Diversity in Data: How Data Collection Decisions Help and Harm the Outcome

Frauke Kreuter – JPSM, University of Maryland; LMU Munich – fkreuter@umd.edu

Threats to Diversity

- Application Examples
- Decisions in Collection
- Decisions in Analysis

Efforts to Consider

- Learning Across Context
- Capturing the Unmeasured
- Keeping the Small Groups

NATIONAL ACADEMY OF SCIENCES

1

1

Application Examples - ADMs

2

Work & Wealth – at the individual level

Job Center

- Interventions to reduce long-term unemployment
- Unnecessary for all with low risk + limited resources
- AI helps prioritize

Job Ads

- AI selects job ads based on likely clicks, past hiring, best matches
- Reproduction of current inequalities
- However, opportunities for 1st screenings
  Hangartner et al. 2021. Nature, Vol 589

Photo by Souvik Banerjee on Unsplash

3

3

Health and Prevention – at the individual level

HEALTHY CHICAGO
CHICAGO DEPARTMENT OF PUBLIC HEALTH

- Less than half of HIV-patients stay in prevention programs
- AI predicts drop out
- Intervention: Resources allocated based on drop out risk

Ramachandran, A., Kumar, A., Koenig, H. et al. Predictive Analytics for Retention in Care in an Urban HIV Clinic. Sci Rep 10, 6421 (2020). https://doi.org/10.1038/s41598-020-62729-x

- Johnson County, Kansas
  Carnegie Mellon University
- Goal: Break spiral of mental health illness and incarceration
- AI predicts incarceration
- Medical intervention prioritized for those with high mental health risk

Rodolfa, K.T., Lamba, H. & Ghani, R. Empirical observation of negligible fairness–accuracy trade-offs in machine learning for public policy. Nat Mach Intell 3, 896–904 (2021). https://doi.org/10.1038/s42256-021-00396-x

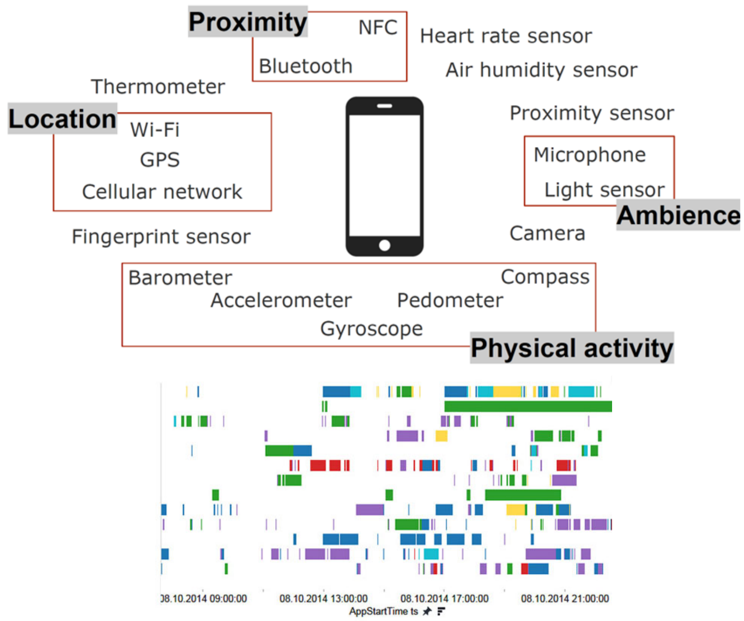
4

4

Health & Society

Goals:
- Increase efficiency
- Reduce subjectivity
- Keep effectiveness

Ultimate goal:
- Better society



5

Justice

Goals:
- Increase efficiency
- Reduce subjectivity
- Maintain effectiveness

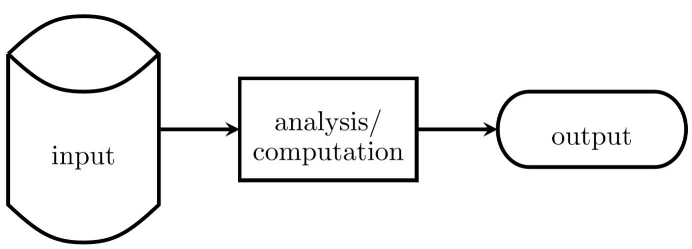
Ultimate goal:
- Fewer inmates

Library of Congress Prints and Photographs Division Washington, D.C. 20540 USA (intermediary roll film) thc 5a38062 https://hdl.loc.gov/loc.pnp/thc.5a38062

6

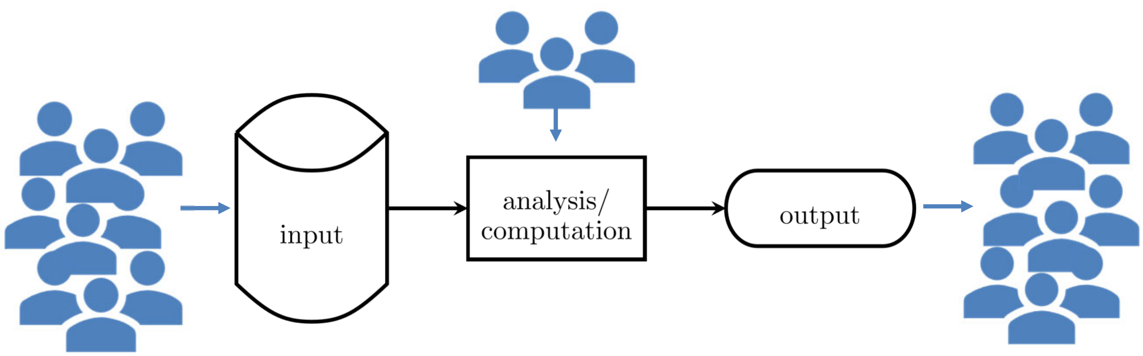
Decisions in the Process

7



8

Threats to Diversity



9

June 12, 2018; July 15, 2019

Google Image Search: “University Professor”

10

July 19, 2019

Google Image Search: “University Professor”

11

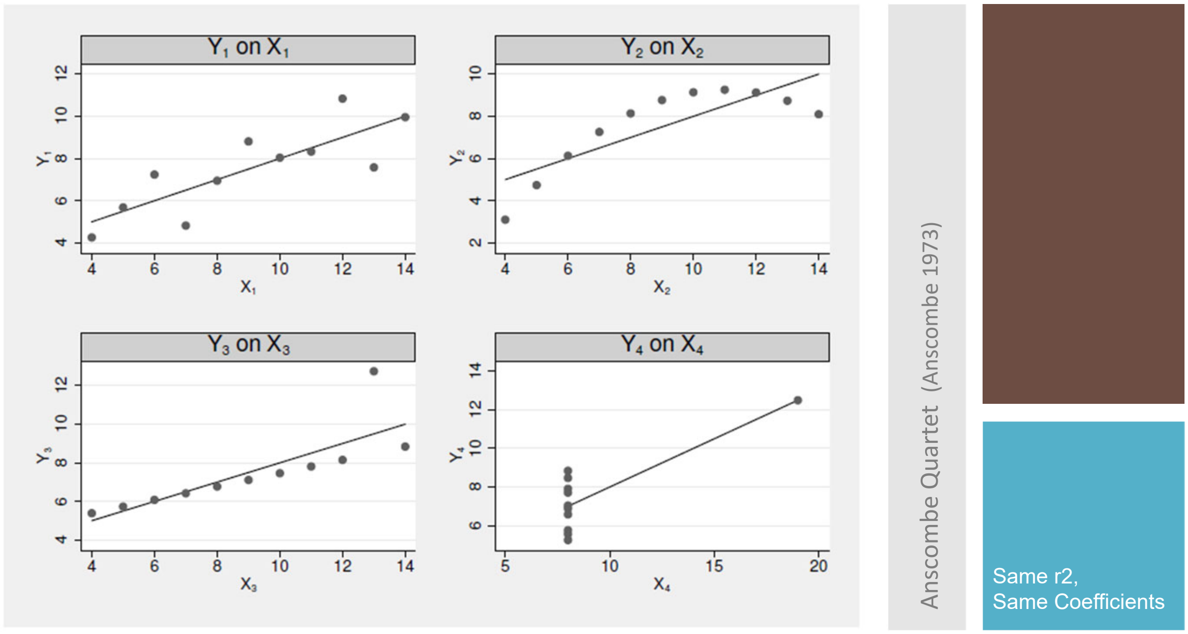
Fairness Challenge

Imperfect training data → many subpopulations

– *Misrepresentation of subpopulations affects accuracy and can have considerable real-world consequences (Buolamwini 2019)*

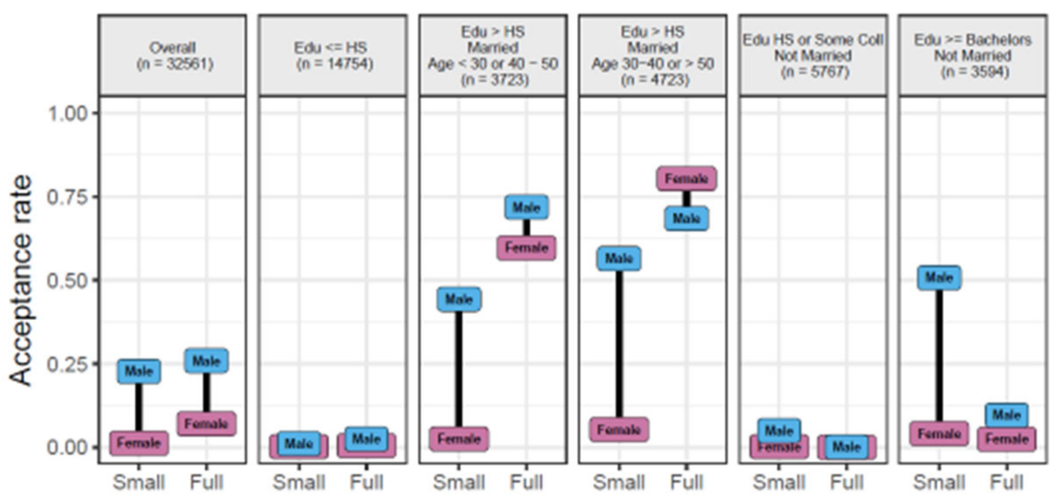
Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

12



13

Credit Scoring -- Models with and without marital status and race

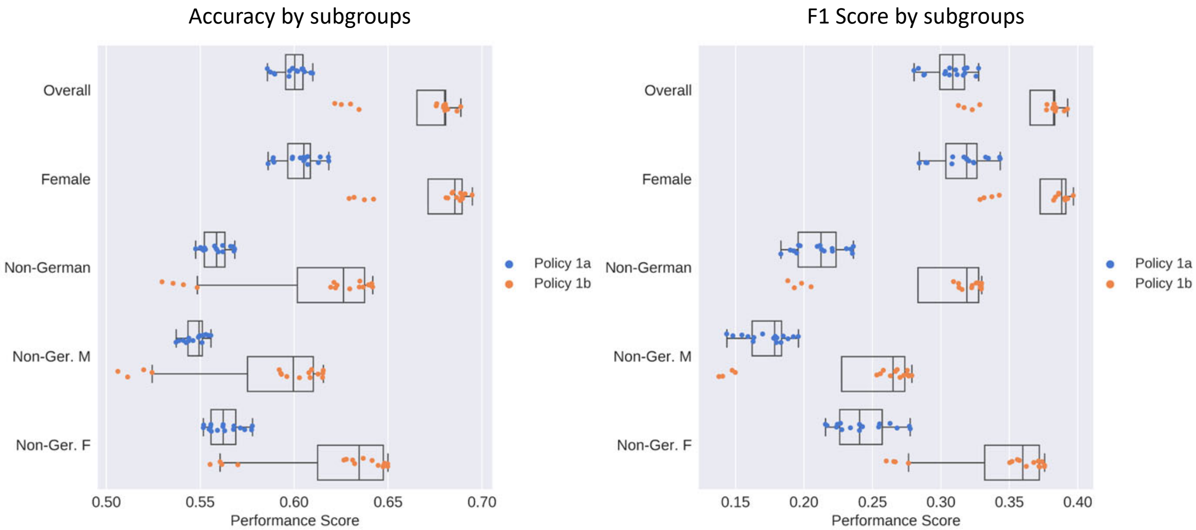


Chouldechova & G'Seel (2017). Fairer and more accurate but for whom? FATML Workshop, 8 2017, Halifax, NS, CA arXiv:1707.00046 CC-by NC 4.0

14

14

Profiling of Jobseekers



Accuracy = correctly classified observations (positive and negative)

F1 score = balancing precision and recall on the positive class

15

15

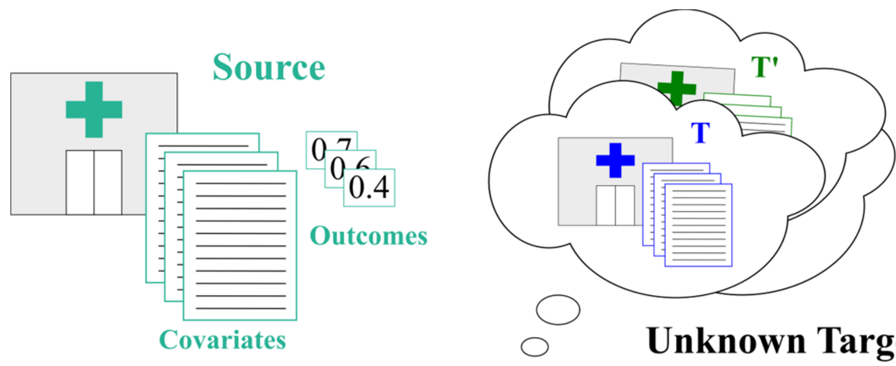
Learning Across Context

16

Inference Challenge: Robust across Populations

Single source → many different targets!

- *s: large longitudinal medical study run by GW hospital*
- *t: different hospital populations across the country*



17

Empirical Evaluation

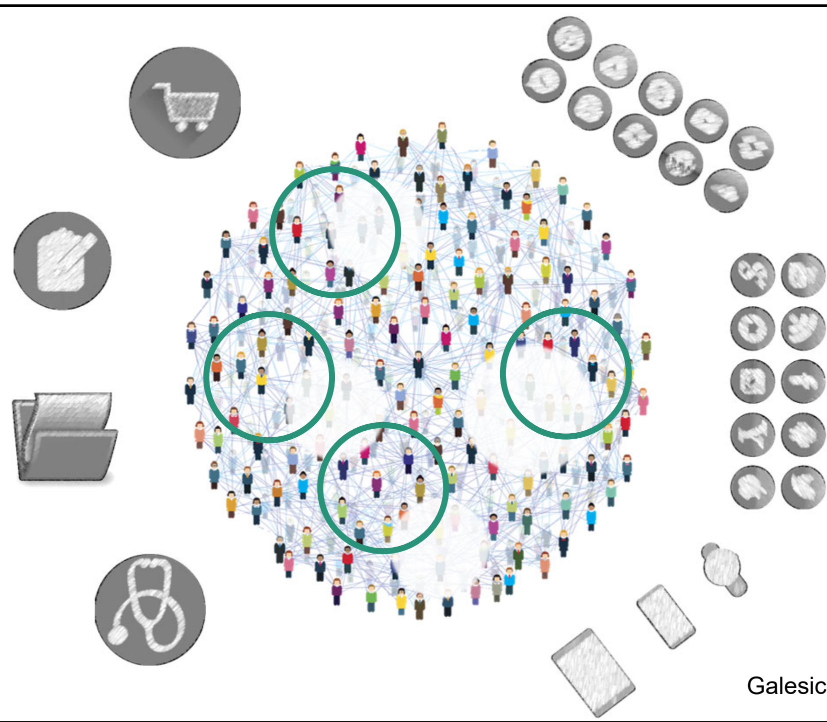
- Setting:
  - Source: US National Health and Nutrition Examination Survey
  - Target: US National Health Interview Survey
  - Estimate 15-year mortality rate across demographic groups
- Results:
  - Imputation with a single multicalibrated predictor
  - Similar performance as demographic-specific PS estimates

Kim, M. P., Kern, C., Goldwasser, S., Kreuter, F. and Reingold, O. (2022). Universal Adaptability: Target-Independent Inference that Competes with Propensity Scoring. Proceedings of the National Academy of Sciences of the United States of America (PNAS) 119(4). doi:10.1073/pnas.2108097119

18

Capturing the Unmeasured

19

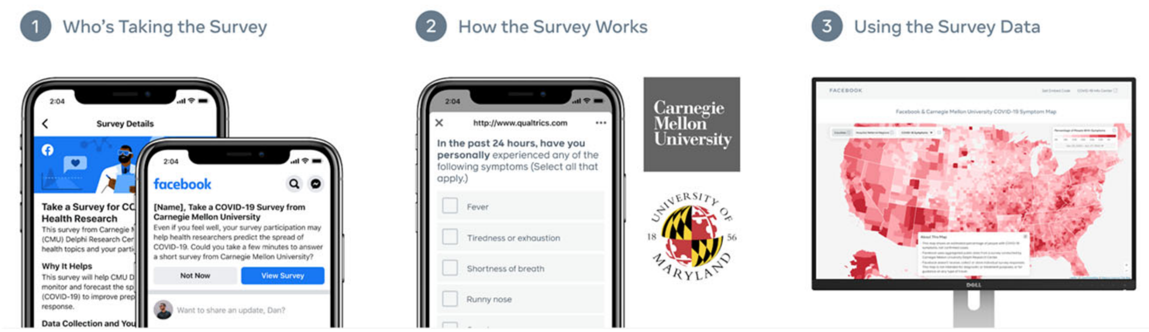


Galesic et al. 2021 Nature

20

Social Sensing and COVID-19 Infections

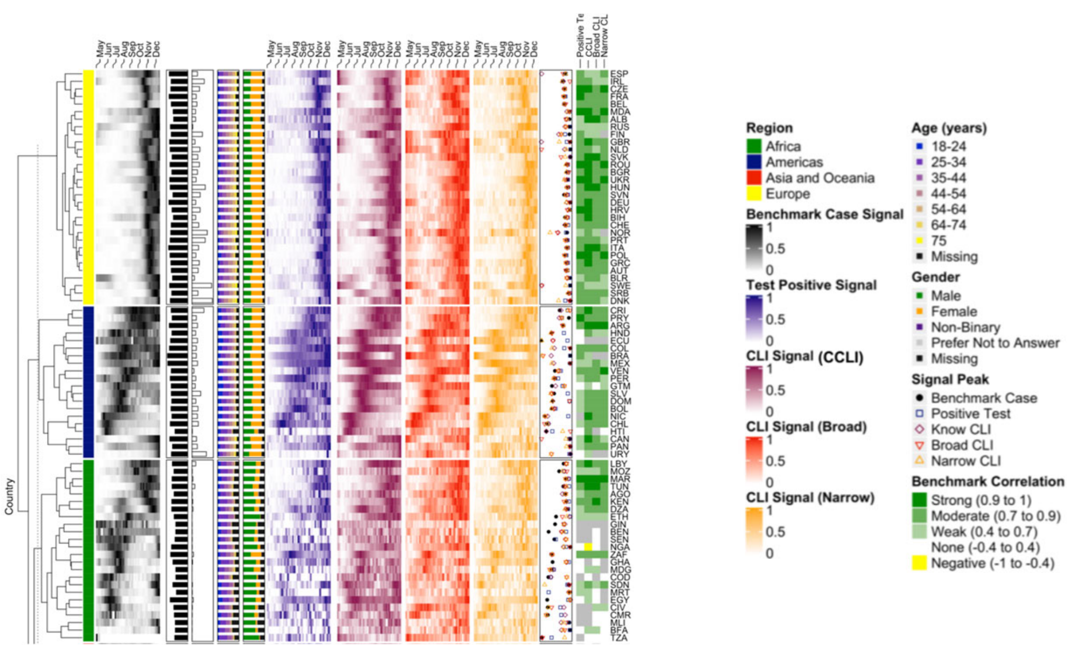
"Do you personally know anyone in your local community who is sick with a fever and either a cough or difficulty breathing?"



Global COVID-19 Trends and Impact Survey (CTIS) covidmap.umd.edu

21

21



Astley et al. 2021 PNAS and - medRxiv 2021.07.05.21259989; doi: https://doi.org/10.1101/2021.07.05.21259989

22

22

Keeping the Small Groups

23

COLERIDGE INITIATIVE

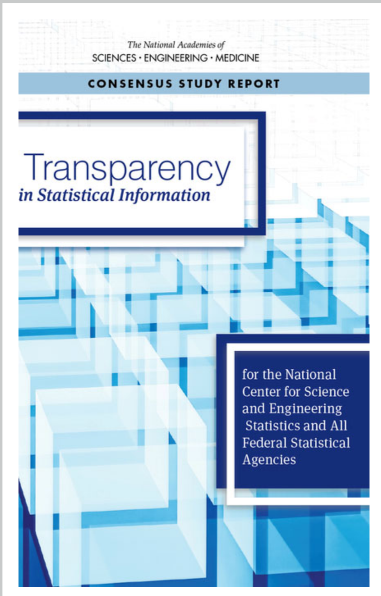
Home · Training · ADRF · Projects & Research · Collaborators · Events

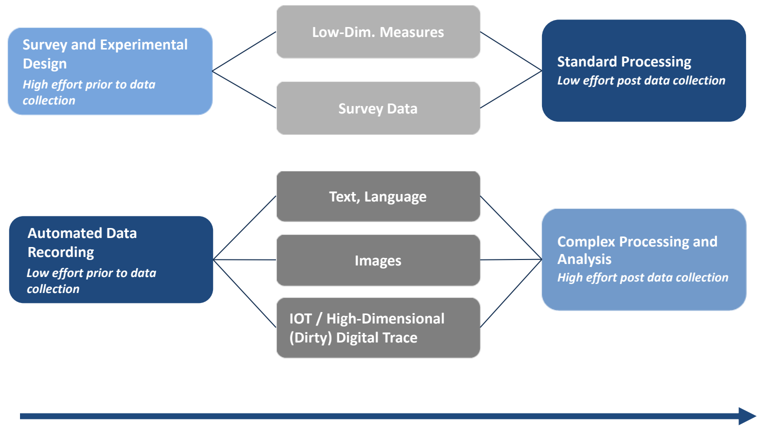
Safe projects · Safe people · Safe settings · Safe data · Safe exports

We follow these principles in Coleridge (see coleridgeinitiative.org/adrf/five-safes/)

24

24





25

25

Summary

- Training data often lack representation. Bad predictions are often not detectable
- Despite large data streams, feature generation not robust. Measurements are erroneous and often do not fit the underlying concept. Humans don't operate in stable environments
- We need to actively work against a lack of diversity and our learned habit of looking at averages. This not only includes thinking of how to represent or include missing cases, but also to allow full transparency to the decisions we make along the way

26

26