Enhancing Data Science Ethics through Statistical Education and Practice

Jessica Utts
Professor Emerita of Statistics
University of California, Irvine

Outline

- Part 1: Setting the stage with some history
- Part 2: The frightening present
- Part 3: How statisticians can help
- Part 4: Encouraging statistical literacy

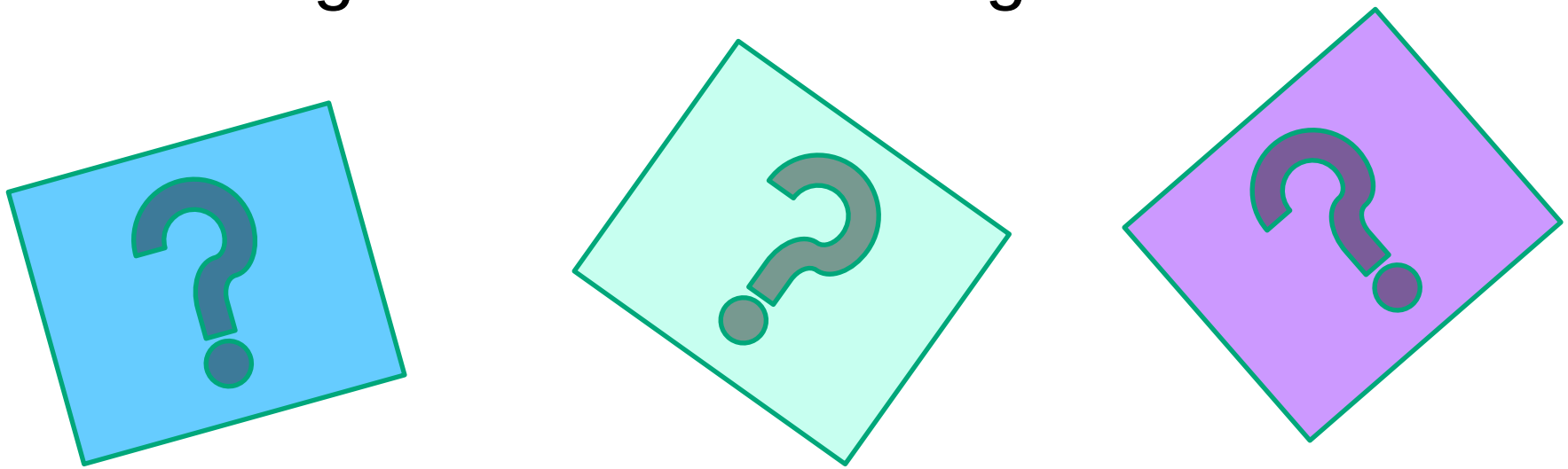
Part 1:
Setting the stage with some history

A story from my early teaching years

- An engineering professor gave students an assignment to design a pipeline to send blood from a poor developing nation to a rich developed one.
- The students got to work, discussing the optimal diameter for the pipe, how to go under a body of water, methods for keeping the blood fresh, etc.
- After letting them discuss for awhile, the professor demanded to know why not one of them had questioned the ethics of the assigned task.
- “This is a class in engineering not ethics,” was the answer the students gave.

Moral of the Story

- Train students (and practitioners) to ask *WHY* before asking *how*.
- Is the task ethical? Are there pros and cons?
- Who might benefit? Who might suffer?



Part 1, continued:
More of my history

Utts, Statistical Science, 1986: Science Fiction?

"This story occurs 30 to 40 years in the future. There are no more statisticians. There are statistical clerks, [who] feed data into the computer and out pops the appropriate model, estimate, or whatever, complete with the associated significance or confidence levels. These are sent to journals, along with a post hoc explanation for the results of any tests which turned out to be "significant." Everyone is quite happy with this arrangement. No one knows how the computer generates these answers, but everyone knows if the computer produced them, they must be right. All sorts of interesting hypotheses are being proved this way, and when they don't agree with common sense, everyone knows common sense must be wrong. Something finally goes wrong... a result which contradicts common sense so much that someone actually has the audacity to question what is happening in the computer."

Utts, Statistical Science, 1986, continued

"Team of scholars attempt to figure out what the computer is doing, but they can't, until they find a few very old retired statisticians who can read the literature from those days.

When the software was being developed most statisticians didn't pay much attention. The packages which were eventually implemented were written by people who were good at selling, but didn't really understand the concepts involved. A few statisticians tried to protest, but since they were advocating the use of their own services, no one took them seriously. After all, the journals were more likely to publish the [black box] version of the results, so why bother with the more cautious and complicated interpretations the statisticians were trying to sell?"

A few recommendations from 1986 paper

- Statisticians need to play a greater role in determining that our work is properly applied. Our techniques are simultaneously becoming more complex and more automated. They are less and less likely to be understood by non-statisticians.
- We should be teaching an undergraduate course on statistical thinking. We will not produce a statistically literate society by merely teaching the mechanics.
- We should reward interdisciplinary collaboration and encourage our best students to do it.
- These recommendations are even more relevant today!

Part 2:
The frightening present

Why this topic? Why now?

Why this topic? Why now?

Why AI Must Be Ethical — And How We Make It So

When artificial intelligence decides who's worthy, who's right, and who's criminal, you can only hope it makes the right call.

Jacob Bergdahl Follow
Jul 14 · 7 min read ★

It's 2021. The rain is pleasantly pouring outside. You're having an afternoon tea, while working on your next project on your laptop. Your workflow is interrupted by a phone call. It's Danielle from Subtling. You were there last Friday on a job interview that you're

Why this topic? Why now?

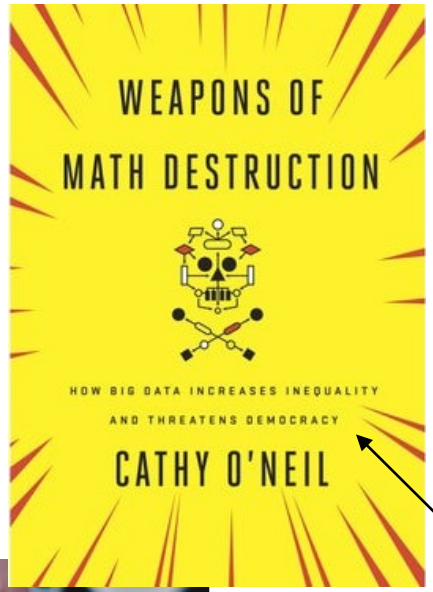
Why AI Must Be Ethical — And How We Make It So

When artificial intelligence decides who's worthy, who's right, and who's criminal, you can only hope it makes the right call.

Jacob Bergdahl Follow
Jul 14 · 7 min read ★

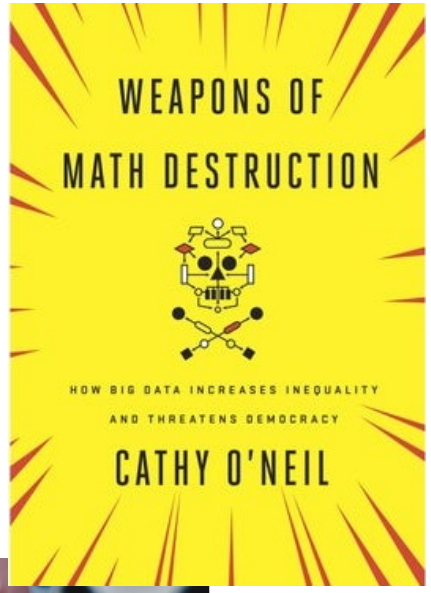
It's 2021. The rain is pleasantly pouring outside. You're having an afternoon tea, while working on your next project on your laptop. Your workflow is interrupted by a phone call. It's Danielle from Subtling. You were there last Friday on a job interview that you're

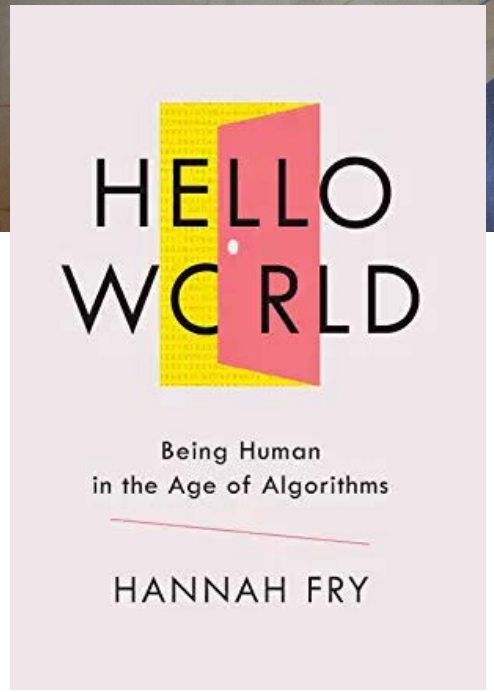
https://towardsdatascience.com/why-ai-must-be-ethical-and-how-we-make-it-so-b52cdb1dd15f 1/8



How big data increases inequality and threatens democracy

Why this topic? Why now?





Why AI Must Be Ethical — And How We Make It So

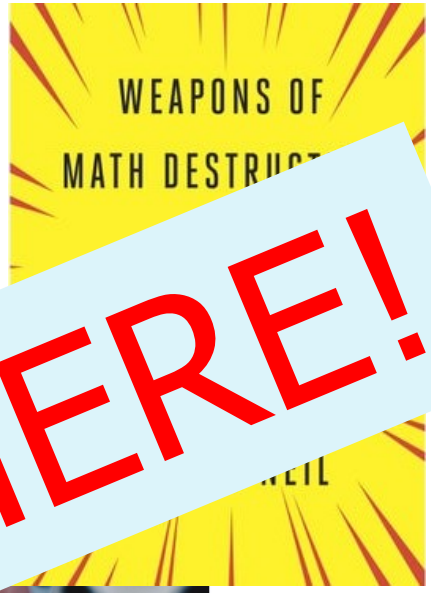
When artificial intelligence decides who's worthy, who's right, and who's criminal, you can only hope it makes the right call.

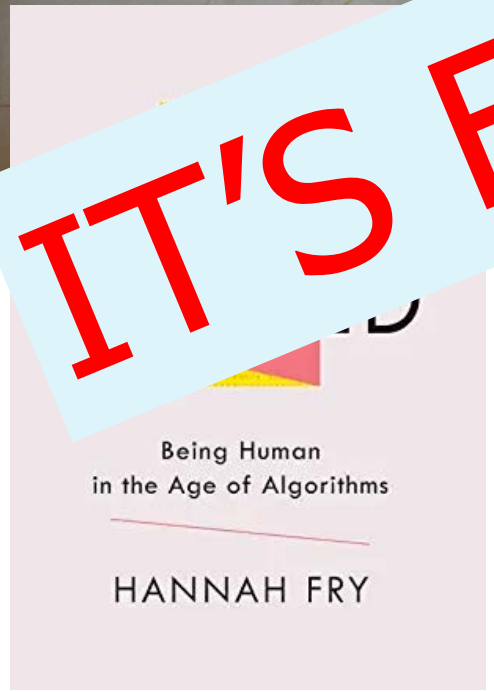
Jacob Bergdahl Follow
Jul 14 · 7 min read ★

It's 2021. The rain is pleasantly pouring outside. You're having an afternoon tea, while working on your next project on your laptop. Your workflow is interrupted by a phone call. It's Danielle from Subtling. You were there last Friday on a job interview that you're

https://towardsdatascience.com/why-ai-must-be-ethical-and-how-we-make-it-so-b52cdb1dd15f 1/8

Why this topic? Why now?





IT'S EVERYWHERE!

Why AI Must Be Ethical — And How We Make It So

When artificial intelligence decides who's worthy, who's right, and who's criminal, you can only hope it makes the right call.

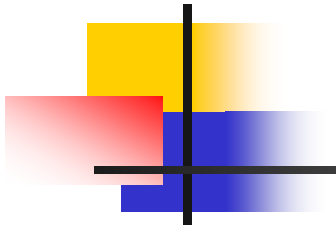
Jacob Bergdahl Follow
Jul 14 · 7 min read ★

It's 2021. The rain is pleasantly pouring outside. You're having an afternoon tea, while working on your next project on your laptop. Your workflow is interrupted by a phone call. It's Danielle from Subtling. You were there last Friday on a job interview that you're

Why this topic? Why now? Continued...

- Lines are blurring for data science:

  Statistics/machine learning/artificial intelligence
- Our students' jobs reflect this cross-over
- Traditional ethical issues for statisticians
  - See for instance "ASA Ethical Guidelines"
  - Latest version approved April 14, 2018
- Not enough. Complexity => new ethical issues
- Need to educate students on ethics of decisions and interpretations
  - As consumers
  - As data scientists



Example: GPS Map program

Is it ethical...

- To clog roads by sending everyone on the same route when leaving a large event?
- To send cars through high-crime areas?
- To even identify high-crime areas?
- To send pedestrians through high-crime areas?
- To increase traffic in residential areas?
- What about school zones?

ACLU Congress Face Recognition Study

- Used facial recognition system Amazon offers to public (Rekognition), which anyone could use.
- Running the entire test cost $12.33.
- Built a face database and search tool using 25,000 publicly available arrest photos. Then searched that database against public photos of every current member of the House and Senate.
- Used default match settings Amazon sets for Rekognition.
- False matches disproportionately people of color; six members of the Congressional Black Caucus.

AI ethics examples: Facial recognition

Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots

By Jacob Snow, Technology & Civil Liberties Attorney, ACLU of Northern California

JULY 26, 2018 | 8:00 AM

TAGS: Face Recognition Technology, Surveillance Technologies, Privacy & Technology

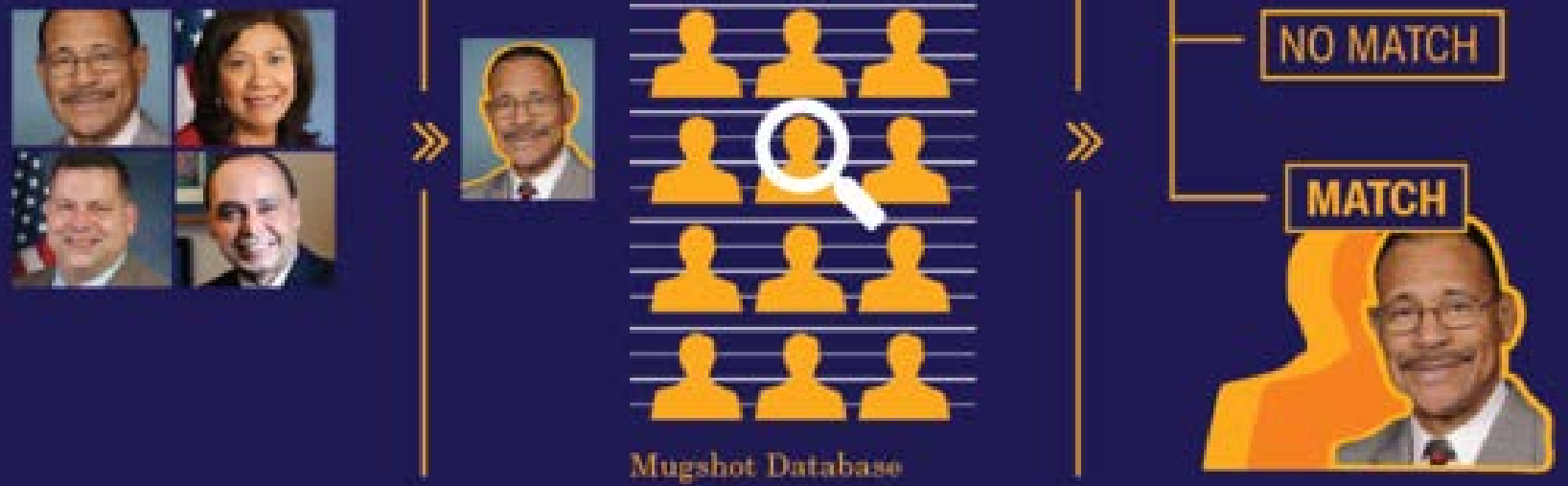
28/535 = 5.2%

https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

https://aws.amazon.com/rekognition/

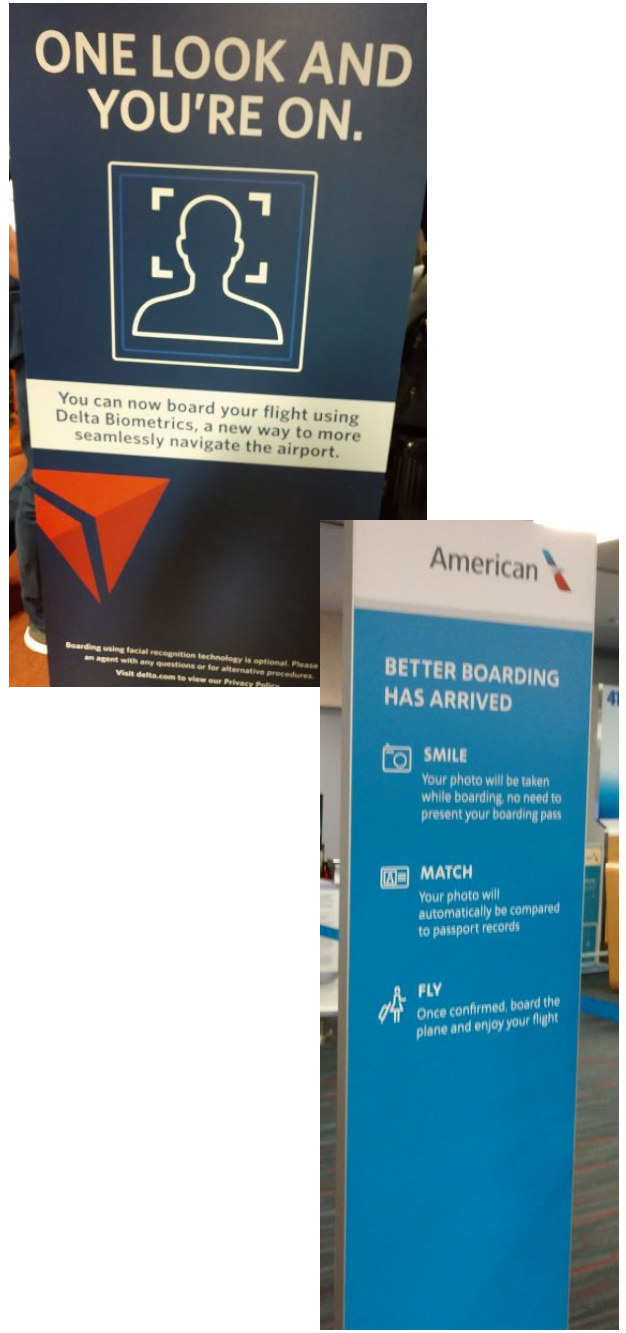
The Rekognition Scan

Comparing input images to mugshot databases

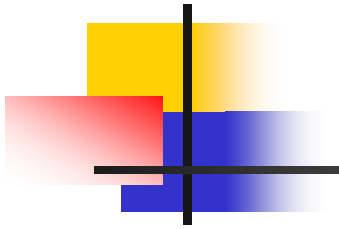


Facial recognition software: Benefit/risk tradeoffs

- Used for people boarding airplanes, finding celebrities in videos, verifying banking customers, etc.
- Used by police to find [possible] criminals
- Has been shown to be less reliable for people of color and women than white men. Especially bad for children.
- Has resulted in false convictions, and even deaths.



Other classic AI ethics examples



- Bias in hiring algorithms, based on bias in training data.
- Algorithms used by judges to decide who is likely to commit (another) crime.
- Deciding who should get loans based on geographic and other aggregate data.
- Medical diagnostic algorithms trained on data excluding certain sub-populations...
- But, are they better or worse than humans??

The New York Times

She Was Arrested at 14. Then Her Photo Went to a Facial Recognition Database.

With little oversight, the N.Y.P.D. has been using powerful surveillance technology on photos of children and teenagers.

A few examples from the media

BUSINESS NEWS OCTOBER 9, 2018 / 8:12 PM / 10 MONTHS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin 8 MIN READ

SAN FRANCISCO (Reuters) - Amazon.com Inc's (AMZN.O) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

amazon

Example: Individual vs Societal Benefit

- DNA tests for genealogy are very popular.
- GEDMatch.com* allows free comparison of your DNA with others, to find relatives.
- Can upload DNA results from multiple testing companies.
- Find matches and contact them to share genealogical ancestor information.

*GED is short for Geneological Data. A GEDCom file is a generic format, shorthand for “genealogical data communication”

Unintended Use of GEDMatch Data!

YOLO COUNTY NEWS

Local News

Arrest made in serial-killer case with ties to Davis

By Lauren Keene

East Area Rapist struck locally in 1978

A former police officer who for decades led a quiet existence in a Sacramento suburb was named by authorities Wednesday as the East Area Rapist — also known as the Golden State Killer — whose dozens of alleged crimes included three sexual assaults in Davis during the summer of 1978.

Joseph James DeAngelo, 72, was taken into custody earlier that morning outside his Citrus Height home and booked into the

East Area Rapist/Golden State Killer suspect Joseph James DeAngelo

The Davis Enterprise online accessed August 4, 2019; original article April 26, 2018

Prognosis

DNA Website Had Unwitting Role in Golden State Manhunt

Kristen V. Brown

May 29, 2018, 11:34 AM PDT
Updated on May 29, 2018, 2:00 PM PDT

- ▶ GEDmatch, founded by a Florida grandfather, pools genetic data
- ▶ Privacy debate swirls after investigators sifted site's users

Curtis Rogers is used to helping people track down their relatives. The 79-year-old Florida grandfather of six founded a genealogy website that helps hobbyists like himself trace the branches on their family tree.

In the past few weeks, though, Rogers has been fielding inquiries from a different kind of user: the police.

https://www.bloomberg.com/news/articles/2018-05-29/killer-app-dna-site-had-unwitting-role-in-golden-state-manhunt

GEDMatch is a public database

- Should law enforcement be allowed to use it to identify criminals?
- Note that they weren't given any special access; they used it like other users, but with crime scene DNA.
- If your DNA was there, would you want it used to find criminals who are your relatives?
- In June users were given opt-in choice.

June email from GEDMatch (excerpts)

- We at GEDmatch recently instituted a change of policy regarding law enforcement (LE) use of our site. We now require specific approval from each user who wishes to have their information available for LE use.
- We are concerned about the handicap our actions have placed on solving cold cases. There are millions of victims, including family and friends of violent crime victims and unidentified remains who need some sense of closure. We have a fast start to rebuilding the LE portion of the GEDmatch database. We encourage everyone who has had a genetic DNA test done to consider helping to build the database for law enforcement use as quickly as possible.

Some Important AI Ethical Guidelines

- Transparency. “Black box” allure gives more credibility than justified. Ignores uncertainty.
- Consider likely biases in data sources.
- Everyone on an interdisciplinary team should take responsibility for ethical issues.
- Humans should always be involved in decisions based on algorithms.
- Correlation does not imply causation. And algorithms can suggest intervention action.

So far, mixed success

- Implementing AI ethics in corporations has not been easy.
- Multiple groups and researchers recommend government oversight.
- No agreement on how to balance corporate secrets and algorithmic transparency.

Technology

Google announces AI ethics panel

Dave Lee
North America technology reporter

26 March 2019

26 March 2019

GETTY IMAGES

Google has launched a global advisory council to offer guidance on ethical issues relating to artificial intelligence, automation and related technologies.

The panel consists of eight people and includes former US deputy secretary of state, and a University of Bath associate professor.

The group will "consider some of Google's most complex challenges", the firm said.

26 March 2019

Technology

Google annc

Dave Lee

North America tech

26 March 2019

Google has laun
relating to artific

The panel consist
University of Bath

The group will "cc

Technology

Google's ethics board shut down

By Jane Wakefield

Technology reporter

5 April 2019

5 April 2019

GETTY IMAGES

An independent group set up to oversee Google's artificial intelligence efforts, has been shut down less than a fortnight after it was launched.

The Advanced Technology External Advisory Council (ATEAC) was due to look at the ethics around AI, machine learning and facial recognition.

One member resigned and there were calls for another to be removed.

AI Now Institute's 2019 Report (NYU)

100 pages, 12 recommendations. For example:

- Machine learning researchers should account for potential risks and harms and better document the origins of their models and data.
- AI bias research should move beyond technical fixes to address the broader politics and consequences of AI's use.
- States should craft expanded biometric privacy laws that regulate both public and private actors.

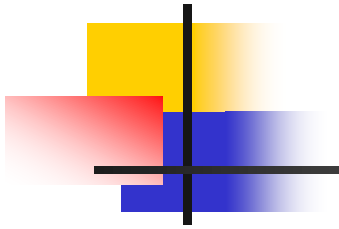
https://ainowinstitute.org/AI_Now_2019_Report.pdf

Part 3:
How statisticians can help

What statisticians can (uniquely?) offer

- Data issues, for example:
  - How to get high-quality data
  - How to assess bias in data
  - How conclusions depend on data sources
- Analysis issues, for example:
  - Intelligent use of modeling; consider assumptions
  - Multi-variable thinking; but pitfalls of multiple analyses
  - Dealing with outliers
- Reporting issues, for example:
  - Practical vs statistical significance; what p-values mean
  - When causal conclusions can be made

Areas of ethical concern for statisticians



- Ethics in data collection, quality and uses
- Ethical implementation of details in a study
- Issues of ethics during the analysis
  - These are the most technical issues; mostly for practitioners and future practitioners.
  - But all students need to know what can go wrong.
- Ethics of reporting results
  - To clients
  - To the media and the public
- Teaching statistical literacy in all introductory statistics courses is an ethical obligation.

Ethics of data issues; questions to ask

- Informed consent for all uses of data and/or all interventions?
- Individual anonymity, and likely to remain so when merged with other datasets?
- Are there likely biases in the data? Sub-populations that are under-represented?
- Missing data or drop outs for reasons related to the research questions?

Examples of ethics in implementing a study

- Ecological validity of intervention
  - Example: Time of day electric use study
- Ethics of interventions without consent
  - Cornell/Facebook emotion study (more next)
- Power analysis to make sure a large enough sample is used to detect a meaningful effect
  - Consider power for sub-groups too.

Facebook/Cornell Emotion Study

- 2012 study, randomly selected 689,003 Facebook users, assigned to 4 groups.
- No informed consent!
- One group had negative news feed reduced; another had positive news feed reduced. Control groups had news feed randomly omitted. Study lasted one week.
- Use of negative and positive words used in subjects' own posts were measured.

Results from Cornell press release

“News feed: Emotional contagion sweeps Facebook”

- “People who had positive content experimentally reduced on their Facebook news feed for one week used more negative words in their status.”
- “When news feed negativity was reduced the opposite pattern occurred. Significantly more positive words were used in peoples’ status updates.”

https://news.cornell.edu/stories/2014/06/news-feed-emotional-contagion-sweeps-facebook

More about the Emotion study

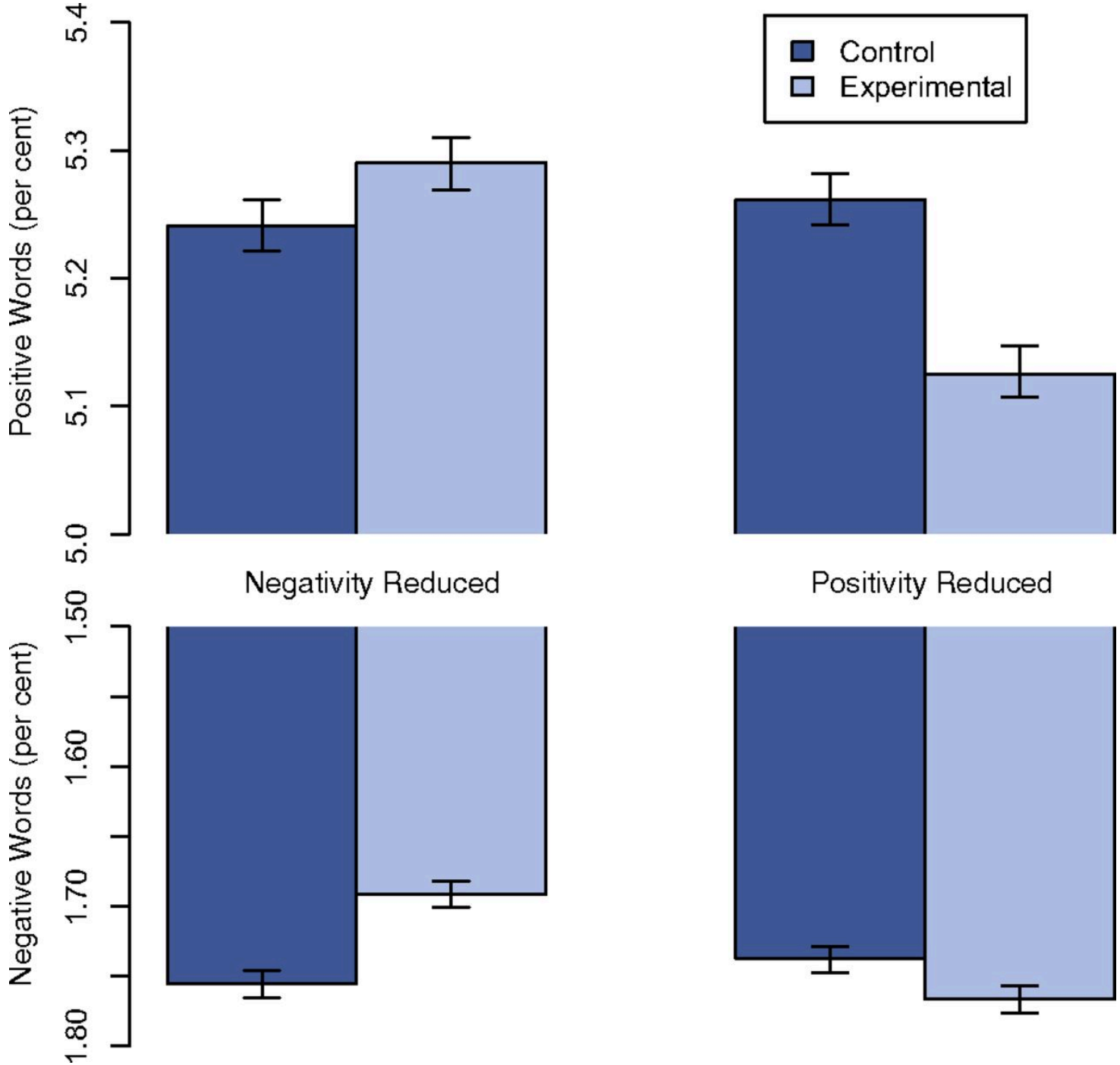
- Published in the *Proceedings of the National Academy of Sciences*, June 2014
- According to Altmetric data:
  - Mentioned by 337 news outlets
  - 136 blogs
  - 4164 tweeters
  - “In top 5% of all research outputs scored by Altmetric”

BUT, the actual results...

"When positive posts were reduced in the News Feed, the percentage of positive words in people's status updates decreased by B = −0.1% compared with control [t(310,044) = −5.63, P < 0.001, Cohen's d = 0.02], whereas the percentage of words that were negative increased by B = 0.04% (t = 2.71, P = 0.007, d = 0.001)."

Results, continued…

“Conversely, when negative posts were reduced, the percent of words that were negative decreased by B = −0.07% [t(310,541) = −5.51, P < 0.001, d = 0.02] and the percentage of words that were positive, conversely, increased by B = 0.06% (t = 2.19, P < 0.003, d = 0.008).”


Control
Experimental
Positive Words (per cent)
5.0
5.1
5.2
5.3
5.4
Negativity Reduced
Positivity Reduced
Negative Words (per cent)
1.50
1.60
1.70
1.80

Authors' justification...

- "The effect sizes from the manipulations are small (as small as d = 0.001).
- Given the massive scale of social networks such as Facebook, even small effects can have large aggregated consequences.
- And after all, an effect size of d = 0.001 at Facebook's scale is not negligible: In early 2013, this would have corresponded to hundreds of thousands of emotion expressions in status updates per day."

Ethical Issues from this Study

- No informed consent.
- Misleading graphs.
- Confusion of statistical significance with practical significance (importance)
- Justification of small effect size as being of practical importance because of large population affected.

Examples of Ethics in Analysis

- Multiple tests, type 1 errors, p-hacking
  - Example: Breakfast cereal and boy babies (next)
- Collinearity and (mis)interpretation of individual regression coefficients
- Hidden, unrealistic Bayesian priors
  - Ex: Bem, Johnson and Utts parapsychology paper in response to Wagenmakers et al
- Ignoring data not missing at random
  - Ex: Dropping out of drug trial due to side effects

Example of multiple tests: Does eating cereal produce boys?

- Headline in *New Scientist*: “Breakfast cereal boosts chances of conceiving boys”
- Numerous other media stories of this study.
- Study in *Proc. of Royal Soc. B* showed of pregnant women who ate cereal, 59% had boys, of women who didn’t, 43% had boys.
- Problem 1: Headline implies eating cereal *causes* change in probability, but this was an observational study!

Cereal Problem 2: Multiple Testing

- The study investigated 132 foods the women ate, at 2 time periods for each food = 264 possible tests! (Stan Young pointed this out in a published criticism.)
- By chance alone, *some* food would show a difference in birth rates for boys and girls.
- Main issue: Selective reporting of results when many relationships are examined, not adjusted for multiple testing. Quite likely that there are false positive results.

Ethics of Reporting Results

- Focus on magnitude, not p-values.
  - With big data, small effects have tiny p-values
- Include clear explanation of uncertainty.
- Don’t overstate the importance of results.
- Graphics should be clear, not misleading.
- Media coverage should include all relevant results, not just most interesting or surprising.
- Don’t imply causal connection if not justified.

Example: Reporting to client & media

- Suppose a client asks you to evaluate an online game for boosting children's math skills.
- Data provided include pre-post math and language scores, time spent studying each.
- Results: Math scores went up but language scores went down, and game was addictive.
- Are you ethically bound to report the negative consequences of using the game...
  - To the client?
  - In media requests?

Example: Hormone replacement therapy

- Women's Health Initiative, randomized study comparing hormones with placebo.
- Surprising result was *increase* in risk of coronary heart disease in hormone group.
- Trial was stopped early, and millions of women were advised to stop taking HRT (hormone replacement therapy) immediately.
- Large scale media attention on risks of heart disease and breast cancer from HRT.

But… Results from the original article

"Absolute excess risks per 10,000 person-years attributable to estrogen plus progestin were 7 more CHD [coronary heart disease] events, 8 more strokes, 8 more PEs [pulmonary embolism], 8 more invasive breast cancers, while absolute risk reductions per 10,000 person-years were 6 fewer colorectal cancers and 5 fewer hip fractures."

More results..

- 231 out of 8506 women taking the hormones died of any cause during the study, which is 2.72%
- 218 of the 8102 women taking placebo died of any cause, which is 2.69%
- Adjusted for the time spent in the study, the death rate was slightly lower in the hormone group, with an annualized rate of 0.52% compared with 0.53% in the placebo group.

Ethical issue for reporting results

- The media and medical community focused on the surprising heart disease results
- In fact the hormone group fared *better* in many ways, including adjusted death rate.
- Were millions of women misled?
- If full results had been reported in the media, women could decide for themselves, for instance based on family or personal medical history.

Part 4:
Encouraging statistical literacy

Ethics in statistics education

- For training statisticians:
  - Include ethical considerations throughout their training
  - Idea: Ask for a discussion of ethical issues as part of all data analysis projects, possibly dissertations as well*
- For educating all students:
  - Statistical literacy involves recognizing ethical issues
  - Emphasize topics students can use in their lives, to help them make informed decisions and recognize statistical errors

*Thanks to Eric Vance for this suggestion.

My Top 10 Important Literacy Topics

1. Observational studies, confounding, causation
2. The problem of multiple testing
3. Sample size and statistical significance
4. Why many studies fail to replicate
5. Does decreasing risk actually increase risk?
6. Personalized risk versus average risk
7. Poor intuition about probability and risk
8. Using expected values to make decisions
9. Surveys and polls – good and not so good
10. Confirmation bias

Observational studies, confounding, causation

Example headline from observational study:

“Breakfast Cereals Prevent Overweight in Children”

Worldhealthme.com, 4/12/13

The article continues:

> *“Regularly eating cereal for breakfast is tied to healthy weight for kids, according to a new study that endorses making breakfast cereal accessible to low-income kids to help fight childhood obesity.”*

Some Details

- Observational study
- 1024 children, only 411 with usable data
  - Mostly low-income Hispanic children in Texas, USA
  - Control group for a larger study on diabetes
- Asked what foods they ate for 3 days, in each of 3 years (same children for 3 years) looked at number of days they ate cereal = 0 to 3 each year.
- Lead author = Vice President of Dairy MAX, a regional dairy council

More Details: The analysis

- Multiple regression was used
  - Response variable = BMI percentile each year (BMI = body mass index)
  - Explanatory variable = days of eating cereal in each year (0 to 3), modeled as linear relationship with BMI!
- Did not differentiate between other breakfast or no breakfast (if not cereal)
- Also adjusted for age, sex, ethnicity and some nutritional variables

Some problems

- Observational study – no cause/effect.
- Obvious possible confounding variable is general quality of nutrition in the home
  - Unhealthy eating for breakfast (non-cereal breakfast or no breakfast), probably unhealthy for other meals too.
- Possible reversed cause/effect: High metabolism could cause low weight and the need to eat breakfast. Those with high metabolism require more frequent meals.

More of my favorite (!) headlines

- 6 cups a day? Coffee lovers less likely to die, study finds *NBC News website*, 5/16/12
- Spanking lowers a child's IQ *LATimes*, 9/25/09
- Joining a Choir Boosts Immunity *Woman's World*, 6/27/16
- Walk faster and you just might live longer *NBC News website*, 1/4/11
  - Researchers find that walking speed can help predict longevity
  - The numbers were especially accurate for those older than 75

Again: My Top 10 Important Topics

1. Observational studies, confounding, causation
2. The problem of multiple testing
3. Sample size and statistical significance
4. Why many studies fail to replicate
5. Does decreasing risk actually increase risk?
6. Personalized risk
7. Poor intuition about probability and risk
8. Using expected values to make decisions
9. Surveys and polls – good and not so good
10. Confirmation bias

Poor intuition about probability & risk

- William James was first to suggest that we have an *intuitive* mind and an *analytical* mind, and that they process information differently.
- Example: People feel safer driving than flying, when probability suggests otherwise.
- Psychologists have studied many ways in which we have poor intuition about probability assessments.
  - Recommended reading: *Thinking, Fast and Slow* by Daniel Kahneman

Which do you think is more probable?

- A massive flood somewhere in North America next year, in which more than 1,000 people drown.
- An earthquake in California sometime next year, causing a dam to burst resulting in a flood in which more than 1,000 people drown.

The Representativeness Heuristic and the Conjunction Fallacy

- **Representativeness heuristic:** People assign higher probabilities than warranted to scenarios that are *representative* of how they *imagine* things would happen.
- This leads to the **conjunction fallacy** … when detailed scenarios involving the conjunction of events are given, people assign *higher* probability assessments to the *combined event* than to statements of one of the simple events alone.
- But *P(A* and *B)* = *can't exceed* *P*(*A*)

Confusion of the inverse: examples

- P(disease | pos. test) vs P(pos. test | disease), especially for rare disease (doctors get this wrong)
- Cell phones and driving (old study):
  - P(Using cell phone | accident) = .015
  - P(Distracted by passenger | accident) = .109
  - Does this mean other occupants should be banned while driving but cell phones are okay??
- In court P(innocent | evidence) vs P(evidence | innocent)

Confusion of the inverse: DNA Example

- Dan is accused of crime because his DNA matches DNA at a crime scene (found through database of DNA). Only 1 in a million people have this specific DNA.
- Suppose there are 6 million people in the local area, so about 6 have this DNA. Only one is guilty!
- Is Dan almost surely guilty??

Is Dan Guilty??

Let’s look at hypothetical 6 million people.

Only 6 have a DNA match

	Guilty	Innocent	Total
DNA match	1	5	6
No match	0	5,999,994	5,999,994
Total	1	5,999,999	6,000,000

How to illustrate conditional probabilities

- P(DNA match | Dan is innocent)
  ≈ 5 out of almost 6 million, extremely low!
  - *Prosecutor would emphasize this*

	Guilty	Innocent	Total
DNA match	1	5	6
No match	0	5,999,994	5,999,994
Total	1	5,999,999	6,000,000

How to illustrate conditional probabilities

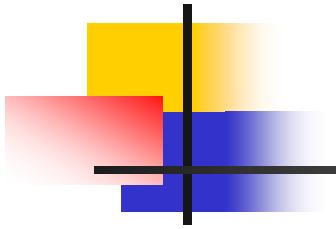
- But… P(Dan is innocent | DNA match)

≈ 5 out of 6, fairly high!

  - *Defense lawyer would emphasize this*

	Guilty	Innocent	Total
DNA match	1	5	6
No match	0	5,999,994	5,999,994
Total	1	5,999,999	6,000,000

DNA Example, continued



- P(DNA match | Dan is innocent) *very low*

  P(Evidence | accused is innocent)

  Prosecutor would emphasize this

- P(Dan is innocent | DNA match) *fairly high*

  P(Accused is innocent | evidence)

  Defense lawyer would emphasize this

- Jury needs to understand this difference!

Again: My Top 10 Important Topics

1. Observational studies, confounding, causation
2. The problem of multiple testing
3. Sample size and statistical significance
4. Why many studies fail to replicate
5. Does decreasing risk actually increase risk?
6. Personalized risk
7. Poor intuition about probability and risk
8. Using expected values to make decisions
9. Surveys and polls – good and not so good
10. Confirmation bias

Expected Values: What would you do?

You are planning an overnight trip but you don't want to go if the weather is bad. You look at hotels and find a room with the following:

- Pay $170 now, nonrefundable *OR*
- Pay $200 when you arrive, but only if you actually go

- What should you do? What additional information would help you decide?

Expected value for your decision

- Define p = probability you go on the trip
- Expected costs for each decision:
  - For advance purchase, E(Cost) = $170
  - If you don't pay in advance

    E(Cost) = $200(p) + $0(1 – p) = $200p
- Which is lower?

  $200p < $170 when p < (170/200) = 0.85.
- Decision: Pay advance purchase if probability of going on the trip is at least 0.85, but not otherwise.

Insurance, lottery, extended warranty

Should you buy an extended warranty?
What about insurance? (e.g. earthquake?)

- *On average* the company wins
- But *some* consumers will be winners, and some will be losers.
- You can use knowledge of your own circumstances to assess which is likely for you.

Probability, Intuition, Expected Value

Examples of Consequences in daily life:

- Assessing probability when on a jury

  Lawyers provide detailed scenarios – people give higher probabilities, even though *less* likely.

- Extended warranties and other insurance

  "Expected value" favors the seller

- Gambling and lotteries

  Again, average "gain" per ticket is negative

- Poor decisions (e.g. driving versus flying)

Summary and Conclusions

- Statisticians have a major role to play in data science ethics.
- Need to speak up as a member of an interdisciplinary team. Ask *why* before how.
- Need to teach our statistics students ethics alongside technical issues.
- Need to teach all students how to identify ethical issues and mistakes in reports based on statistical studies.

THANK YOU

Contact info:

jutts@uci.edu

http://www.ics.uci.edu/~jutts

UCIrvine
UNIVERSITY OF CALIFORNIA, IRVINE