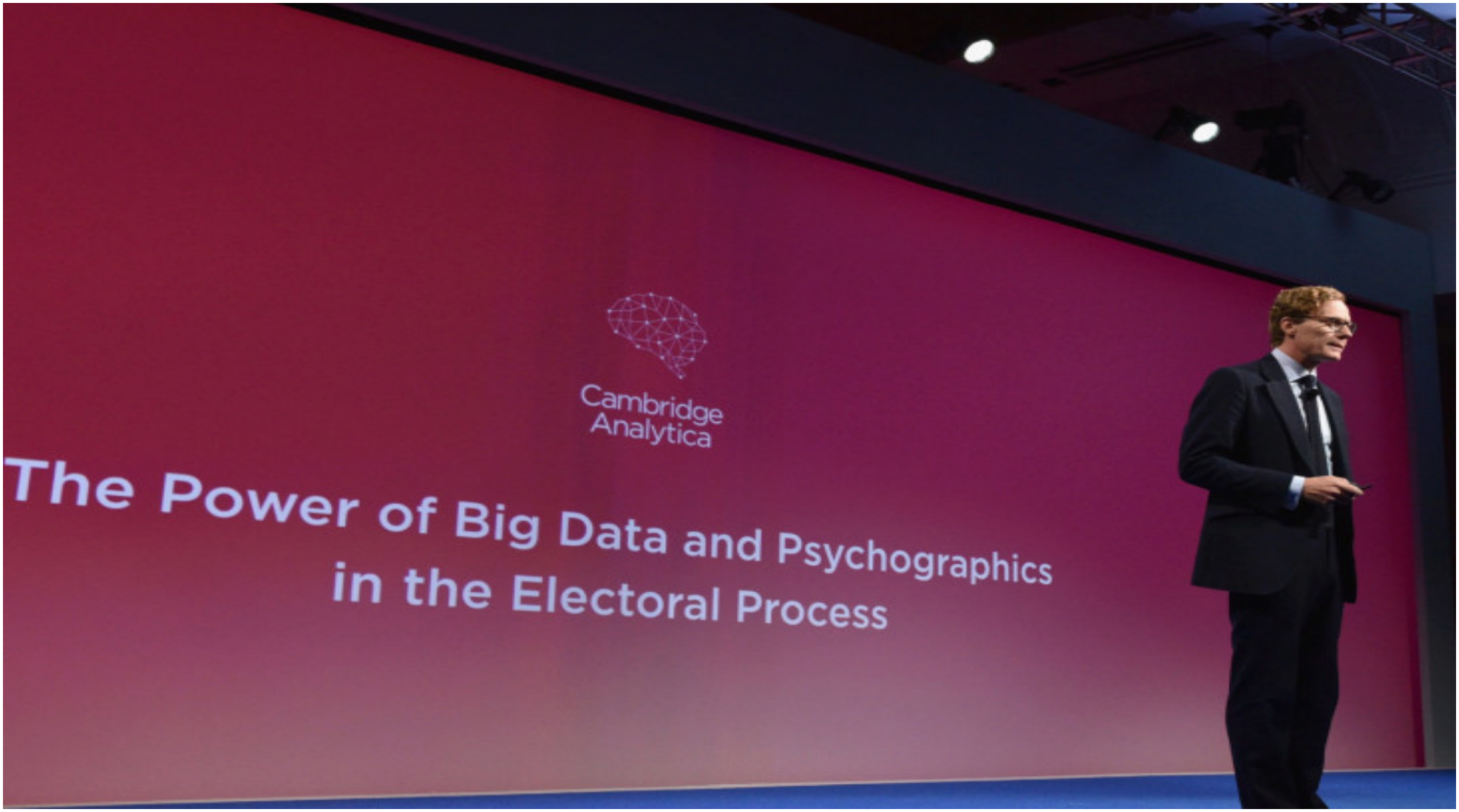
USED THROUGHOUT CRIMINAL JUSTICE SYSTEM

pretrial, bail and sentencing



Overpredicts recidivism for African Americans; underpredicts recidivism for whites

Propublica, May 2016



The Great Hack, Netflix Documentary

Image Credits: Bryan Bedder / Getty Images / Getty Images

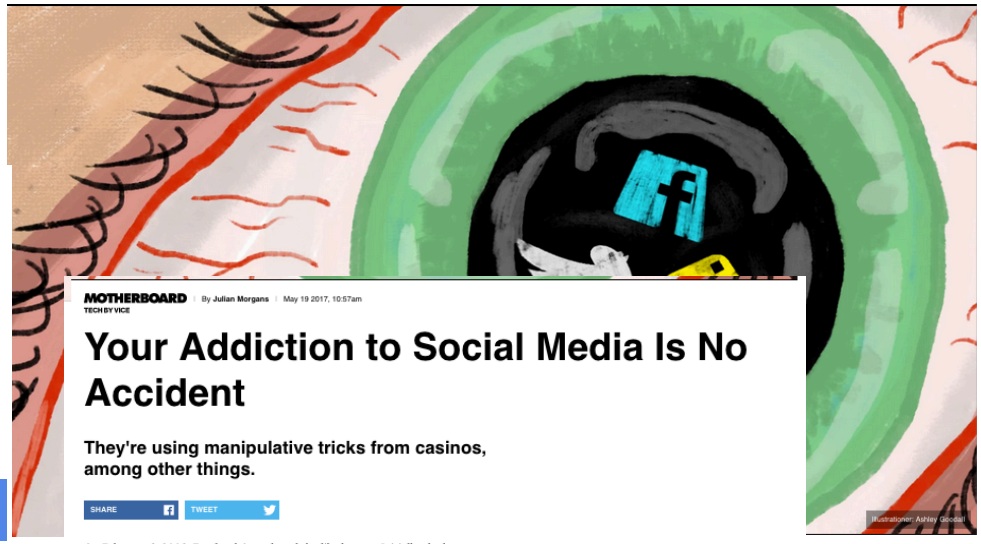
Technology Designed for Addiction

What are the dangers of digital feedback loops?

Posted Jan 04, 2018

One of the things that Wade L. Robinson discusses in his book, *Engineering Ethics*, is the importance of avoiding "error-provocative" designs, in which the technological artifact not only allows for the possibility of human fallibility but actively steers the user into the direction of harm. He spends a great deal of time discussing stove knobs, for example. I imagine most of us have had the frustration of thinking that we are activating one burner of the stove, say the left front, only to find that another burner, the back right, has actually heated. It turns out that every stove manufacturer has a slightly different way of aligning the knobs to the burners, and this design problem is surprisingly complex. Most of the time, the mismatch between knobs and burners simply causes frustration, but it can cause
it only takes a minute or two for a house fire to rea

Psychology Today



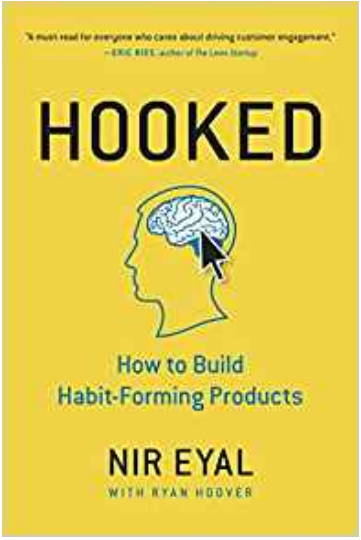
MOTHERBOARD | By Julian Morgans | May 19 2017, 10:57am
TECH BY VICE

Your Addiction to Social Media Is No Accident

They're using manipulative tricks from casinos, among other things.

SHARE TWEET

On February 9, 2009, Facebook introduced the like button. Initially, the button

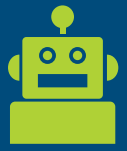


Wikimedia

ROADMAP

MACHINE LEARNING

RELATIONAL LEARNING



ARTIFICIAL INTELLIGENCE

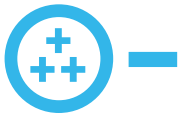
ROADMAP

MACHINE LEARNING

RELATIONAL LEARNING

ARTIFICIAL INTELLIGENCE

ML TYPES



CONCEPT LEARNING

learn a logically consistent hypothesis

STATISTICAL LEARNING

learn a hypothesis that maximizes probability

OPTIMIZATION-BASED

hypothesis minimizes some loss function

NEURAL-INSPIRED LEARNING

represent hypothesis as a neural network

ML OPTIMIZATION

Definition: Given data D = {x_1, x_2, …, x_n}, target labels L = {y_1, y_2, …, y_n} find a hypothesis *H* such that

$$\underset{H}{\operatorname{argmin}} \sum_{x_i \in D} \ell(H(x_i), y_i)$$

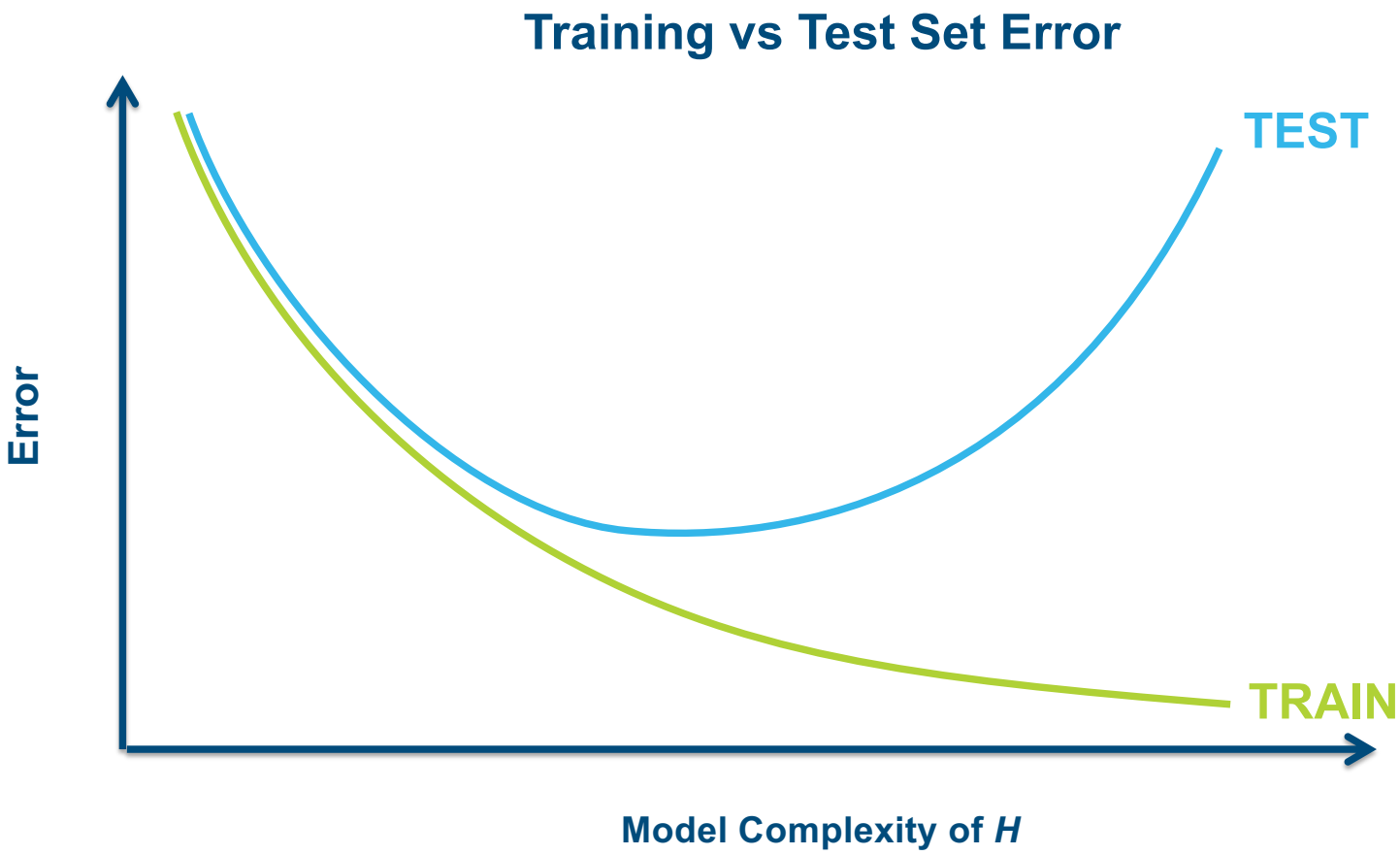
where ℓ is some loss function.

ESTIMATING ERROR

Definition: Given test data T = {x_1, x_2, …, x_m}, test labels L' = {y_1, y_2, …, y_m} the error is:

$$\sum_{x_i \in T} \ell(H(x_i), y_i)$$

where ℓ is some loss function.


Training vs Test Set Error
TEST
TRAIN
Error
Model Complexity of H

WHAT CAN GO WRONG?

#1

FORMALIZATION

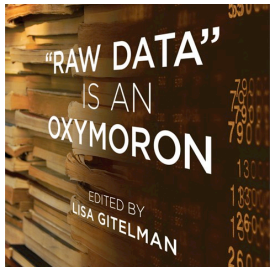
FRAME OF REFERENCE

Definition: Given data $D = \{x_1, x_2, \ldots, x_n\}$, target labels $L = \{y_1, y_2, \ldots, y_n\}$ find a hypothesis *H* such that

$$\underset{H}{\operatorname{argmin}} \sum_{x_i \in D} \ell(H(x_i), y_i)$$

where ℓ is some loss function.

ABSTRACTIONS

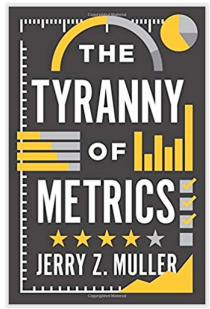


DATA

Transformation of raw data into feature vectors determines **u**nit of analysis and quantification
Data (especially human data) occurs in context

LABELS

Discretization matters; number and boundaries affect results
Who defines the labels, external "expert", self-report?
Measurement model



LOSS

Interact in non-obvious, domain-specific ways
Penalize errors differently
Often trade-offs are simplified and baked into the loss function

All optimization problems are abstractions, none of them are "correct"

STRUCTURAL PLAUSIBILITY

Mapping from features to labels must pass the "smell test"

There should be plausible scientific connection between the input features and the output label

If there is not, **no matter how well your classifier performs,** you should reject the hypothesis

Automated Inference on Criminality using Face Images

Nov 2016

Abstract

We study, for the first time, automated inference on criminality based solely on still face images. Via supervised machine learning, we build four classifiers (logistic regression.

corner distance, and the so-called nose-mouth

Doesn't pass the plausibility test

Measuring the worth or character of a person by their physical traits is generally thought to be a sad staple of the quasi-scientific 19th and 20th centuries. But some of these ideas seem to be making a comeback, apparently dressed in full scientific garb, and accessorizing in trendy buzzphrases like machine-learning and artificial intelligence. That's the case with a new "scientific" paper, currently waiting to be published, which reports that researchers in China managed to teach a machine to identify criminals just by looking at their faces.

Criminal machine learning

For those who prefer video, this case study is described in the April 26th lecture of our Spring 2017 course.

ember of 2016, engineering researchers Xiaolin Wu and Xi Zhang posted entitled "Automated Inference on Criminality using Face Images" to a online repository of research papers known as the arXiv. In their and Zhang explore the use of machine learning to detect features of an face that are associated with "criminality"—and they claim to have oped algorithms that can use a simple headshot to distinguish criminals om non-criminals with high accuracy. If this strikes you as frighteningly close to Philip K. Dick's notion of pre-crime, the film Minority Report, and other dystopian science fiction, you're not alone. The media thought so, too. A number of technology-focused press outlets [1,2,3] picked up on the story and explored the ethical implications.

If one could really detect criminality from the structure of a person's face, we

#2

HIGH DIMENSIONAL DATA

BIG DATA

High dimensional data is problematic because danger of overfitting is significantly higher than low dimensional data

Curse of dimensionality: our intuitions break down in high dimensions

Likelihood of finding random subset of features that are predictive is high

Required **sample size** for generalization proportional to dimension

Multiple testing, p-value hacking, reproducibility crisis

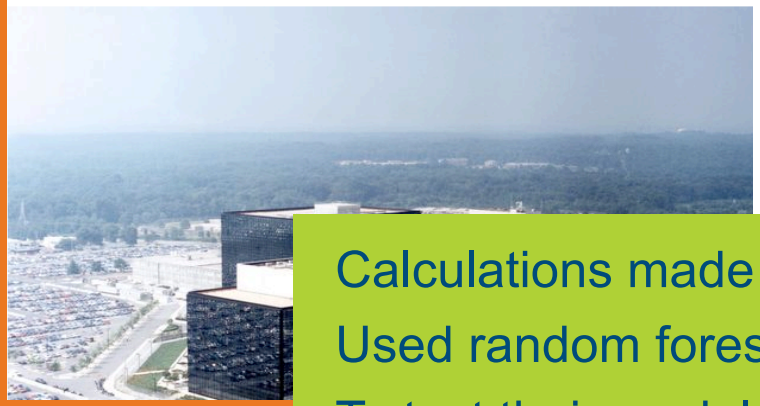
Models learned from high dimensional data treated with extreme care!

MILITARY

Can The NSA's Machines Recognize A Terrorist?

The big problem with little data

By Dave Gershgorn *February 16, 2016*



Calculations made on 80 variables for each cell phone user

Used random forests

To test their model, project uses data from just 7 known terrorists, plus random sample of 100,000 mobile phone users

When run it the wild, it identified an Al Jazeera reporter covering Al Qaeda as a potential terrorist

#3

MEASURING ERROR

MY ALGORITHM HAS AN ACCURACY OF 78.6%**

MY ALGORITHM HAS AN ACCURACY OF 78.6%**

ACCORDING TO OUT OF SAMPLE TEST WHICH ASSUMES

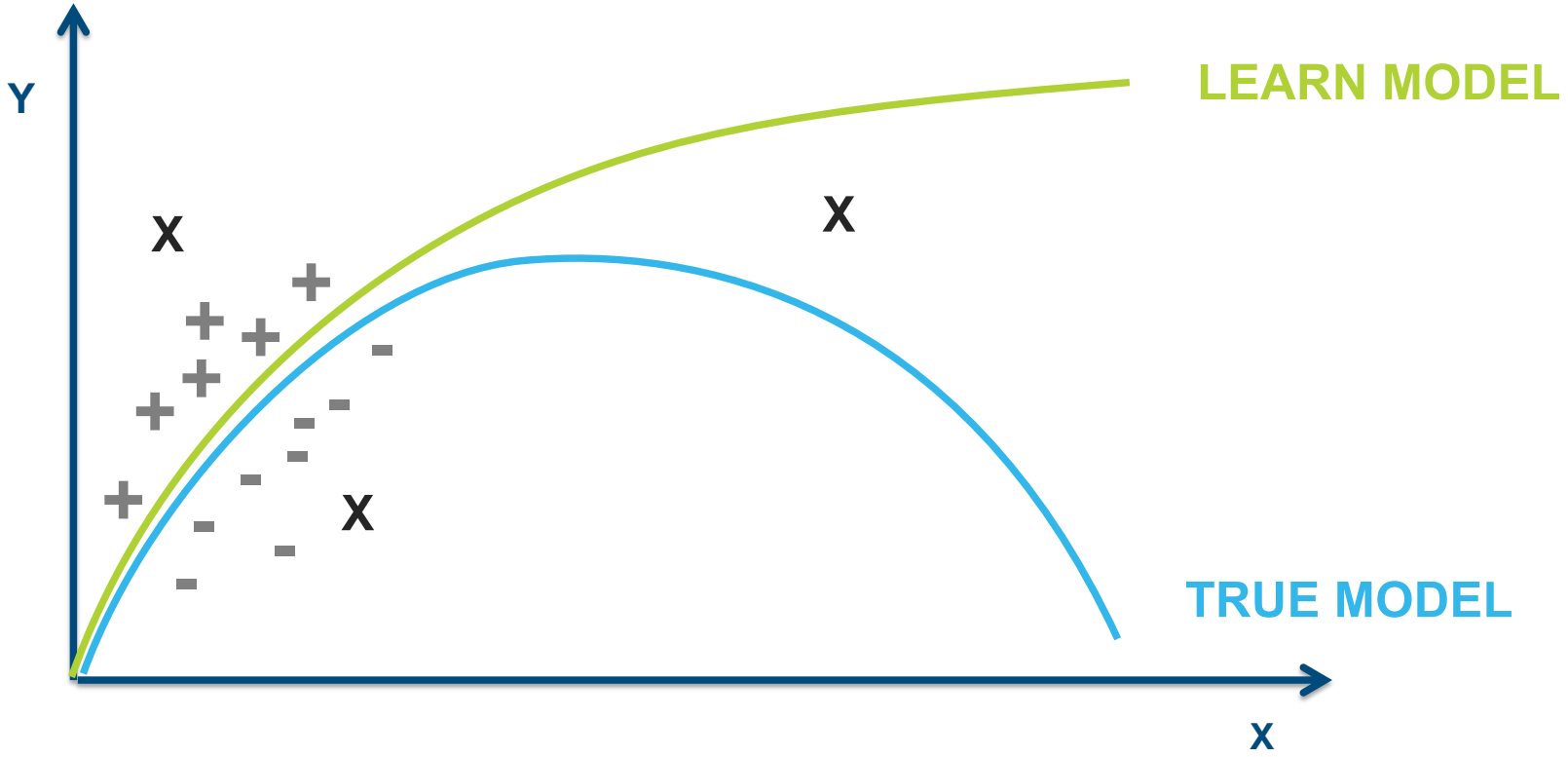
A well-defined population

The training data is a representative sample of the population

The test data is a representative sample of the population

These almost never hold in practice

NON-REPRESENTATIVE SAMPLES



MY ALGORITHM HAS AN ACCURACY OF 78.6%**

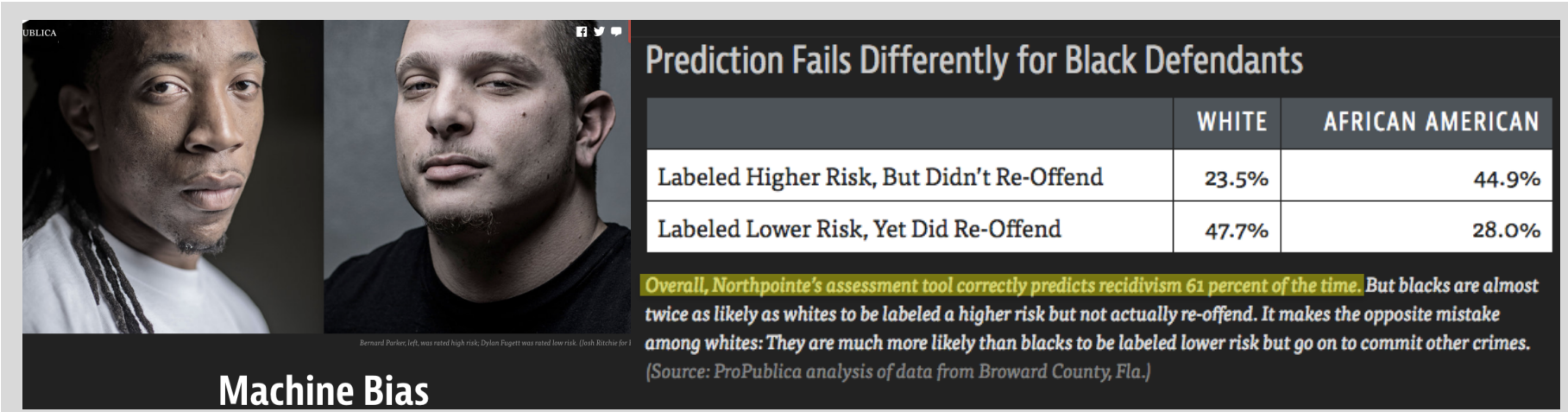
Doesn’t quantify model uncertainty!

RECIDIVISM RISK

defendant's likelihood of committing a crime

USED THROUGHOUT CRIMINAL JUSTICE SYSTEM

pretrial, bail and sentencing



ONLY 61% ACCURACY … that's really bad!

#4

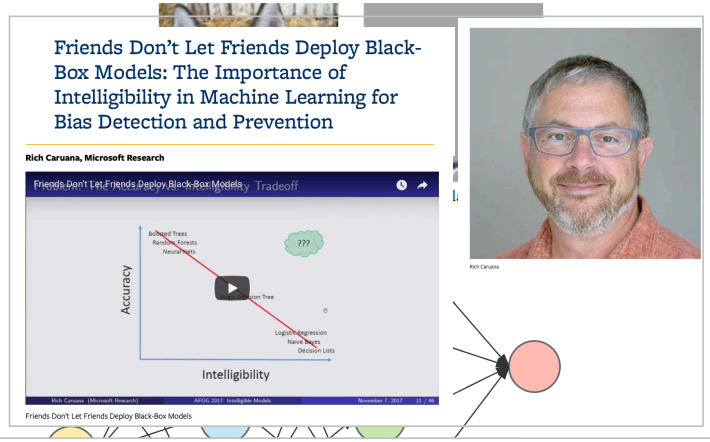
INTREPRETABILITY

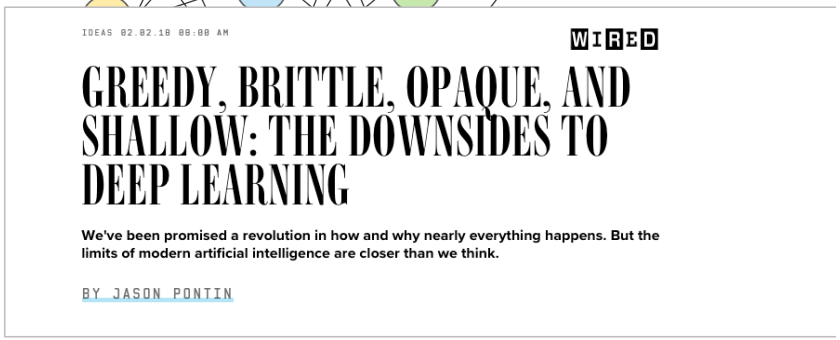
DEEP LEARNING

Deep learning, or **deep neural networks**, build a abstract, simplified neural network.

Given a bunch of observed input/output pairs, they are really good at constructing an abstract representation.

The models are black-box – it is difficult to interpret what the models are capturing.

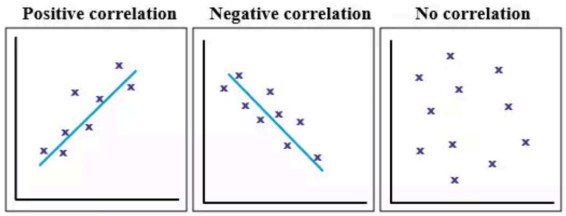




#5

CORRELATION vs. CAUSATION

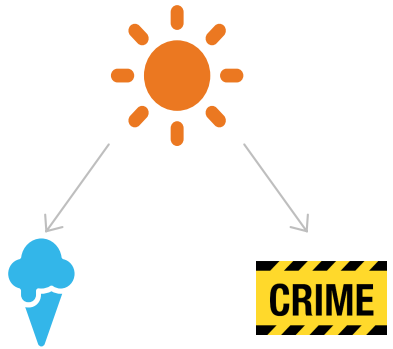
CORRELATION helps with prediction; if X and Y are positively correlated, then if observe high X expect to see high Y



are positively correlated

CAUSATION needed for decision making; if X and Y are causally connected, then if I manipulate value of X, keeping everything else constant, the value of Y will change.

CONFOUNDER
correlation is often due to confounding latent variable that is hidden cause of both X and Y.



CAUSAL QUERIES

?

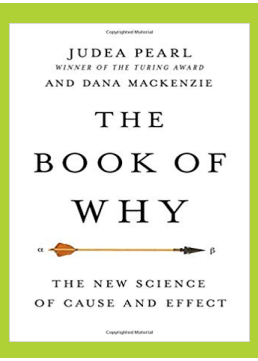
INTERVENTION:
What is the effect of treatment T on Y?

?

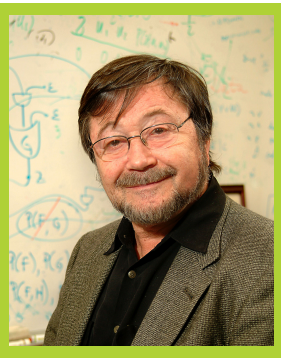
EXPLANATION:
Why did Y happen?

?

COUNTERFACTUAL:
What would happen if instead of X, ¬X?



Learn More This Afternoon!

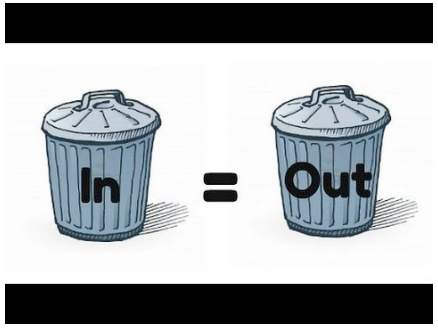


#6

BIAS

DATA BIAS

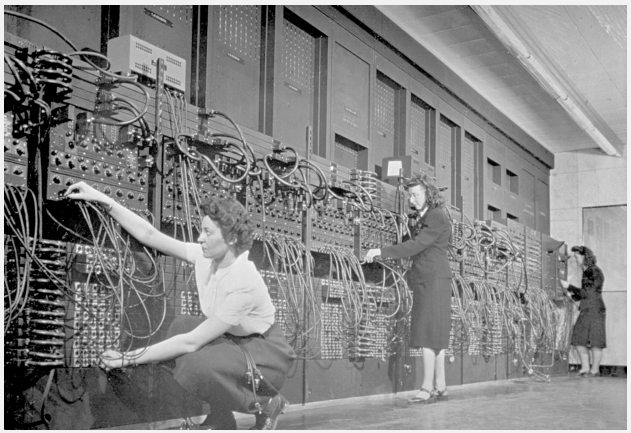
If the input to the system is biased,
due to selection bias, institutional bias, or societal bias,
then the output will be biased



Famous Computer Science phrase: GARBAGE IN, GARBAGE OUT

Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women

Isobel Asher Hamilton Oct. 10, 2018, 5:47 AM



Between 30 and 50 percent of programmers were women in the 1950s, and it was seen as natural career for them, as evidenced by a 1967 *Cosmopolitan* feature about "Computer Girls."

Automated Inference on Criminality using Face Images

Key to understanding problems with the paper is the training sets —1800 photos of Chinese men 18-55, no distinguishing facial hair, scars, or tattoos

1100 photos of non-criminals scraped from on the web
700 criminal ID photographs

Non-criminal images likely chosen by the subject to convey a positive impression

By contrast, for ID photos, it's likely that these were selected neither by the individual depicted, nor with aim of casting individual in a favorable light

☺ or ☹ ?

Thank goodness no one judges our character based upon our driver's license photos!

sons controlled for race, gender, age and facial expressions, nearly half of whom were convicted criminals, for discrimi-

dency of inferring innate traits and social attributes (e.g., the trustworthiness, dominance) of a person from his/her facial

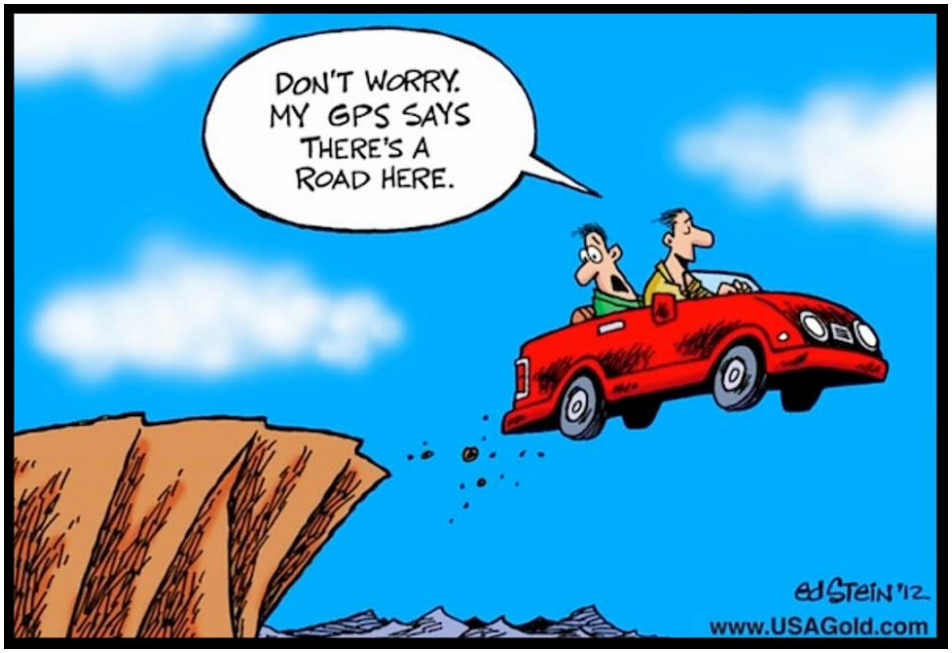
For those who prefer video, this case study is described in the April 26th lecture of our Spring 2017 course.

In November of 2016, engineering researchers Xiaolin Wu and Xi Zhang posted an article entitled "Automated Inference on Criminality using Face Images" to a widely used online repository of research papers known as the arXiv. In their article, Wu and Zhang explore the use of machine learning to detect features of the human face that are associated with "criminality"—and they claim to have developed algorithms that can use a simple headshot to distinguish criminals

AUTOMATION BIAS

AUTOMATION BIAS

Humans tend to favor suggestions from automated decision-making systems and often ignore contradictory information



ABDICATION & ACCOUNTABILITY

- Decision makers may abdicate decision responsibility to algorithms
- Especially for hard decisions, tempting to rely on an algorithm
- This affects risks and accountability

Andrew Grossman

ALGORITHMIC DISCRIMINATION

ALGORITHMIC DISCRIMINATION

COMPAS software has been used in over 2M cases since 2000

Algorithms can

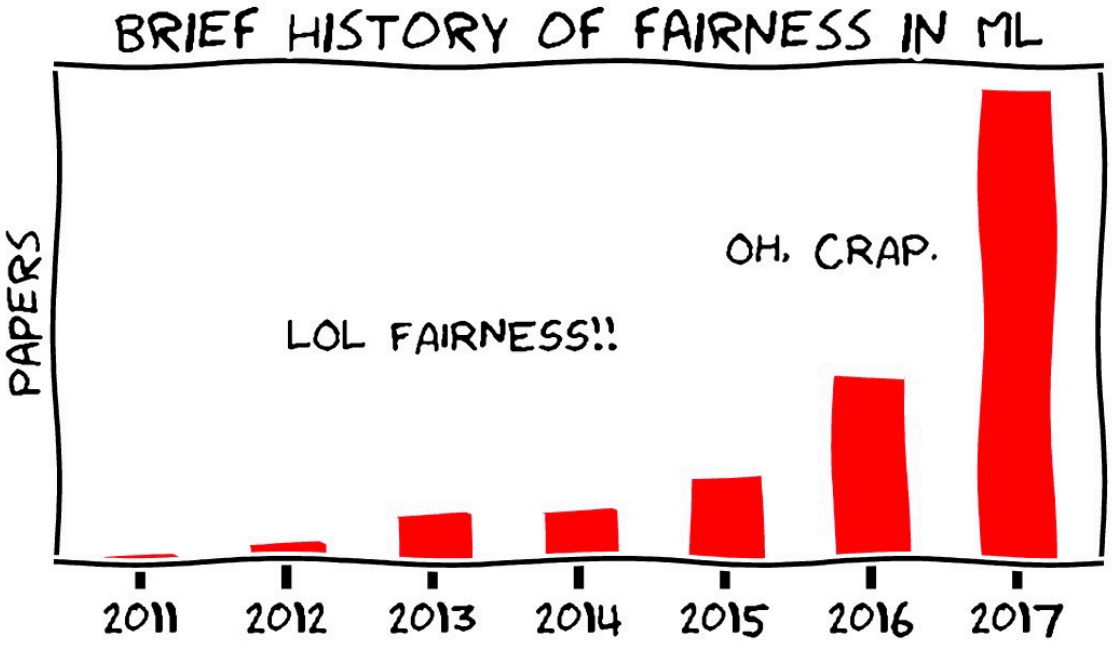
amplify

operationalize

legitimize

institutional bias

FAIRNESS



Very active area of research

From Barocas & Hardt, 2017

ACM Conference on Fairness, Accountability, and Transparency (ACM FAT*)

A multi-disciplinary conference that brings together researchers and practitioners interested in fairness, accountability, and transparency in socio-technical systems.

Algorithmic systems are being adopted in a growing number of contexts. Fueled by big data, these systems filter, sort, score, recommend, personalize, and otherwise shape human experiences of socio-technical systems. Although these systems bring myriad benefits, they also contain inherent risks, such as codifying and entrenching biases; reducing accountability and hindering due process; and increasing the information assymmetry between data producers and data holders.

ACM FAT* is an annual conference dedicating to bringing together a diverse community to investigate and tackle issues in this emerging area. Topics of interest include, but are not limited to:

- The theory and practice of fair and interpretable Machine Learning, Information Retrieval, NLP, and Computer Vision
- Measurement and auditing of deployed systems
- Users' experience of algorithms, and design interventions to empower users

MANY DEFINITIONS OF FAIRNESS

Tutorial: 21 fairness definitions and their politics

6,469 views

Arvind Narayanan
Published on Mar 1, 2018

Computer scientists and statisticians have devised numerous mathematical criteria to define what it means for a classifier or a model to be fair. The proliferation of these definitions represents an attempt to make technical sense of the complex, shifting social understanding of fairness. Thus, these definitions are laden with values and politics, and seemingly technical discussions about mathematical definitions in fact implicate weighty normative questions. A core component of these technical discussions has been the discovery of trade-offs between different (mathematical) notions of fairness; these trade-offs deserve attention beyond the technical community.

This tutorial has two goals. The first is to explain the technical definitions. In doing so, I will aim to make explicit the values embedded in each of them. This will help policymakers and others better understand what is truly at stake in debates about fairness criteria (such as individual fairness versus group fairness, or statistical parity versus error-rate equality). It will also help computer scientists recognize that the proliferation of definitions is to be celebrated, not shunned, and that the search for one true definition is not a fruitful direction, as technical considerations cannot

Definition	Citation #
Group fairness or statistical parity	208
Conditional statistical parity	29
Predictive parity	57
False positive error rate balance	57
False negative error rate balance	57
Equalised odds	106
Conditional use accuracy equality	18
Overall accuracy equality	18
Treatment equality	18
Test-fairness or calibration	57
Well calibration	81
Balance for positive class	81
Balance for negative class	81
Causal discrimination	1
Fairness through unawareness	14
Fairness through awareness	208
Counterfactual fairness	14
No unresolved discrimination	14
No proxy discrimination	14
Fair inference	6

Verma, Sahil, and Julia Rubin. "Fairness definitions explained."
IEEE/ACM International Workshop on Software Fairness (FairWare). IEEE, 2018.

GROUP COMPARISONS

GROUP A

GROUP B

Is classification algorithm C fair to both groups?

STATISTICAL CONCEPTS

DEMOGRAPHIC PARITY (INDEPENDENCE):

$P(C = 1 \mid G = A) = P(C = 1 \mid G = B)$

EQUALIZED ODDS:

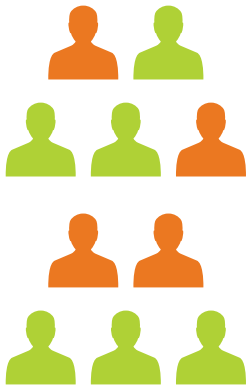
$P(C = 1 \mid G = A, Y = y) = P(C = 1 \mid G = B, Y = y), y \in \{0, 1\}$

PREDICTIVE PARITY (PRECISION, CALIBRATION):

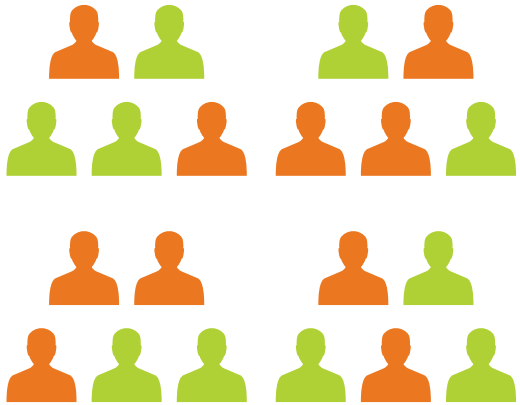
$P(Y = 1 \mid C = 1, G = A) = P(Y = 1 \mid C = 1, G = B)$

C is output of classifier, G is group, Y is (true) target variable

BASE RATE DIFFERENCES IN Y



GROUP A

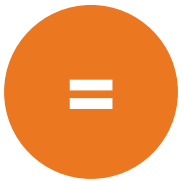


GROUP B

TRADE-OFFS ARE NECESSARY

If base rates differ across groups,
it is impossible to achieve any two of:

DEMOGRAPHIC PARITY



EQUALIZED ODDS

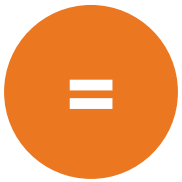
PREDICTIVE PARITY

Chouldechova 2017, Kleinberg et al. 2017, Eliassi-Rad & Fitelson 2018

TRADE-OFFS ARE NECESSARY

If base rates differ across groups,
it is impossible to achieve any two of:

DEMOGRAPHIC PARITY



EQUALIZED ODDS
(Propublica complaint)

PREDICTIVE PARITY
(Northpointe response)

Chouldechova 2017, Kleinberg et al. 2017, Eliassi-Rad & Fitelson 2018

FAIRNESS FOR WHOM?

Decision-maker: of those I've labeled low-risk, how many will recidivate?

Defendant: how likely am I to be incorrectly classified high-risk?

	labeled low-risk	labeled high-risk
did not recidivate	TN	**FP**
recidivated	**FN**	TP

Different metrics matter to different stakeholders

Slide credit: Julia Stoyanovich, Arvind Narayanan

FAIRNESS IN ALGORITHMS

Fairness is a **social and ethical concept**, not a statistical concept

Correcting for algorithmic bias generally requires:

- knowledge of how the measurement process is biased
- judgments about properties to satisfy in an “unbiased” world

Bias is **subjective** and must be considered **relative** to task

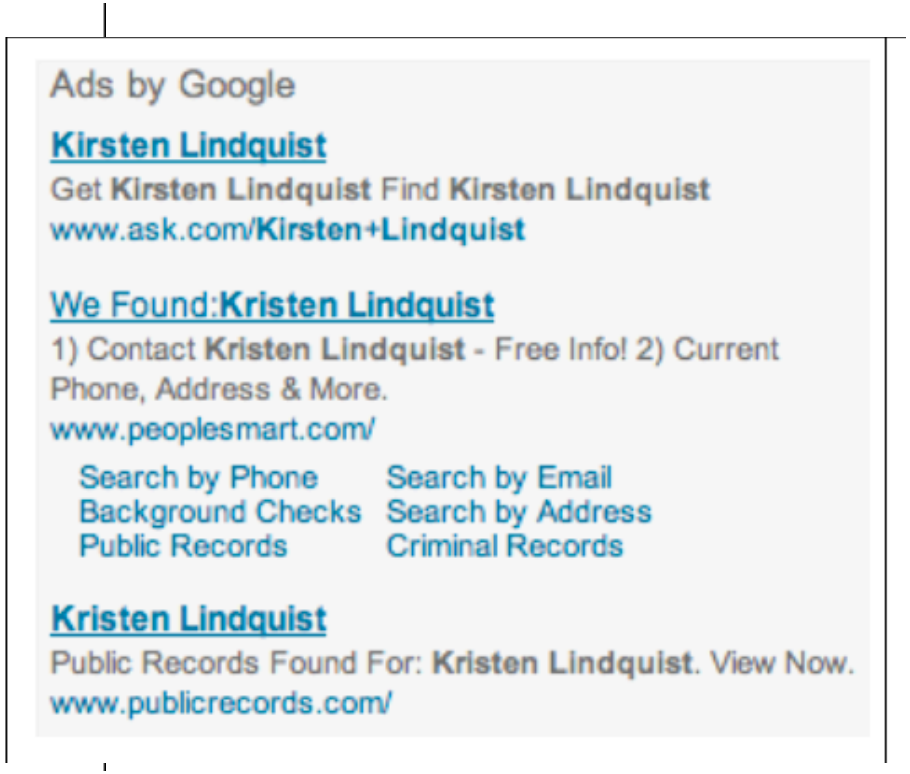
Slide credit: David Parkes

#7

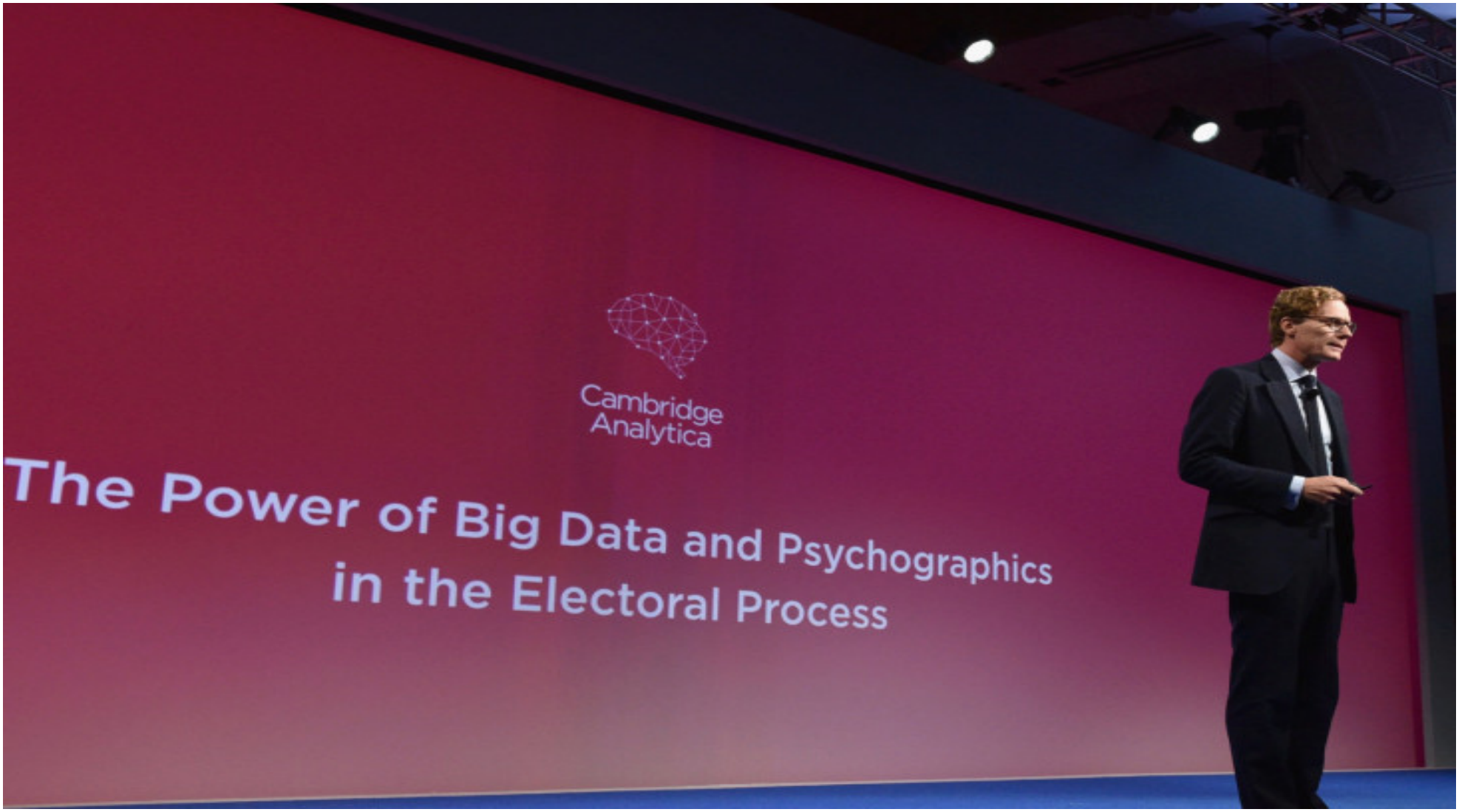
DATA DIGNITY

What is not Data Dignity?

DISCRIMINATION IN AD DELIVERY



Discrimination In Online Ad Delivery, Sweeney, ACMQueue, 2013



The Great Hack, Netflix Documentary

Image Credits: Bryan Bedder / Getty Images / Getty Images

DATA DIGNITY

having the ability to understand and control how your data is used

DATA AS LABOR

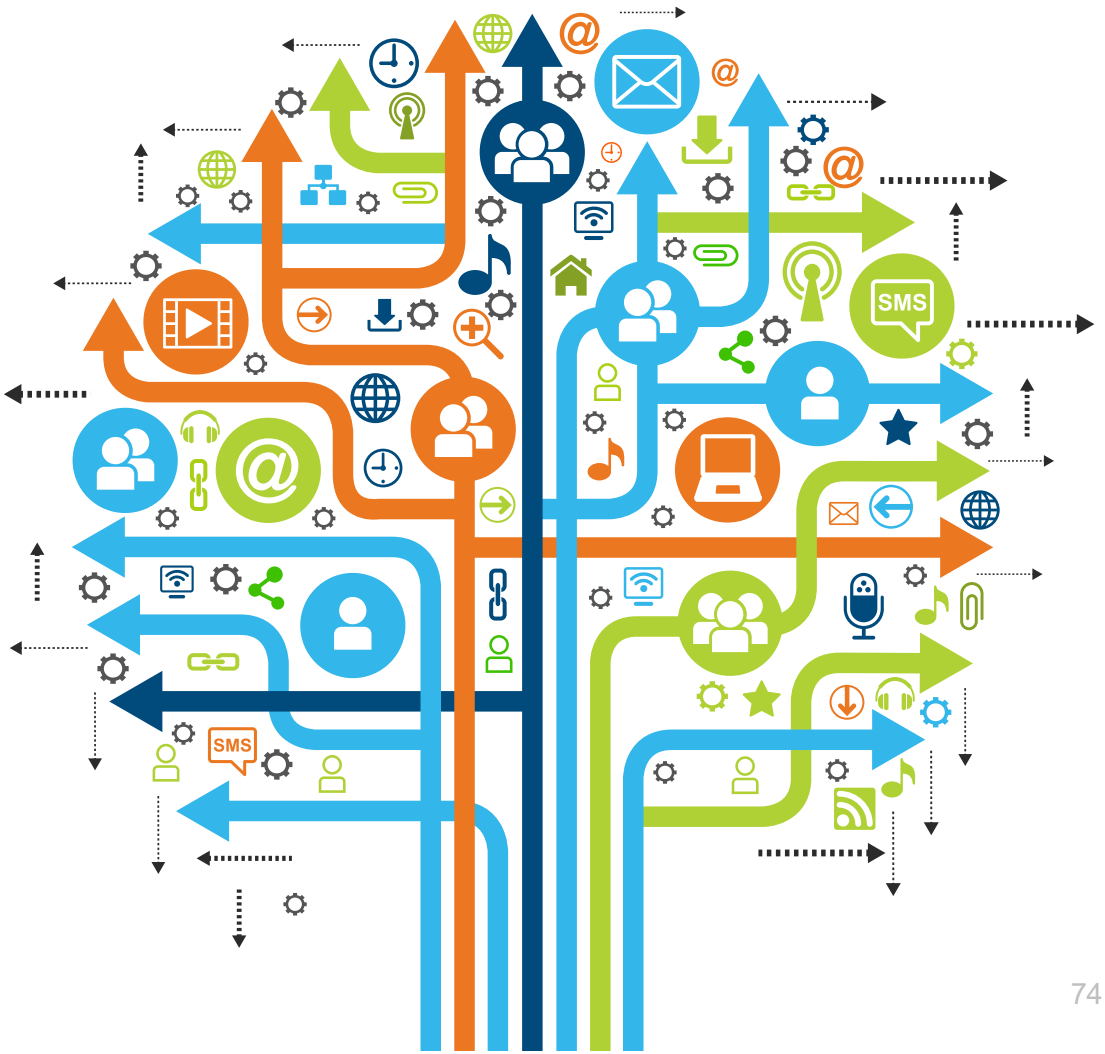
Ability to get paid for use of your data

Jaron Lanier, *Who Owns the Future*, Glen Weyl & Eric Posner, *Radical Markets*

Data Governance

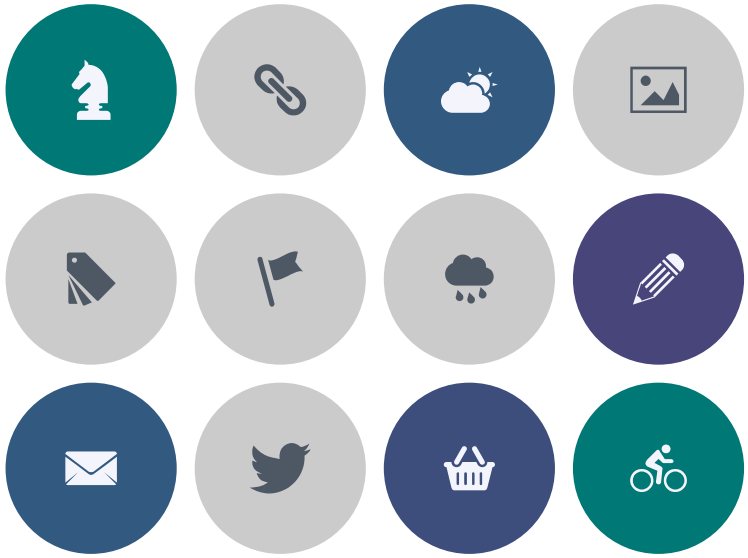
why is it so hard?

NETWORK EFFECTS



WHEN VALUE COMES FROM NETWORK EFFECTS, WHO OWNS THE NETWORK?

- Data about people is the output of a web of social activity.
- “Individual” data (shopping habits, personal travel) also contains information about other people.
- *People’s data can increase in value nonlinearly when combined with other peoples’ data.*



ISSUES SUMMARY

1. Formalization
2. High Dimensional Data
3. Measuring Error
4. Interpretability
5. Causal Modeling
6. Bias & Fairness
7. Data Dignity

ROADMAP

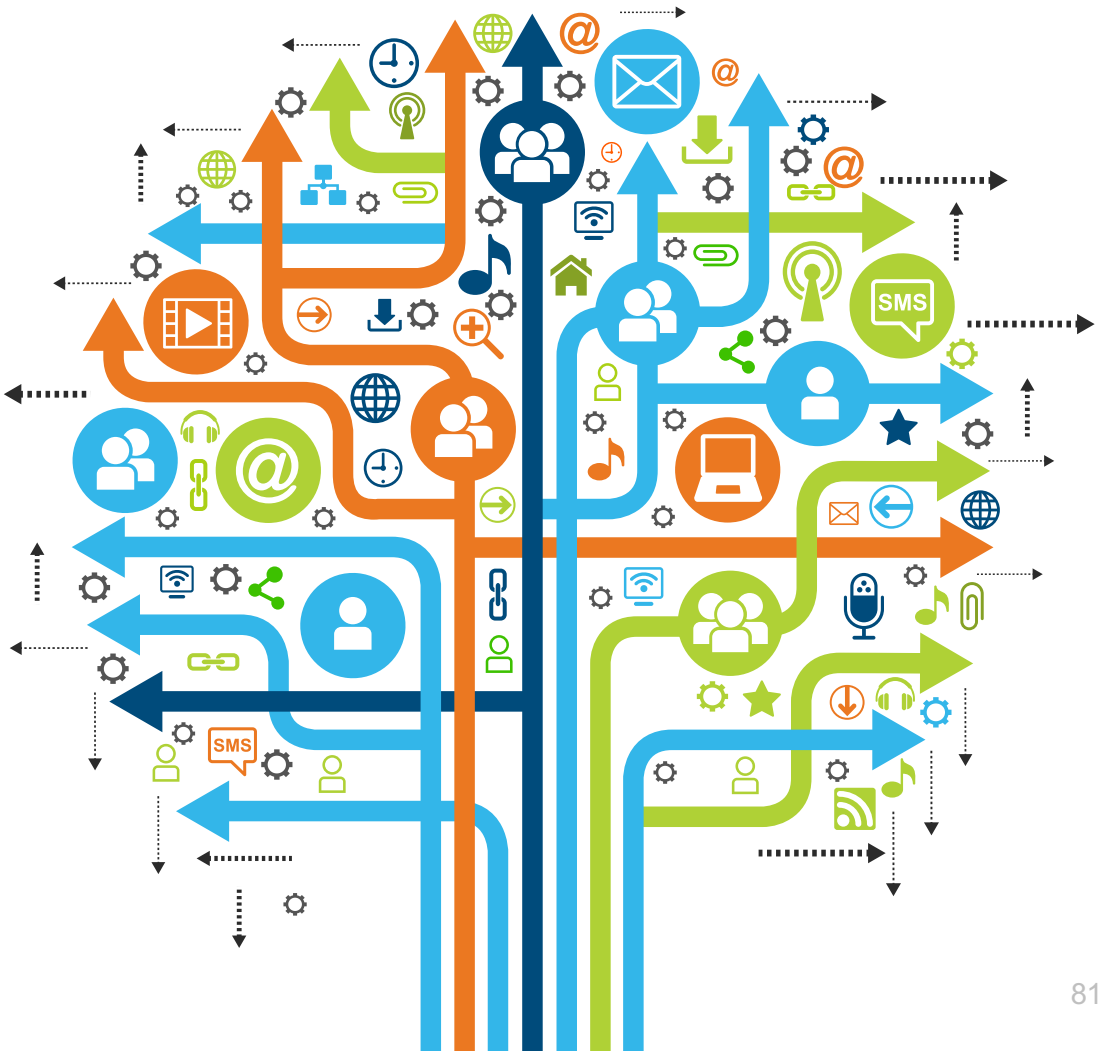
MACHINE LEARNING

RELATIONAL LEARNING

ARTIFICIAL INTELLIGENCE

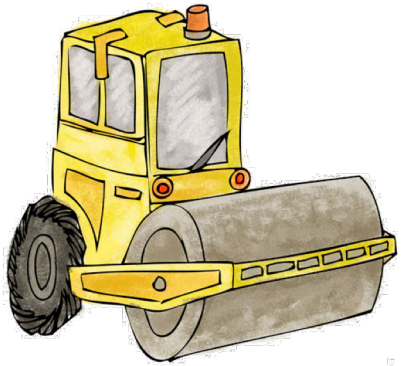

250 ml
ANTIDOTE

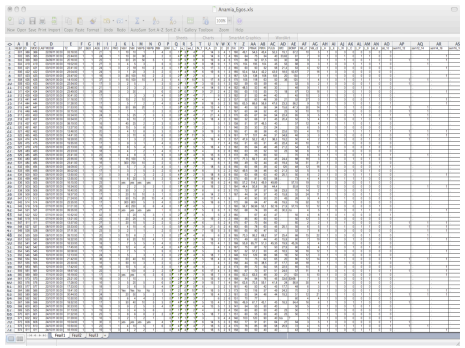
SOCIO-TECHNICAL SYSTEMS



FLATTEN

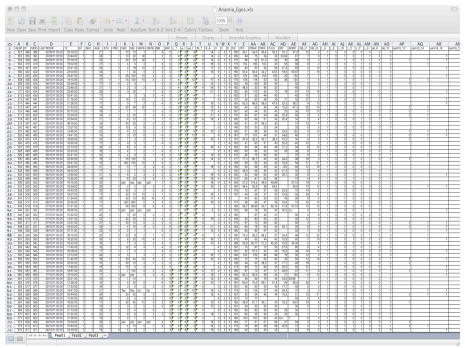
**Most machine learning algorithms take this rich structure,
and flatten it into tables
where each row is treated independently**

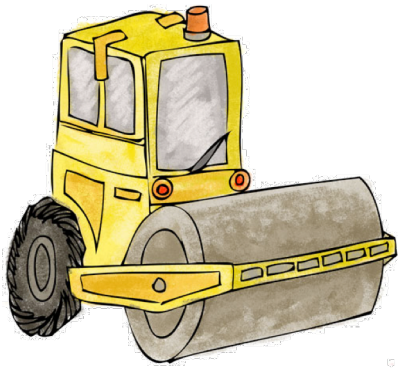




ISSUES

Incorrect independence assumptions
Models are not declarative
Doesn't support collective reasoning





NEED

machine learning methods
that take into account
RELATIONAL STRUCTURE

MY RESEARCH

Takes a more nuanced approach which takes into account logical relationships but at the same time takes into account context and probabilistic dependencies

[insert one hour talk on Relational Learning here]

PERILS OF IGNORING STRUCTURE

PERILS OF IGNORING STRUCTURE

Privacy

- Many approaches consider only individuals' attribute data
- Don't take into account what can be inferred from relational context

Fairness

- Impartial decision making without discrimination
- Need to take into account structural patterns

Algorithmic Discrimination

- Key structural pattern: feedback loop

PERILS OF IGNORING STRUCTURE

Privacy

- **Many approaches consider only individuals' attribute data**
- **Don't take into account what can be inferred from relational context**

ELENA ZHELEVA

Fairness

- Need to take into account structural patterns

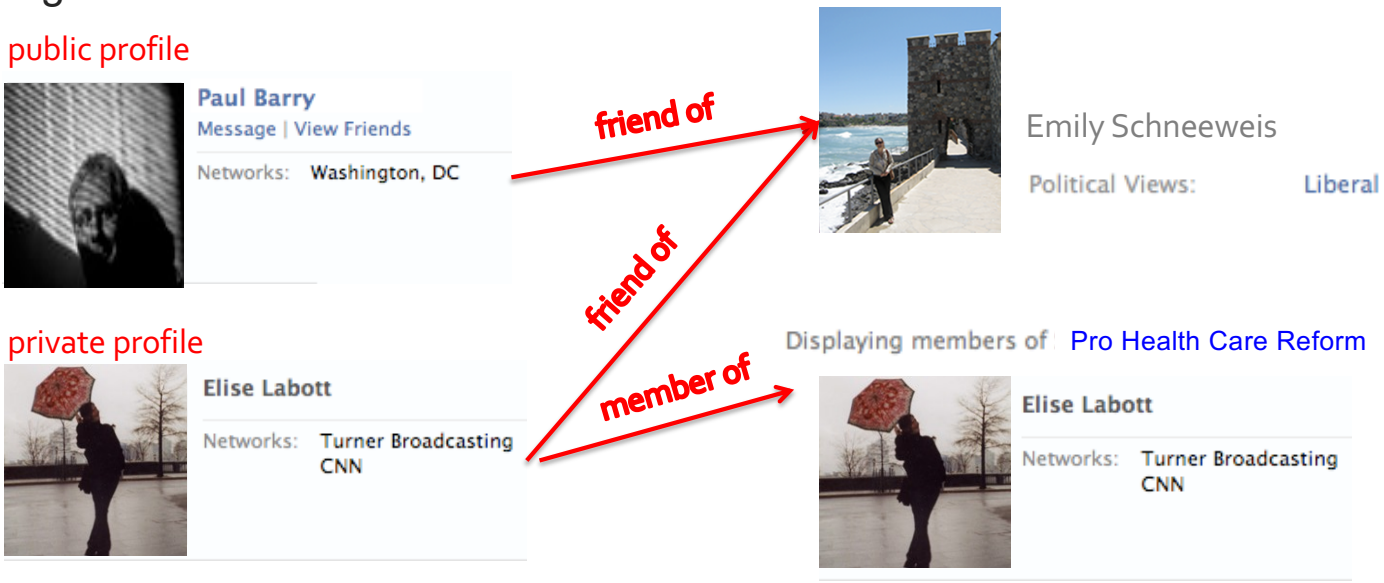
Algorithmic Discrimination

- Key structural pattern: feedback loop

Zheleva and Getoor, *To Join or Not to Join: the Illusion of Privacy in Online Social Networks,* WWW 2009

DISCLOSURE IN NETWORKS

- E.g. Is Paul liberal? Is Elise liberal?



Even when individual is able to control their privacy settings, information leaks, and control determined by group owners

To Join or Not to Join the Illusion of Privacy in Social Networks, Zheleva & Getoor, WWW2009

PERILS OF IGNORING STRUCTURE

Privacy

- Many approaches consider only individuals' attribute data
- Don't take into account what can be inferred from relational context

Fairness

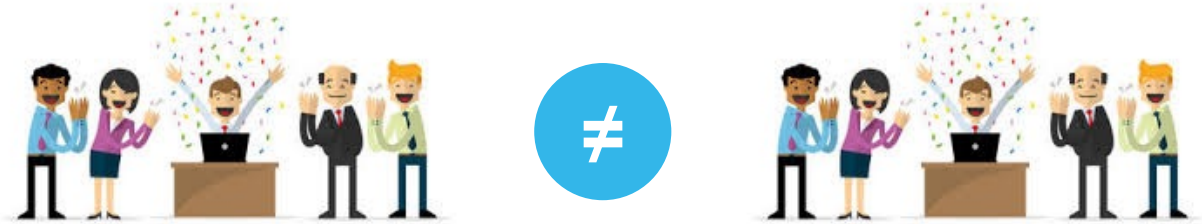
- **Need to take into account structural patterns**

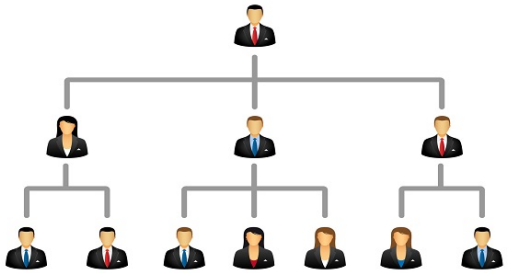
Algorithmic Discrimination

- Key structural pattern: feedback loop

RELATIONAL FAIRNESS

Employees with supervisors from different group
have different promotion chances
than employees from the matching group





Need to take into account
organizational hierarchy

RELATIONAL FAIRNESS

Existing literature on fairness in machine learning is limited to non-relational setting

Taking into account the social, organizational, and other connections between individuals is important

A Declarative Approach to Fairness in Relational Domains, Farnadi, Behrouz & Getoor, Data Engineering Bulletin 2019.

GOLNOOSH FARNADI

PERILS OF IGNORING STRUCTURE

Privacy

- Many approaches consider only individuals' attribute data
- Don't take into account what can be inferred from relational context

Fairness

- Impartial decision making without discrimination

Algorithmic Discrimination

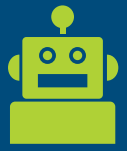
- **Key structural pattern: feedback loop**

Understanding (relational causal) structure can be key to mitigating effects!

ROADMAP

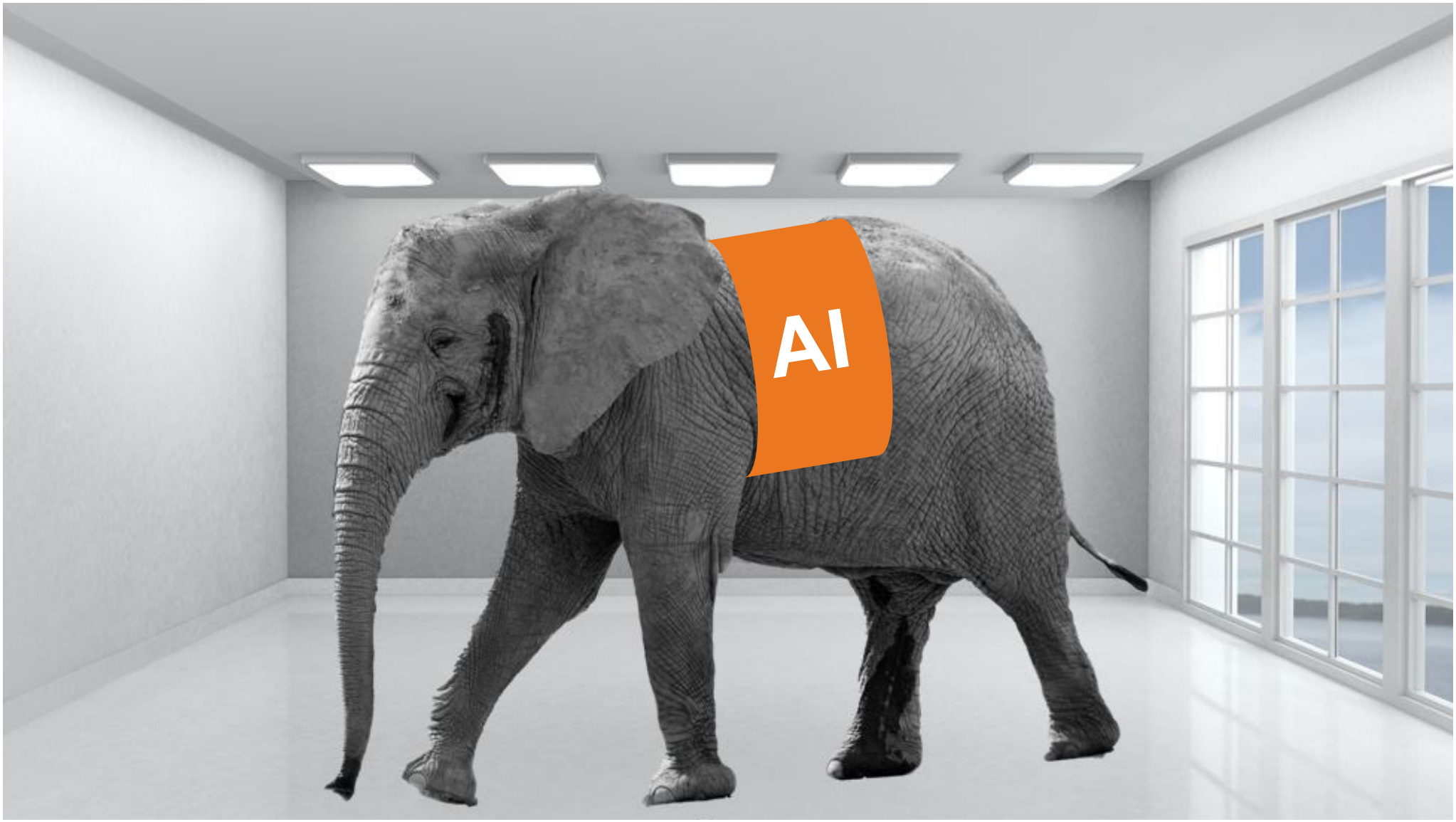
MACHINE LEARNING

RELATIONAL LEARNING



ARTIFICIAL INTELLIGENCE


iStockphoto.com | IlkerErgun


AI


Shutterstock.com

DATA SCIENCE REBRANDED AI

Data science has been rebranded as “AI” over the past few years. This rebranding is worthy of some scrutiny.

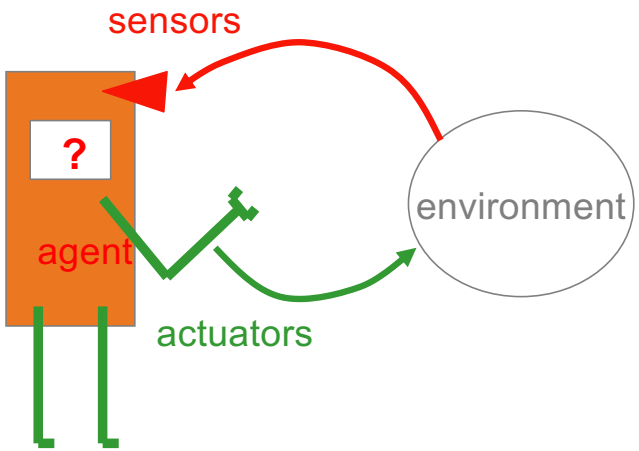
Artificial intelligence: The revolution hasn't happened yet. M. I. Jordan. *Medium*, 2018.

REAL DANGER *NOT* ARTIFICIAL GENERAL INTELLIGENCE OR SUPER INTELLIGENCE

REAL DANGER UNDUE TRUST PUT INTO LIMITED, FLAWED SYSTEMS

AI: MAXIMIZING EXPECTED UTILITY

AI: build rational agents that act to maximize their expected utility (MEU)



$$\operatorname*{argmax}_{a} \sum_{S'} P(S'|S,a)U(S')$$

**Reinforcement learning looks at delayed rewards,
and sequences of actions, but MEU is still at the core**

WHAT CAN GO WRONG?

#1

FRAME PROBLEM

FRAME PROBLEM

AI systems have very crude models of the world. Necessary limitations as to what is taken into account and what is left out

Consider the state transition function:

qualification problem: the preconditions for an action

ramification problem: the effects of an action

AI works for limited, specific, and constrained situations (“weak AI” or “narrow AI”)

#2

VALUE ALIGNMENT PROBLEM

VALUE ALIGNMENT PROBLEM

Whose utility?

Humans are very bad at specifying utility functions

Machines are ***beneficial*** to the extent that ***their*** actions can be expected to achieve ***our*** objectives. We need machines to be ***provably beneficial.*** Russell & Norvig, Artificial Intelligence: A Modern Approach, Fourth Edition upcoming

Proposed solution: learn utility function

#3

ATOMISTIC INTELLIGENCE

ATOMISTIC

Modern AI focuses on the construction of individual, autonomous agents

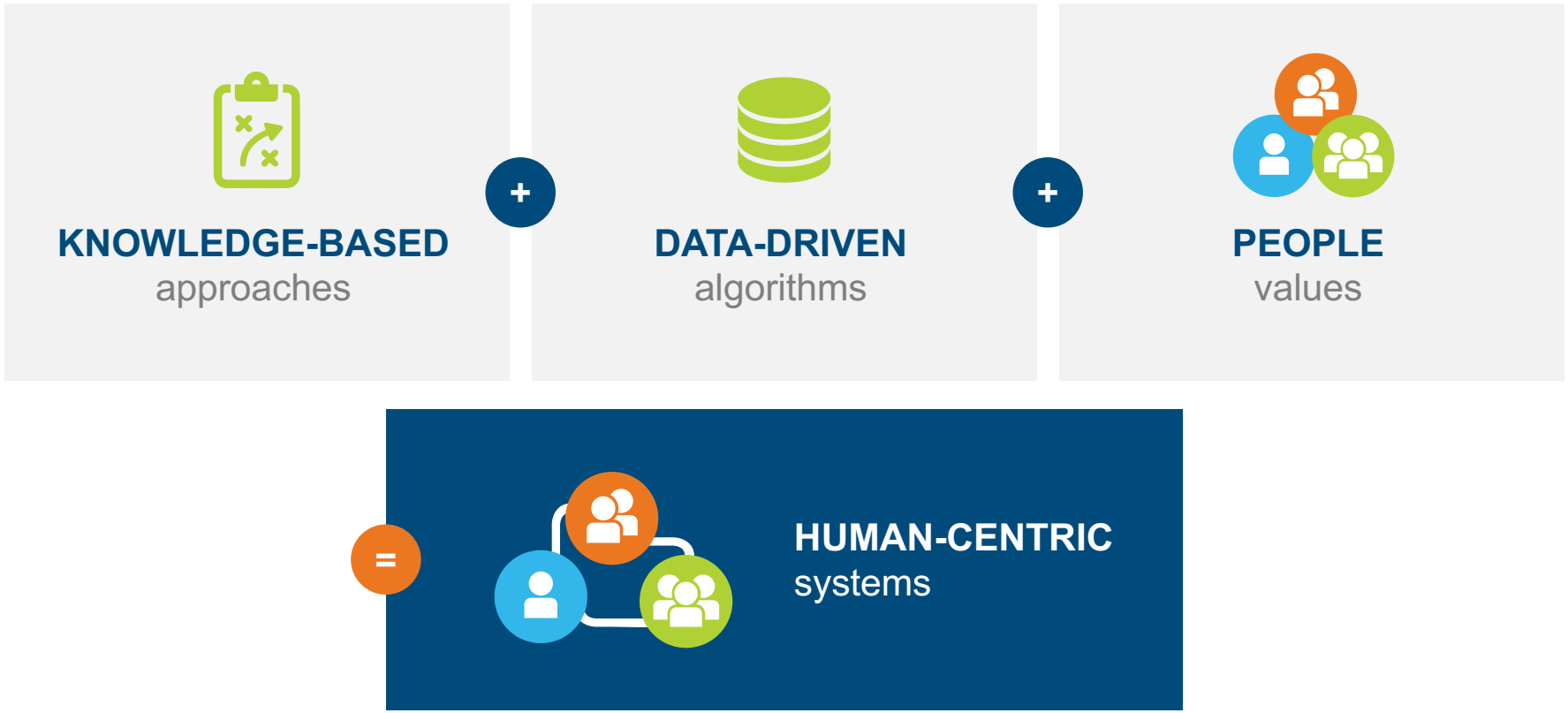
Assigns individual as the basic unit of analysis

This is a reductionist way of looking at the problem

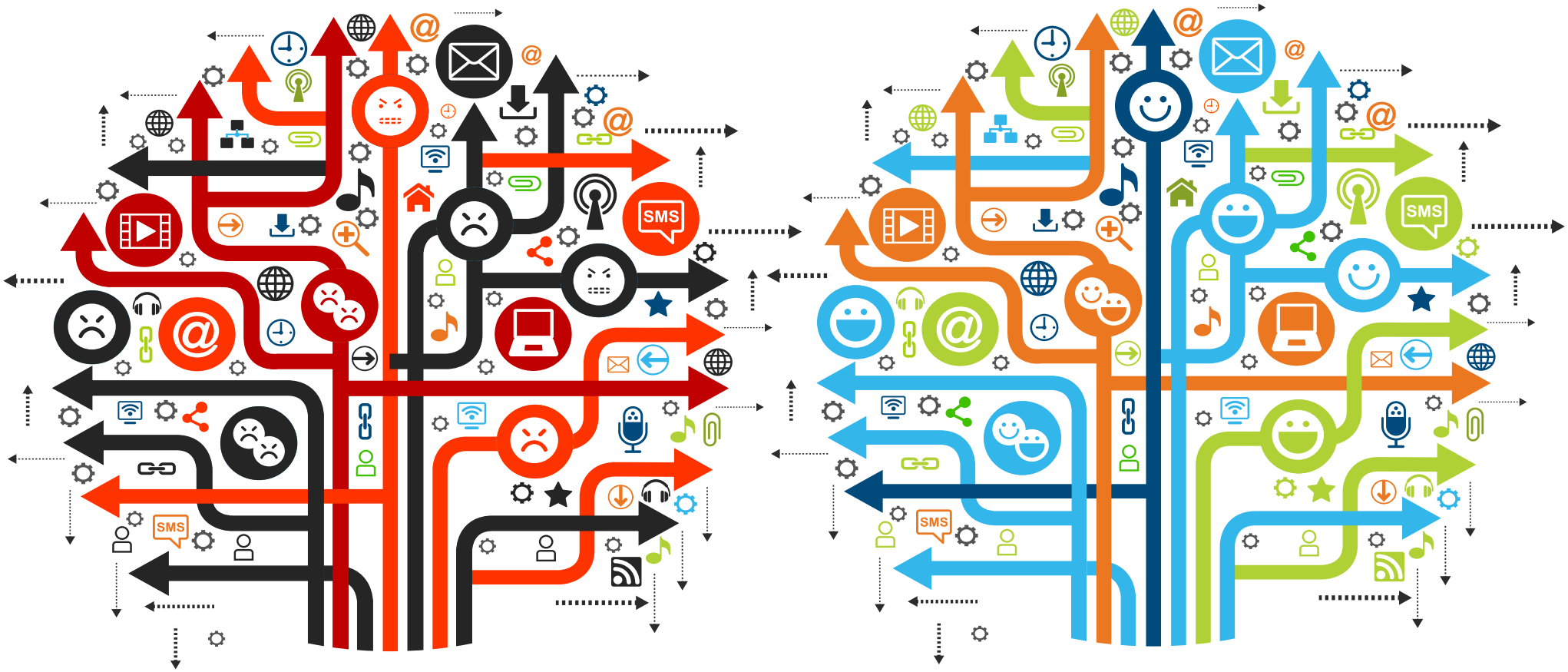
Ethics not just about individual moral/right action, takes into account social and political spheres

RATHER THAN AI AUTONOMY

MAINTAIN HUMAN AUTONOMY


KNOWLEDGE-BASED
approaches
+
DATA-DRIVEN
algorithms
+
PEOPLE
values
=
HUMAN-CENTRIC
systems

WHICH WILL WE CHOOSE?



CLOSING

ACKNOWLEDGEMENTS

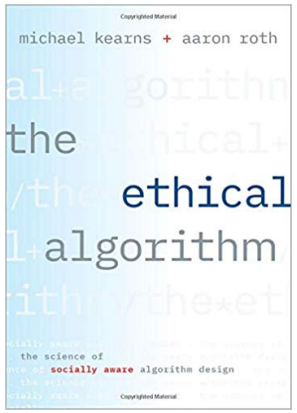
LINQS | LISE'S INQUISITIVE STUDENTS

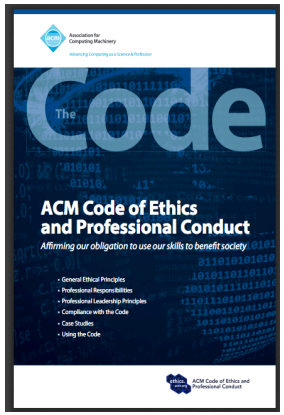
ACKNOWLEDGEMENTS

SPONSORS




EXPLOSION OF INTEREST IN ETHICS





Forbes Billionaires Innovation Leadership Money Consumer Industry Lifestyle Brand

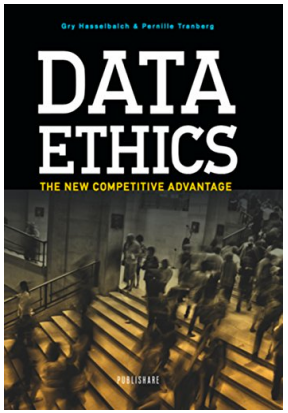
2,131 views | Oct 8, 2018, 01:42pm

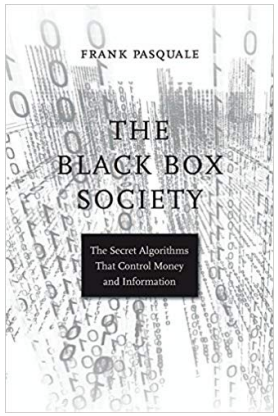
Do We Need To Teach Ethics And Empathy To Data Scientists?

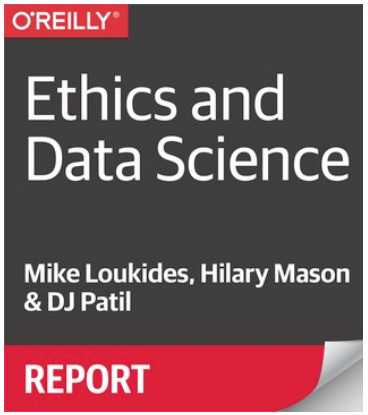
Kalev Leetaru Contributor ⓘ

AI & Big Data

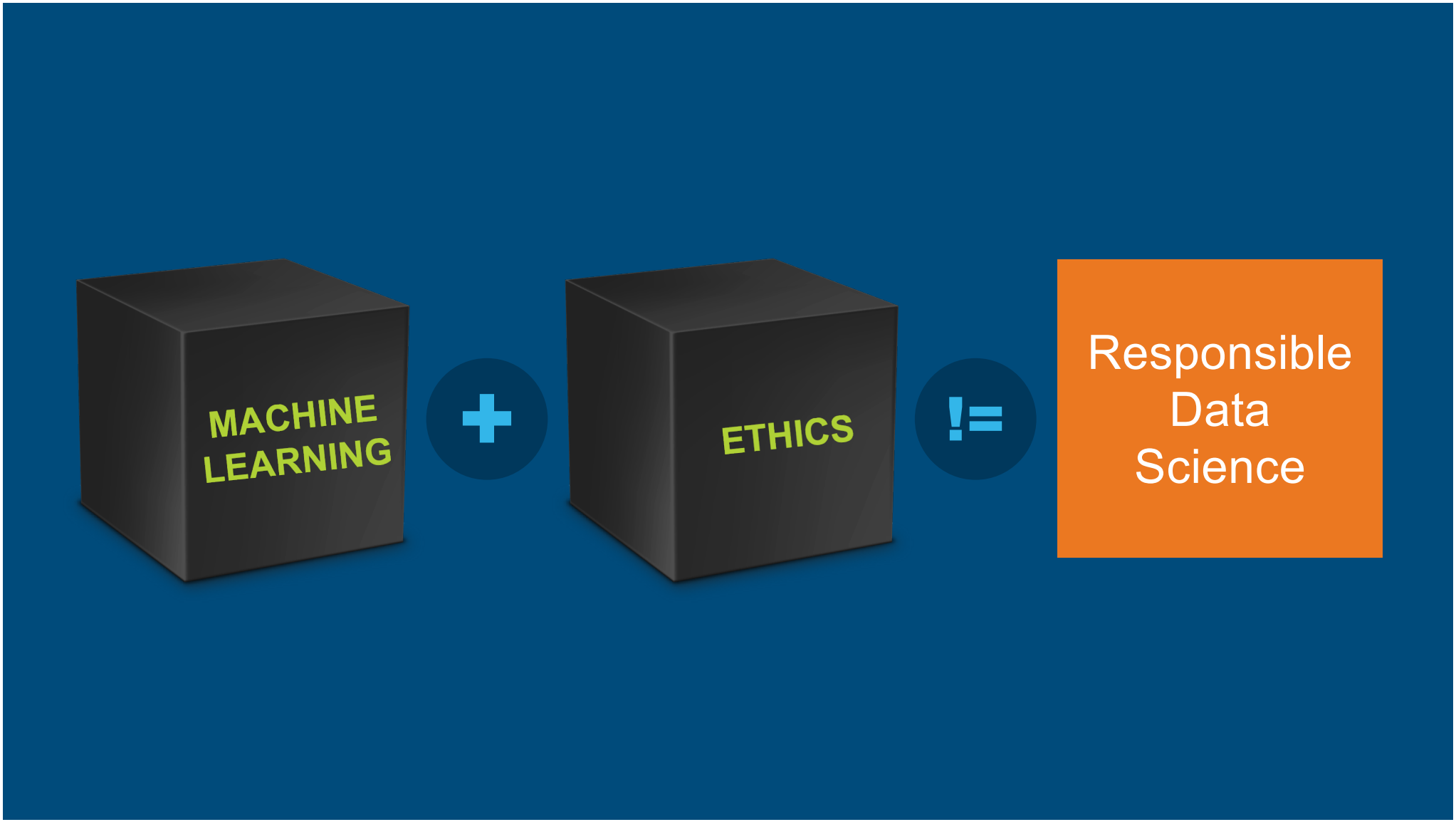
I write about the broad intersection of data and society.





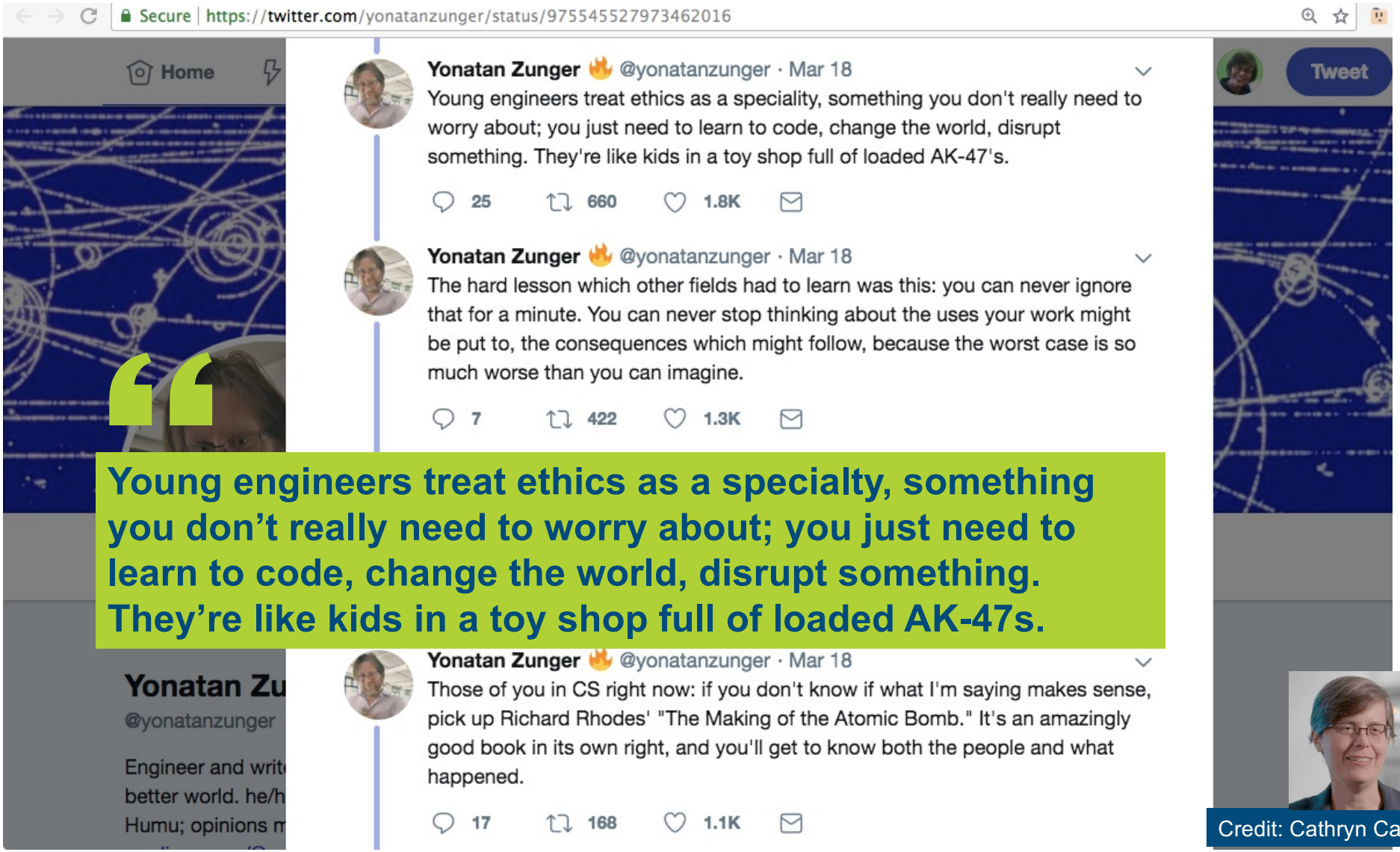


ETHICAL OS - HTTPS://ETHICALOS.ORG/


MACHINE LEARNING
ETHICS
!=
Responsible Data Science

BILL HOWE
UW

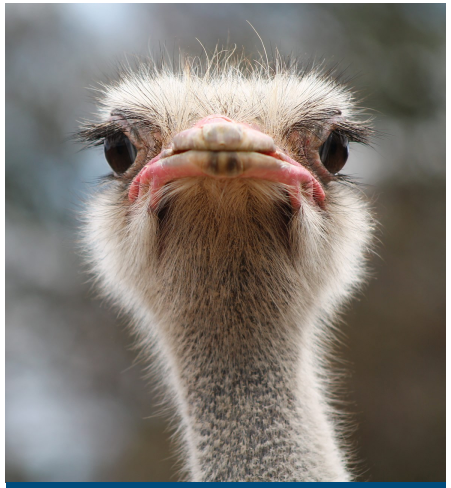
“Responsibility means going beyond technical or technocratic solutions to also involve substantive debate about ethics, values, and competing interests. How is ethical expertise defined? Who gets to be at the table? What are the limits of certain kinds of solutions?”



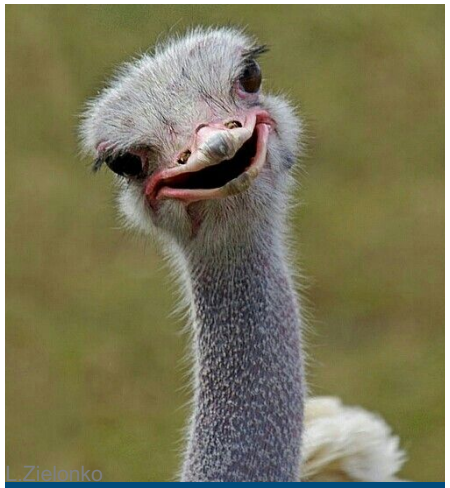
Credit: Cathryn Carson

CONCLUSION

NOT IN DENIAL



SKEPTIC



CURIOUS

ENGAGED

RESPONSIBLE

Images: Shutterstock.com

THANK YOU !

PROBABILISTIC
SOFT
LOGIC

psl.linqs.org

DATA SCIENCE
D³ RESEARCH CENTER
DATA • DISCOVERY • DECISIONS


getoor@ucsc.edu


@lgetoor