Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Bias in AI

Davide Cirillo

Machine Learning for Biomedical Research

Life Sciences Department

03/05/2022

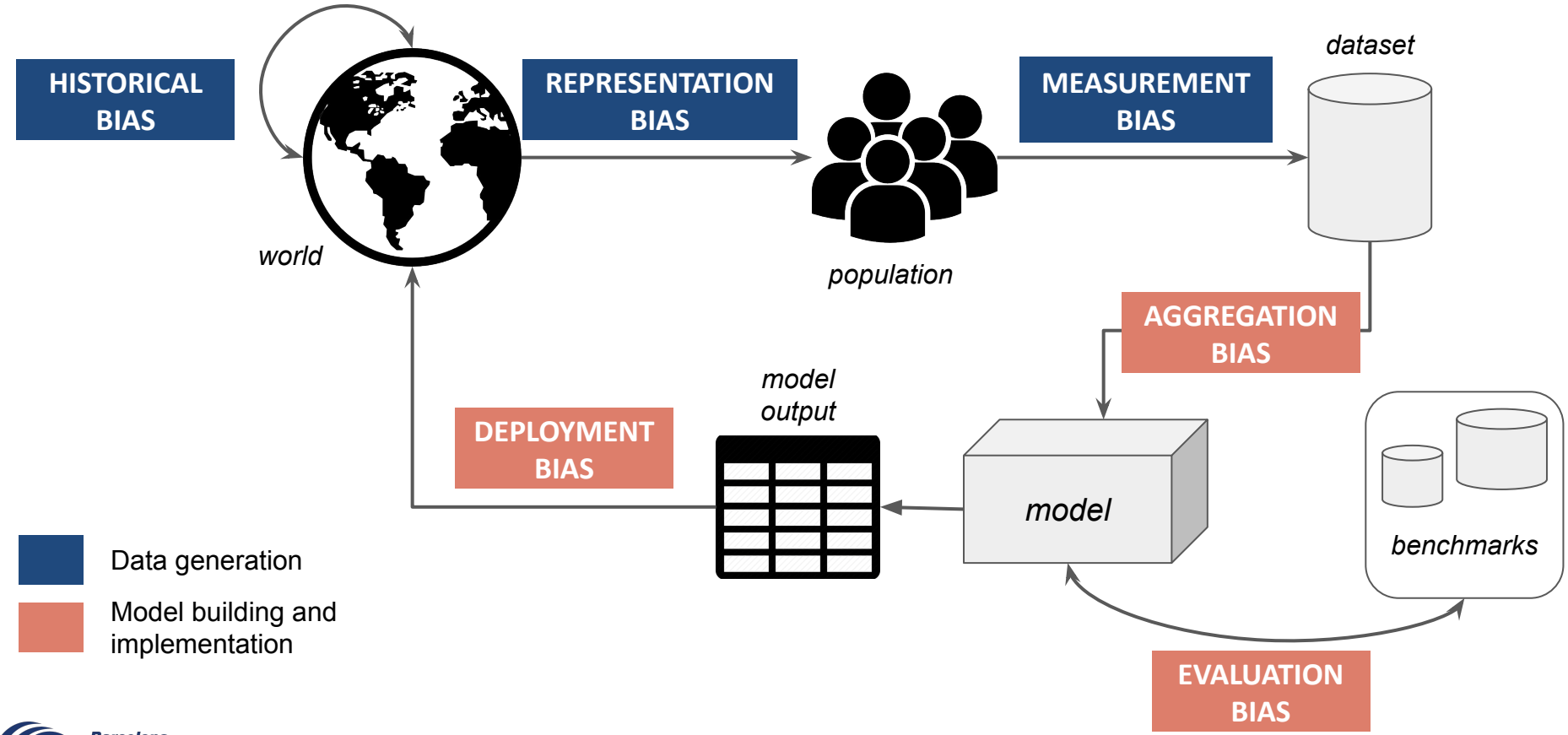
Depression in men and women

Sex, gender and social roles

- **1 in 8 women** experiences depression in her lifetime - twice as many as men, regardless of race or ethnicity.
- Men represent **3.5 times the number of suicides** than women
- **Sex and gender differences** in signs of depression:
  - women: crying, sadness, loss of interest, verbal expression of suicidal thoughts
  - men: outbursts of anger and / or aggression, substance abuse, and risk-taking

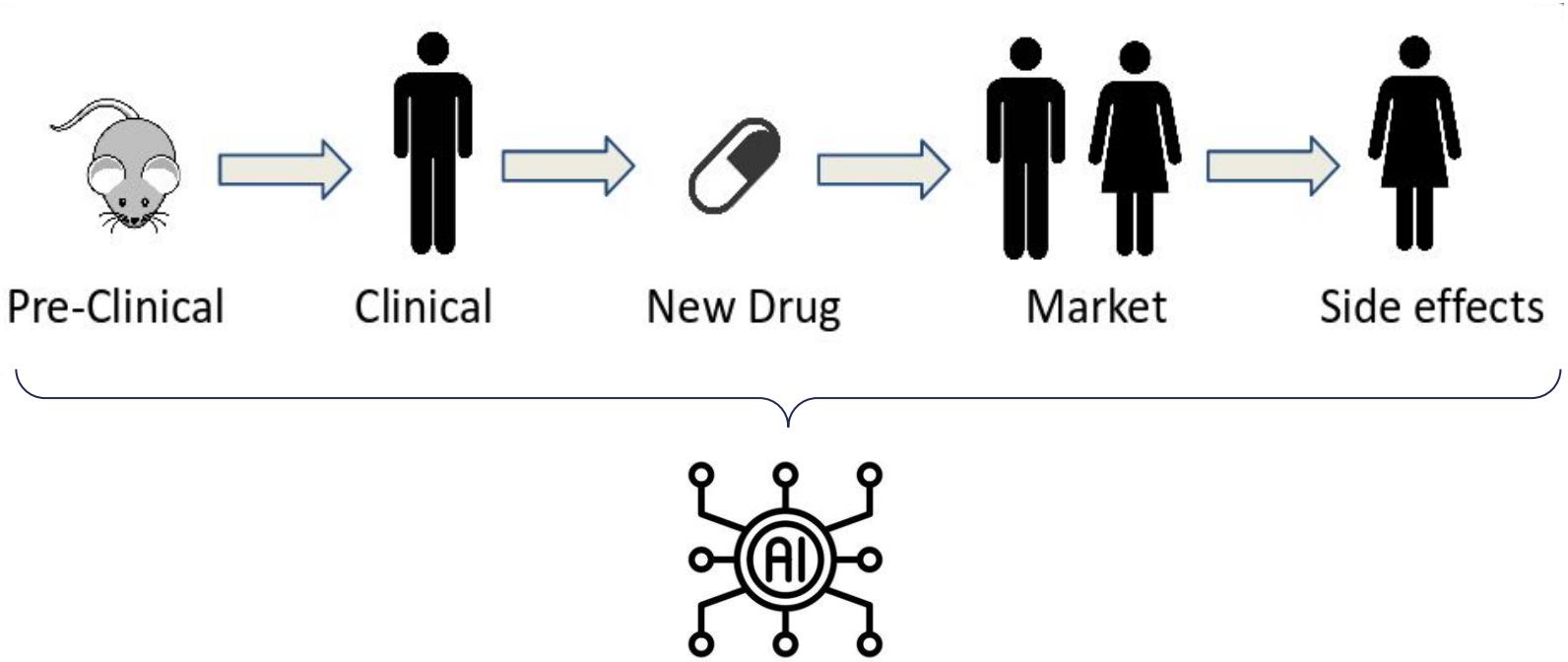
source: The National Alliance on Mental Illness (NAMI)

Biases in data generation and modeling

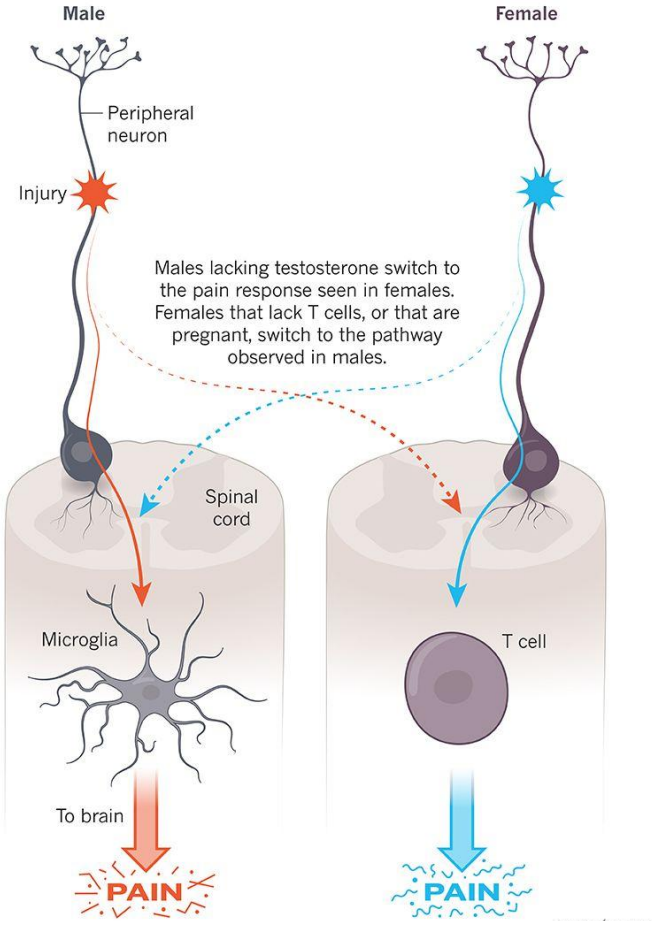


Adapted from: Surech & Guttag. arXiv:1901.10002 (2020)

Generation of biased data in medicine

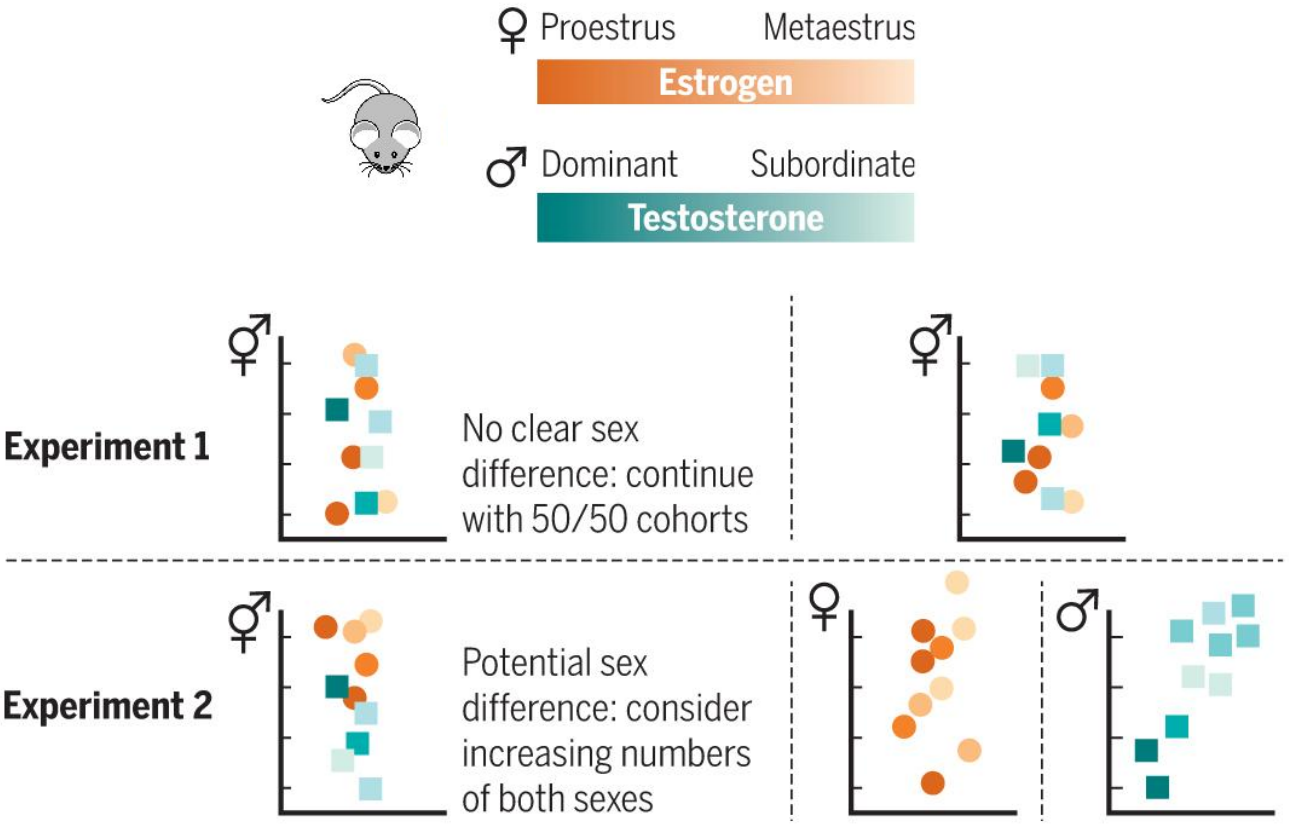


Pre-clinical studies



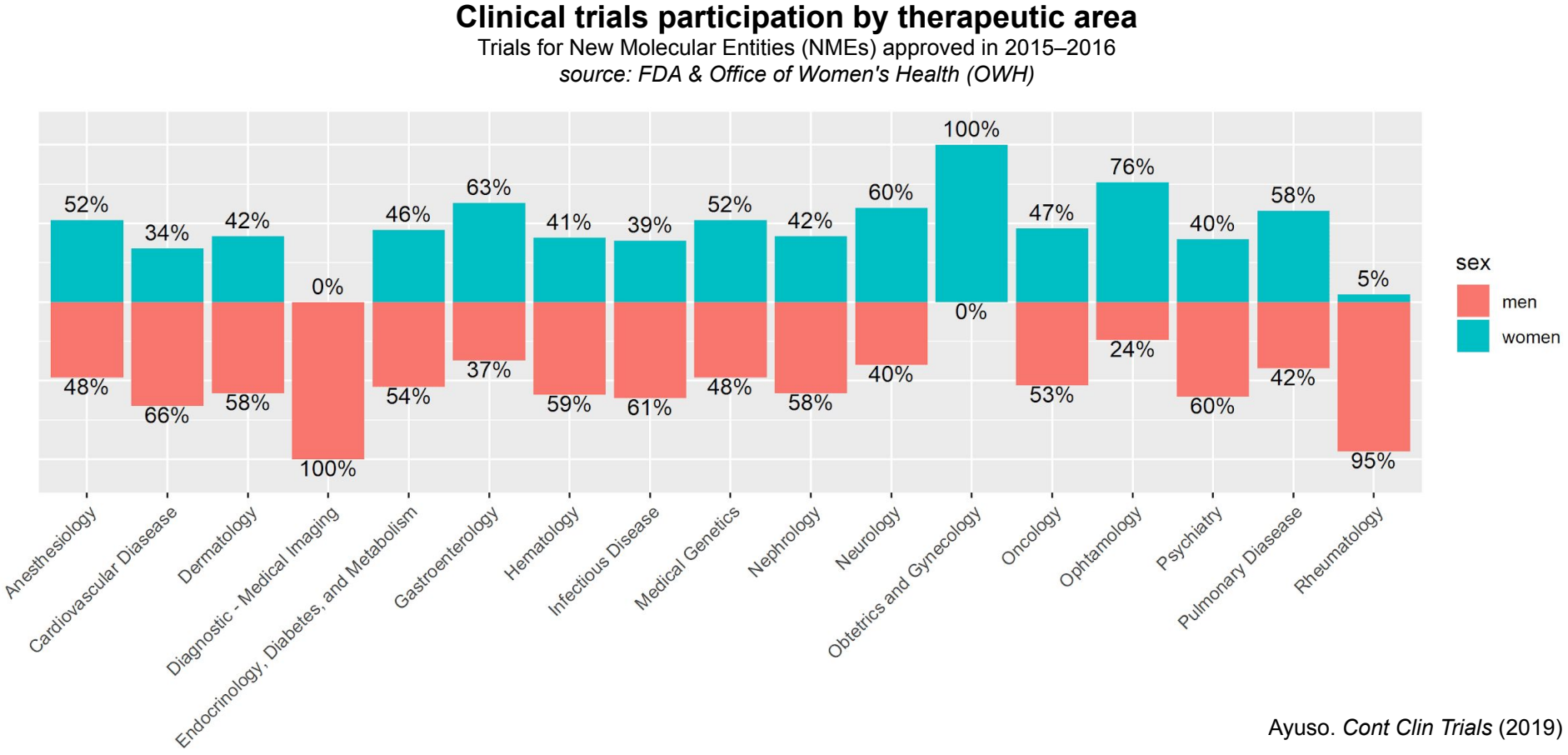
Sorge et al. *Nature Neuroscience* (2015)

Pre-clinical studies



Shansky. *Science* (2019)

Clinical trials



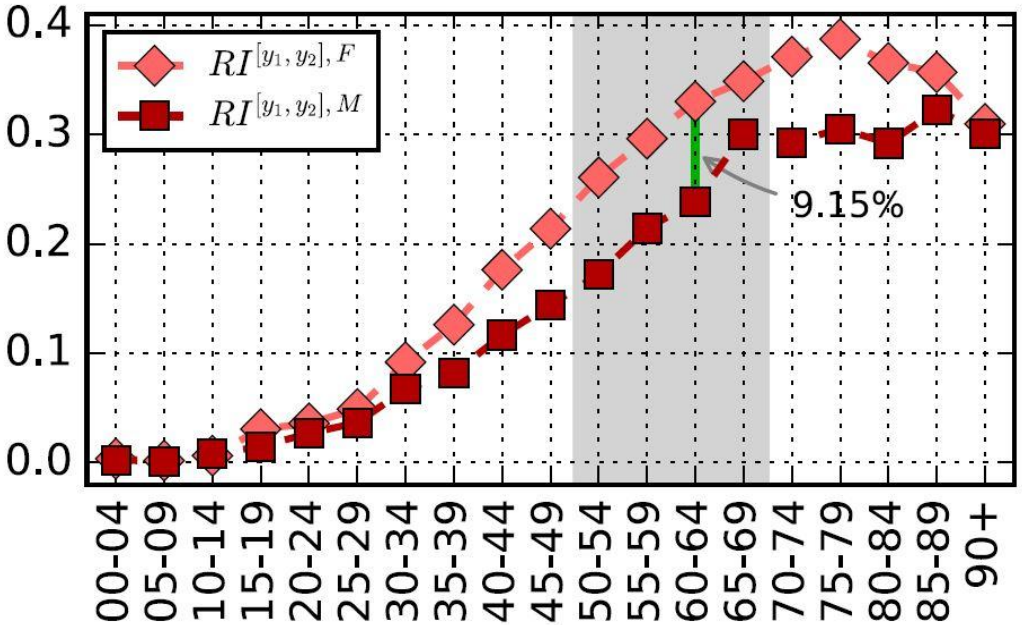
Ayuso. *Cont Clin Trials* (2019)

Pharmacovigilance

Drug–drug interactions in EHRs

Large-scale longitudinal study (18 months)

Population 338k (Blumenau, Brazil)



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

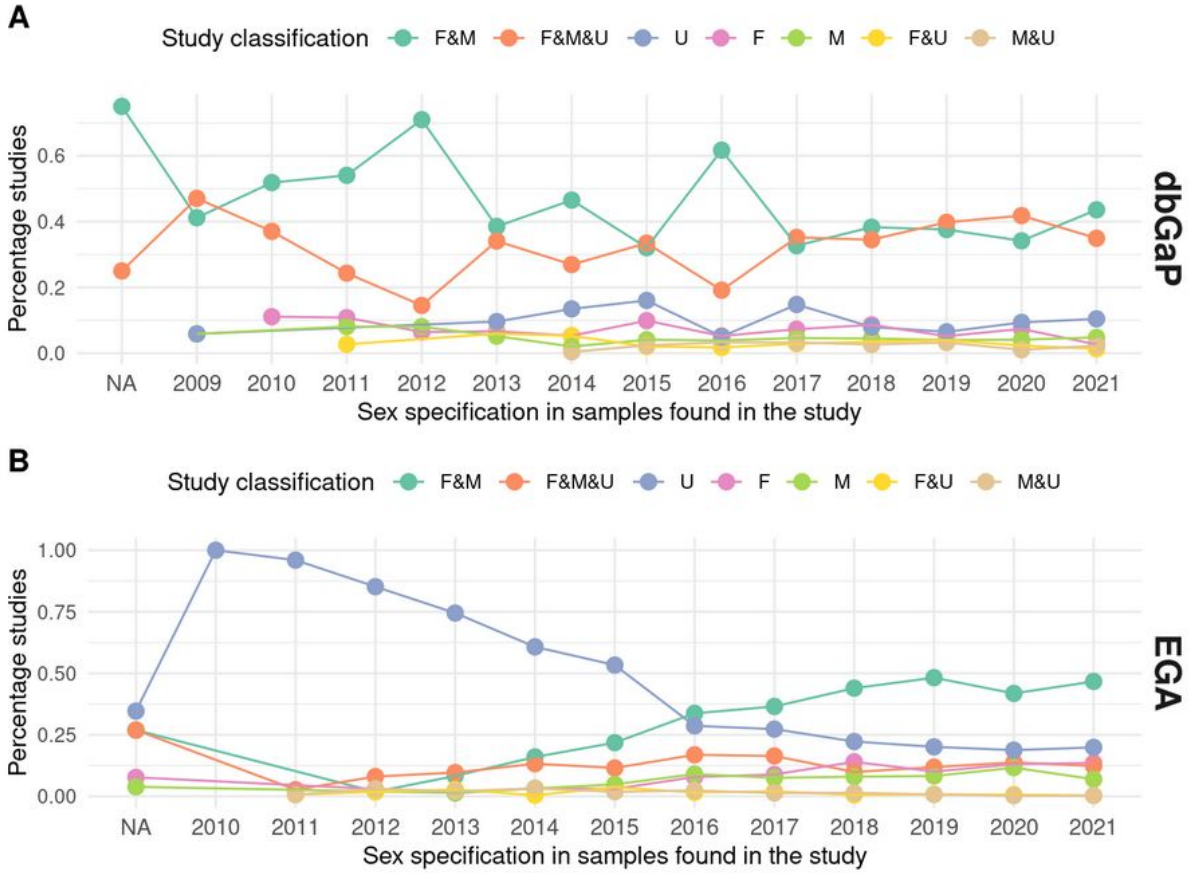
Correia et al. *NPJ Digital Medicine* (2019)

Table 3. Prescription drugs withdrawn from the US market (1997–2000).

Drug	Type of drug	Patient population	Primary health risk
Prescription drugs with evidence of greater health risks in women			
Pondimin®	Appetite suppressant	Women	Valvular heart disease
Redux®	Appetite suppressant	Women	Valvular heart disease
Rezulin®	Diabetic	Women	Liver failure
Lotronex®	Gastrointestinal	Women	Ischemic colitis
Seldane®	Antihistamine	Women and men	Torsades de Pointes
Posicor®	Cardiovascular	Women and men	Lowered heart rate in elderly women and adverse interactions with 26 other drugs
Hismanal®	Antihistamine	Women and men	Torsades de Pointes
Propulsid®	Gastrointestinal	Women and men	Torsades de Pointes
Prescription drugs without evidence of greater health risks in women			
Raxar®	Antibiotic	Women and men	Torsades de Pointes
Duract®	Analgesic and anesthetic	Women and men	Liver failure

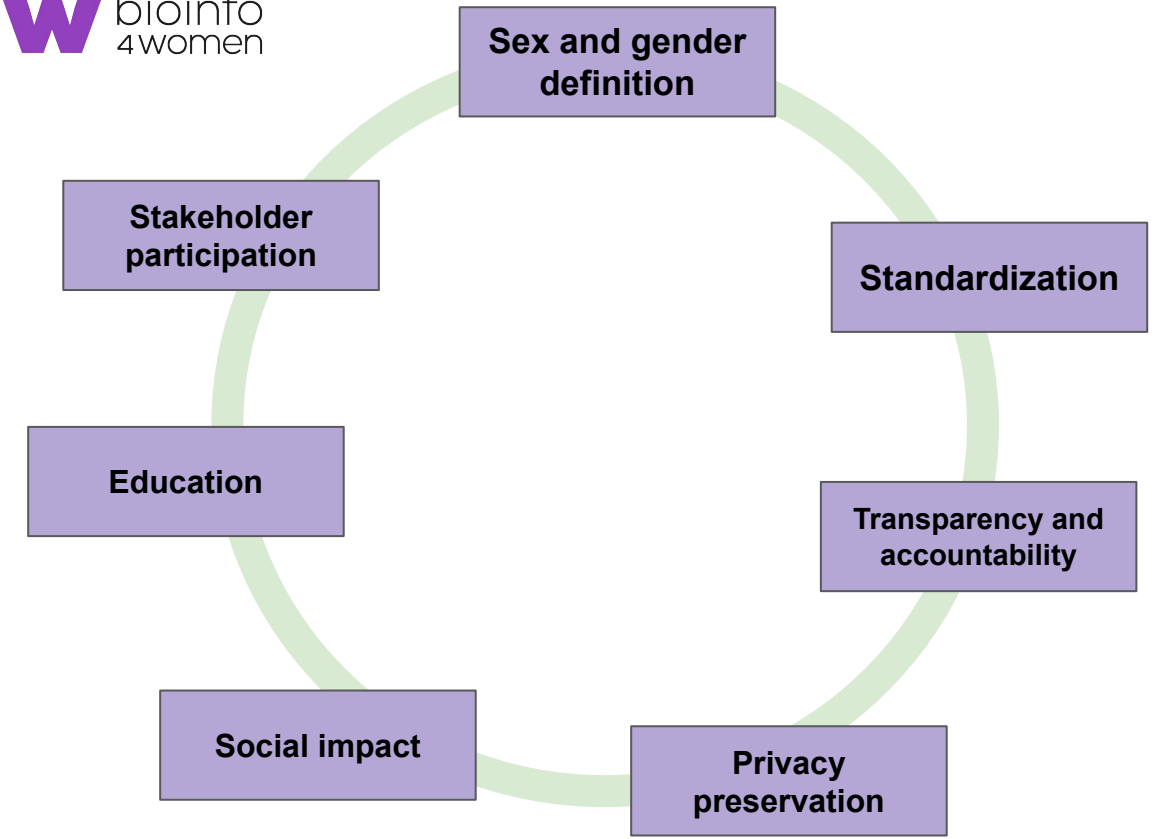
Parekh et al. *Expert Rev. Clin. Pharmacol.* (2011)

Sex bias in large-scale human data repositories worldwide



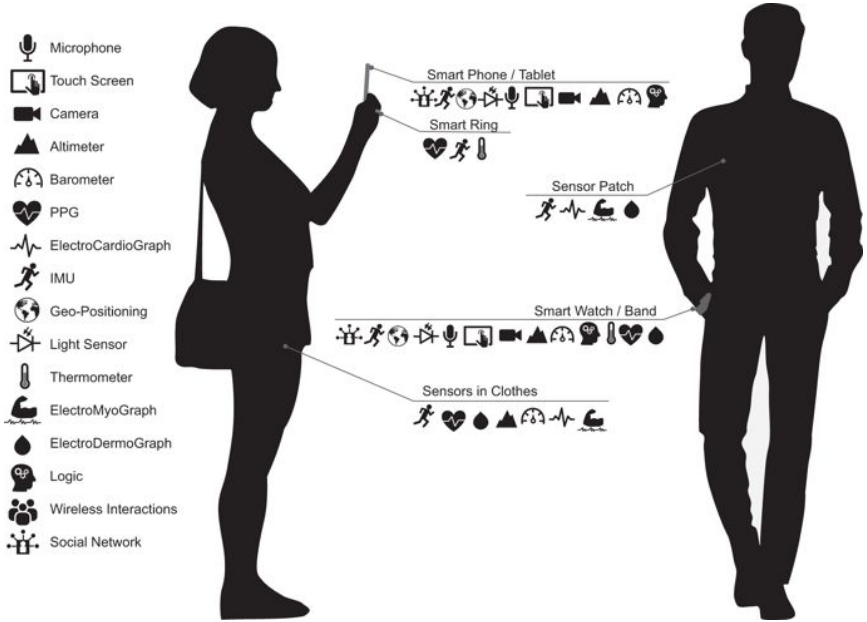
Ruiz Serra et al. 2022 [in preparation]

bioinfo 4women

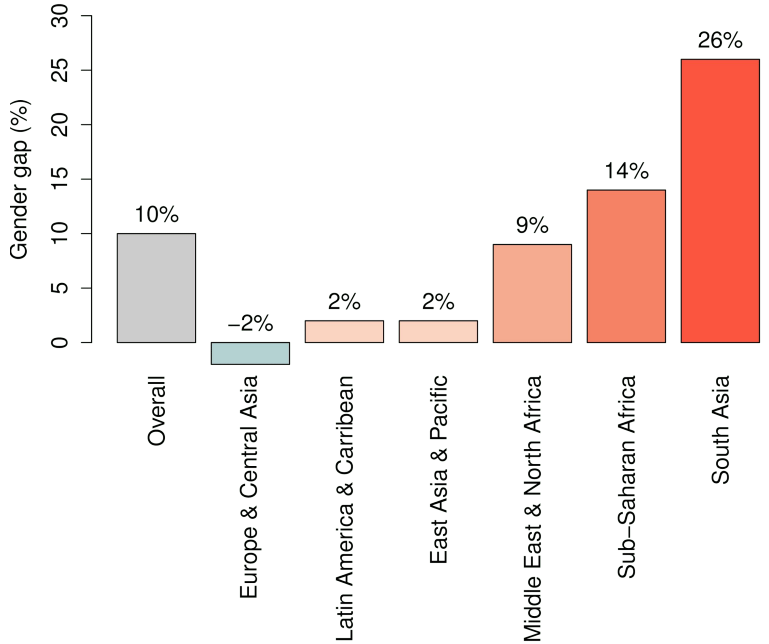


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Digital biomarkers and digital divide



How less likely a woman is to own a mobile phone than a man

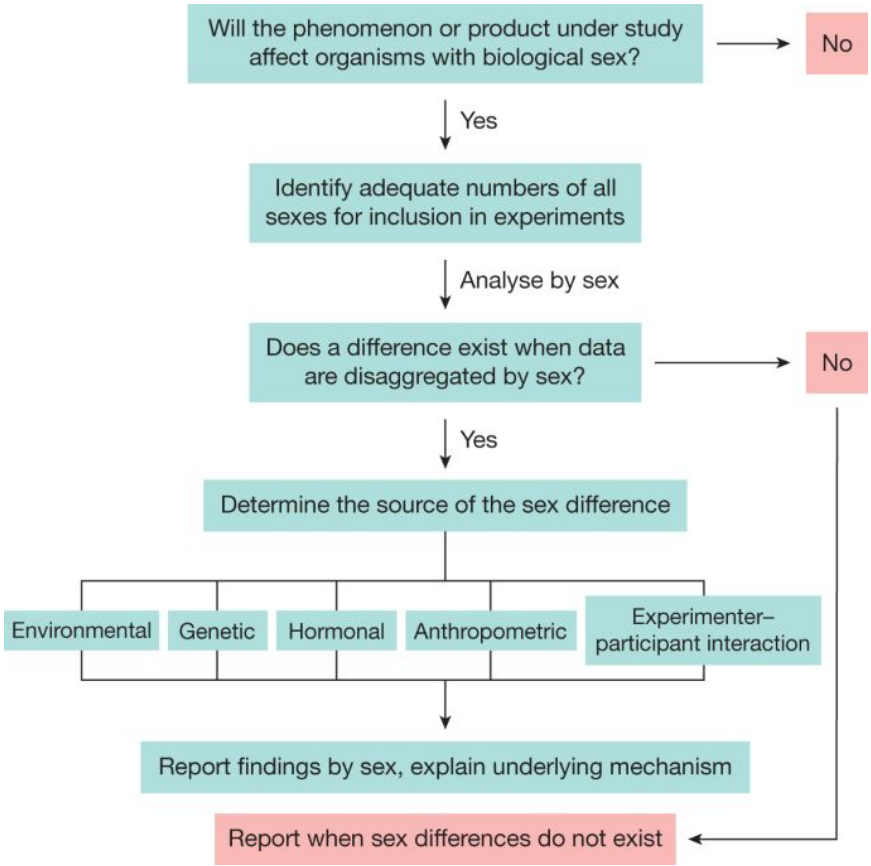


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

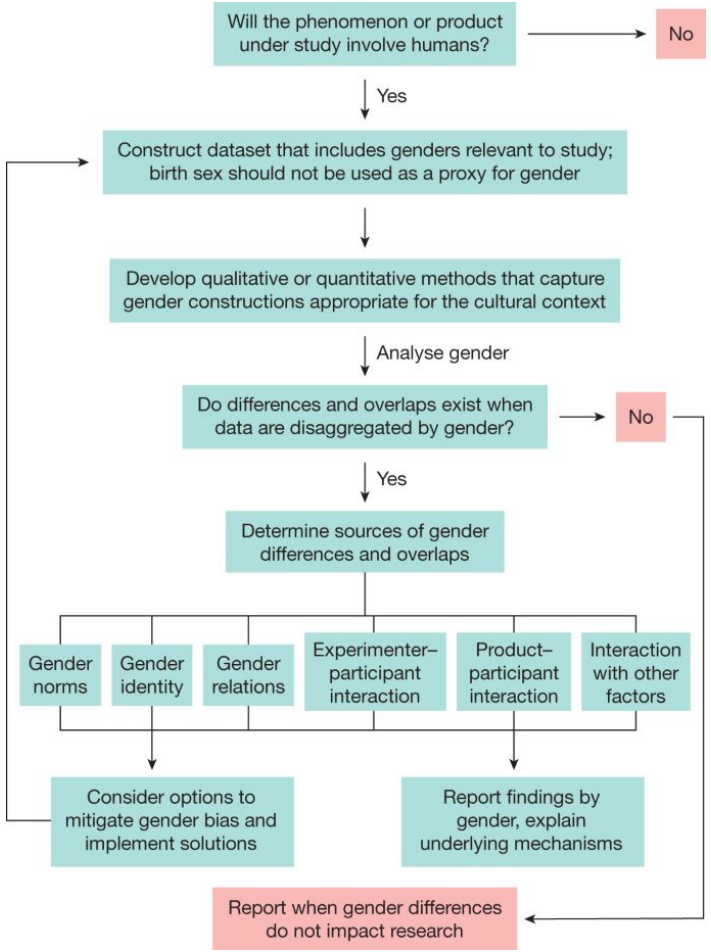
Rowntree. *GSMA* 2018

Gendered Innovations in Science, Health & Medicine, Engineering, and Environment

Sex analysis



Gender analysis



Tannenbaum, C., Ellis, R.P., Eyssel, F. et al. Sex and gender analysis improves science and engineering. Nature 575, 137–146 (2019). https://doi.org/10.1038/s41586-019-1657-6

bioinfo 4women

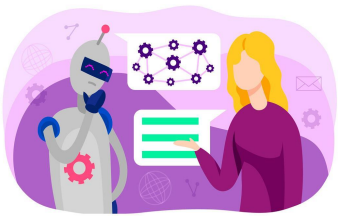
WOMEN'S BRAIN PROJECT®

Sex and gender differences and biases in Artificial Intelligence for Biomedicine and Healthcare

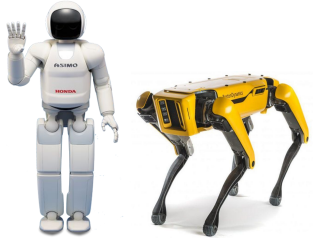
Cirillo, Catuara-Solarz et al. NPJ Digit Med. 2020

Big Data Analytics

NLP



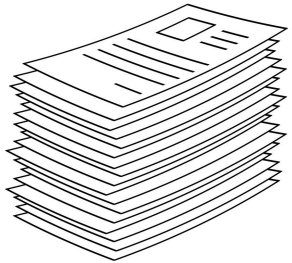
Robotics

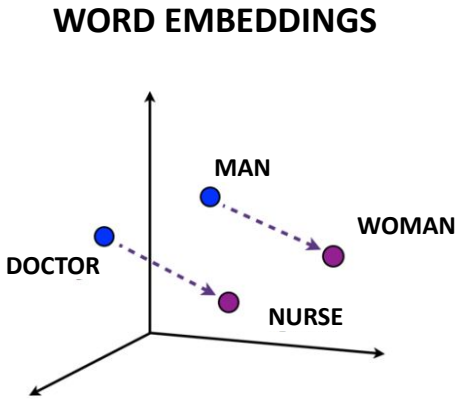


Digital medicine

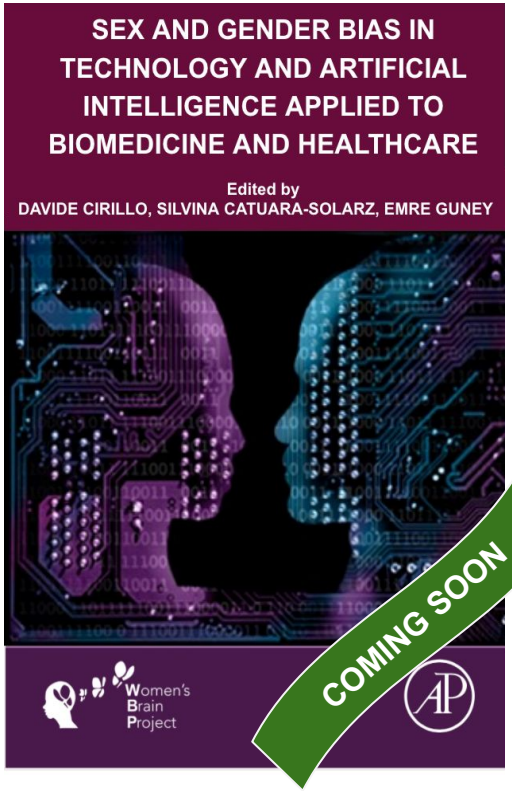
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Natural Language Processing





Caliskan et al. *Science* (2017)
Bolukbasi et al. *NIPS* (2016)
Marta R. Costa-jussá. *Nature Medicine* (2019)



"It's a must read for anyone working in this area; the materials presented here should be integrated into medical school curricula."

Londa Schiebinger
John L. Hinds Professor of History of Science, Stanford University
Founding director, Gendered Innovations in Science, Health & Medicine, Engineering, and Environment

ISBN: 9780128213926

ELSEVIER

bioinfo 4women

Palau Macaya

"la Caixa" Foundation

LA VANGUARDIA

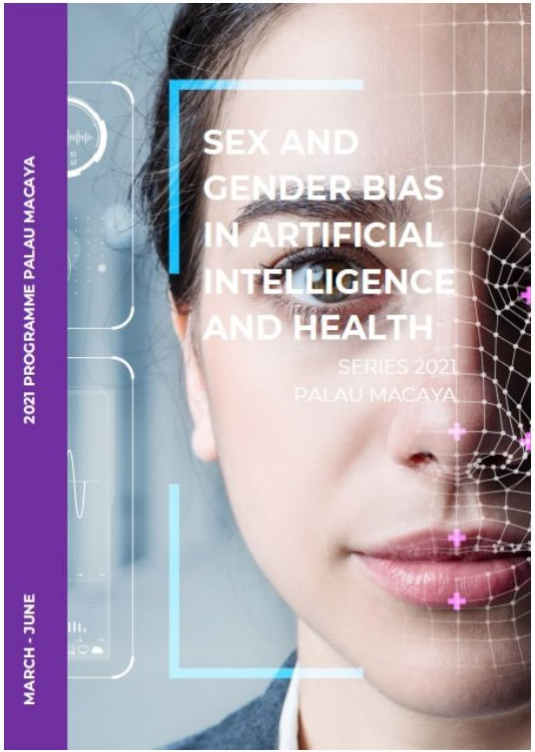
CAT-FUNDACIÓN LA CAIXA

El Palau Macaya impulsará once proyectos sobre innovación y sostenibilidad

REDACCIÓN
27/07/2020 16:52

Barcelona, 27 jul (EFE).- El Palau Macaya de la Fundació "la Caixa" albergará once proyectos innovadores y sostenibles seleccionados por la fundación bancaria y que se desarrollarán en el centro durante el curso que va de septiembre de 2020 a junio de 2021.

Palau Macaya en Barcelona (Archivo) - OBRA SOCIAL LA CAIXA - Archivo
BARCELONA, 27 Jul. (EUROPA PRESS) -



United Nations

10 REDUCED INEQUALITIES

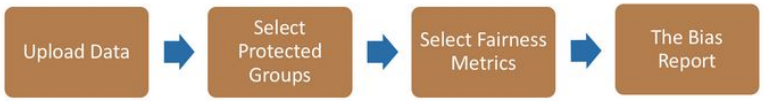
3 GOOD HEALTH AND WELL-BEING

5 GENDER EQUALITY

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Buslón, N,; Rementeria, M.; Bioinfo4Women Programme. (2021). https://doi.org/10.5281/zenodo.5513939

Open source toolkit for bias detection and mitigation



Aequitas
Bias & Fairness Audit

Home Code About

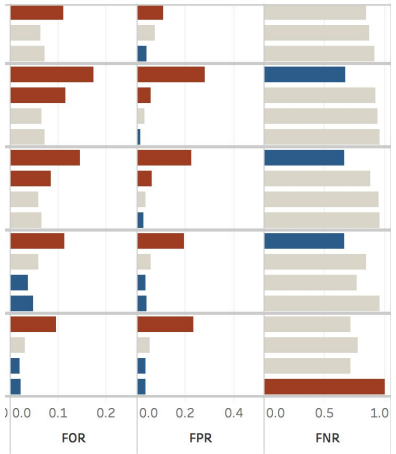
The Bias Report in Action

Using a clean version of the COMPAS dataset, we demostrate the use of The Bias Report web app. Click below for background on the dataset, a description of the process, and analysis.

Details about the report

The Bias Report

Audit Date:	04 Jun 2018
Data Audited:	7214 rows
Attributes Audited:	race
Audit Goal(s):	False Positive Rate Parity - Ensure all protected groups have the same false positive rates as the reference group).
	False Discovery Rate Parity - Ensure all protected groups have equally proportional false positives within the selected set (compared to the reference group).
	False Negative Rate Parity - Ensure all protected groups have the same false negative rates (as the reference group).
Reference Groups:	Custom group - The reference groups you selected for each attribute will be used to calculate relative disparities in this audit.
Fairness Threshold:	80%. If disparity for a group is within 80% and 125% of the value of the reference group on a group metric (e.g. False Positive Rate), this audit will pass.



Web Audit Tool

Python Library

Command Line Tool

https://dsapp.uchicago.edu/projects/aequitas/

Open source toolkit for bias detection and mitigation

AI Fairness 360 - Demo

Data Check Mitigate Compare

3. Choose bias mitigation algorithm

A variety of algorithms can be used to mitigate bias. The choice of which to use depends on whether you want to fix the data (pre-process), the classifier (in-process), or the predictions (post-process). Learn more about how to choose.

Reweighing

Weights the examples in each (group, label) combination differently to ensure fairness before classification.

Data → Classifier → Predictions

Optimized Pre-Processing

Learns a probabilistic transformation that can modify the features and the labels in the training data.

Data → Classifier → Predictions

Adversarial Debiasing

Learns a classifier that maximizes prediction accuracy and simultaneously reduces an adversary's ability to determine the protected attribute from the predictions. This approach leads to a fair classifier as the predictions cannot carry any group discrimination information that the adversary can exploit.

Data → Classifier → Predictions

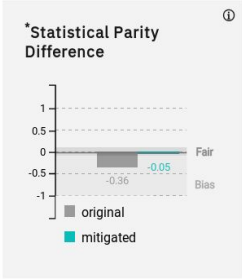
Reject Option Based Classification

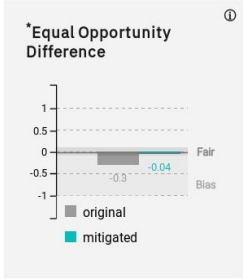
Changes predictions from a classifier to make them fairer. Provides favorable outcomes to unprivileged groups and unfavorable outcomes to privileged groups in a confidence band around the decision boundary with the highest uncertainty.

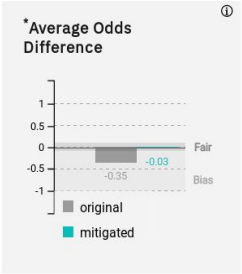
Data → Classifier → Predictions

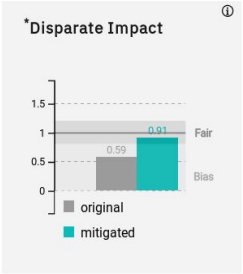
Protected Attribute: Sex

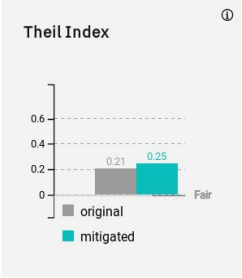
Privileged Group: ***Female***, Unprivileged Group: ***Male***





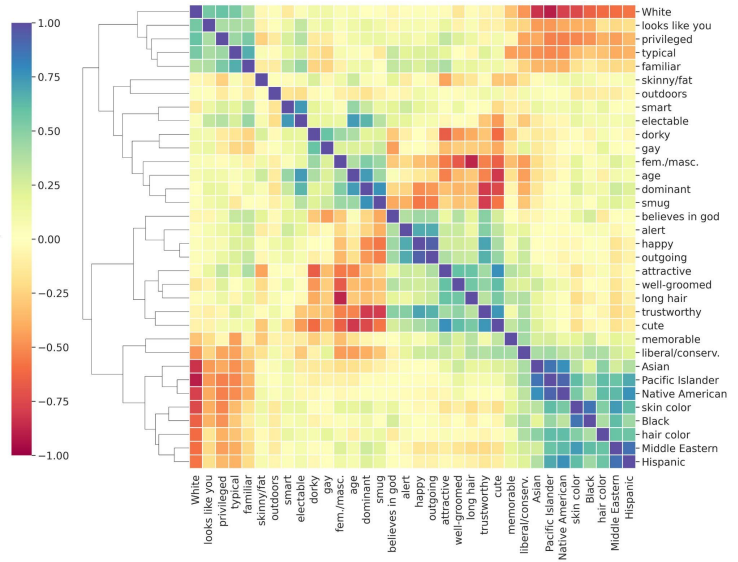






https://aif360.mybluemix.net

Creating a deliberately biased algorithm



Peterson JC et al. “Deep models of superficial face judgments.” PNAS vol. 119,17 (2022): e2115228119. doi:10.1073/pnas.2115228119

Recommended Twitter thread by Iris Van Rooij (Radboud University Nijmegen)

https://twitter.com/IrisVanRooij/status/1519055634607616003


BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación
EXCELENCIA SEVERO OCHOA
Thank you
bioinfo 4women
Juan de la Cierva Programa
GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD
davide.cirillo@bsc.es