Expanding on Gender Diversity Report :
NamSor algorithms for classification of names by "Race"/ethnicity or cultural origin/diasporas

Gender, 'race'/ethnicity or origin bias in AI ?

Algorithms are used to 'assist' human decision in funnel-based processes, ex.

- recruitment,

- credit allocation,

…

AI especially used in the early stage of the selection process (ex resume sourcing or screening) : search, scoring, tagging …
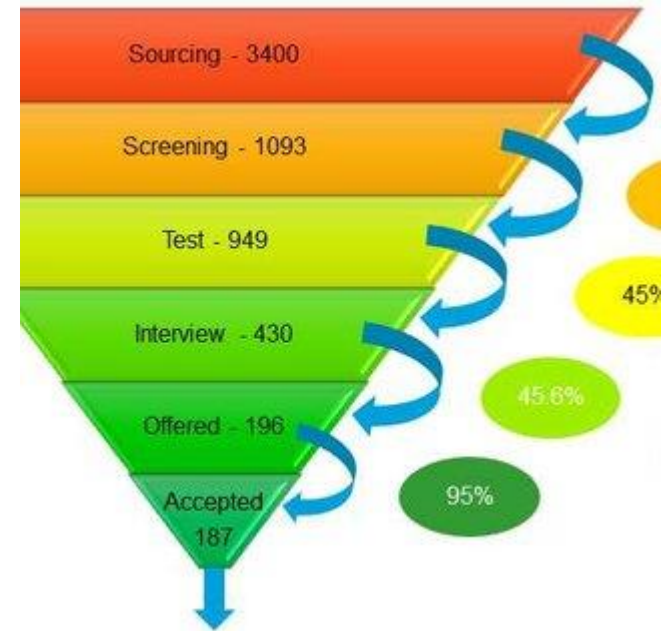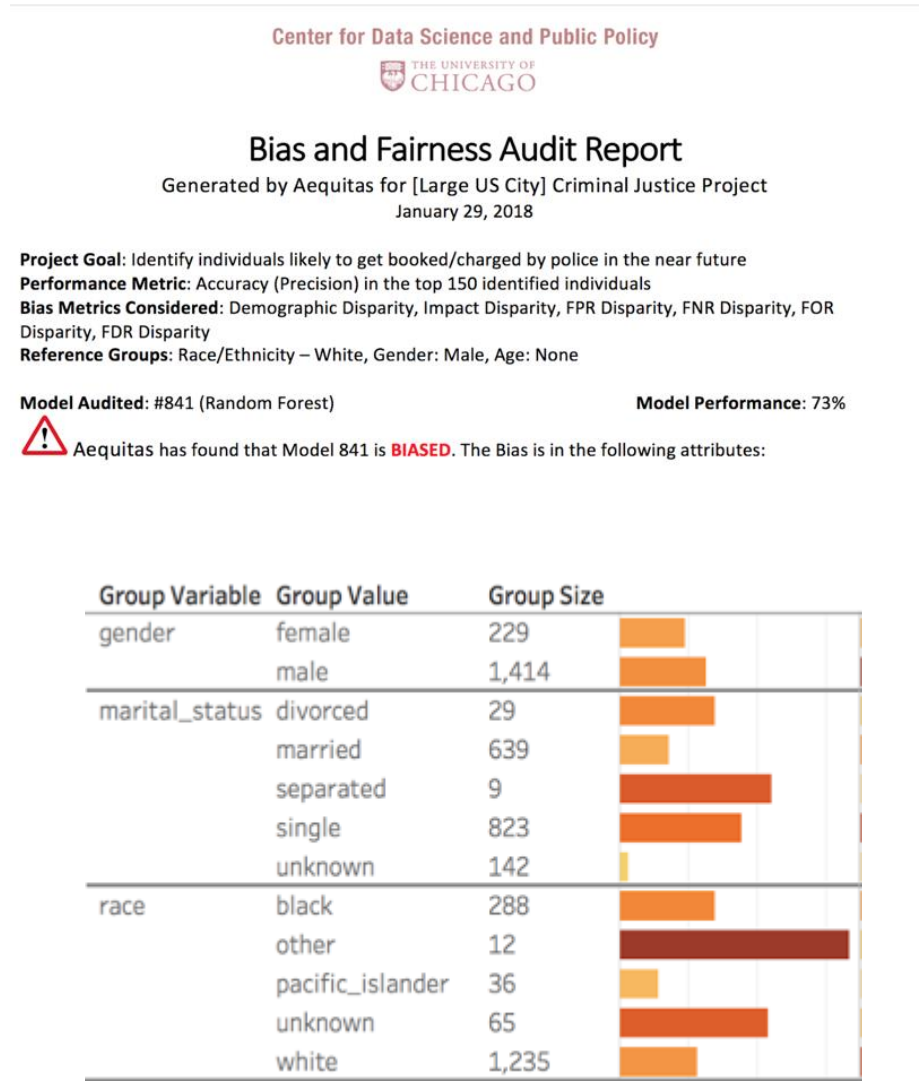
Is the algorithm FAIR?

Estimating gender, racial/ethnic bias in algorithms ex. recruitment

Two approaches :

1) Use Aequitas, an open source bias audit toolkit developed by the Center for Data Science and Public Policy at University of Chicago

2) Measure changes in diversity index (Shannon or Simpson) at each selective step

What taxonomy for diversity analytics? What is "race"/ethnicity ?

Center for Data Science and Public Policy

THE UNIVERSITY OF CHICAGO

Bias and Fairness Audit Report
Generated by Aequitas for [Large US City] Criminal Justice Project
January 29, 2018

Project Goal: Identify individuals likely to get booked/charged by police in the near future
Performance Metric: Accuracy (Precision) in the top 150 identified individuals
Bias Metrics Considered: Demographic Disparity, Impact Disparity, FPR Disparity, FNR Disparity, FOR Disparity, FDR Disparity
Reference Groups: Race/Ethnicity – White, Gender: Male, Age: None

Model Audited: #841 (Random Forest)   **Model Performance**: 73%

⚠ Aequitas has found that Model 841 is **BIASED**. The Bias is in the following attributes:

Group Variable	Group Value	Group Size	
gender	female	229	
	male	1,414	
marital_status	divorced	29	
	married	639	
	separated	9	
	single	823	
	unknown	142	
race	black	288	
	other	12	
	pacific_islander	36	
	unknown	65	
	white	1,235	

NamSor sorts *Names*

Names reflect cultural *Identity*

Since 2012, **NamSor** data mining software recognizes the **linguistic or cultural origin of names** in any alphabet / language, using both supervised and unsupervised machine learning (ie. clustering).

2014 : launch Gender API v1

2018 : software is re-written from scratch with standard ML frameworks : 1/ name embedding + neural networks 2/ naïve bayes classifier

2019 : launch NamSor API v2 with Gender, US 'Race'/Ethnicity, Country/Origin/Diaspora classifiers

Our proud contribution to Gender Reports

- NamSor Gender API (v1) was used independently by both by Science-Metrix and Elsevier in 2015 and 2017
- NamSor Gender API V2 was used for 'The Researcher Journey Through a Gender Lens' and we've made specific improvements :
  - Enhanced probability estimates for gender inference
  - Improved support for East-Asian names (Chinese, Korean, Japanese)

Gender diversity is just one dimension, there are many other …

An artistic illustration of ethnic diversity / diversity of origin among COVID-19 scientists
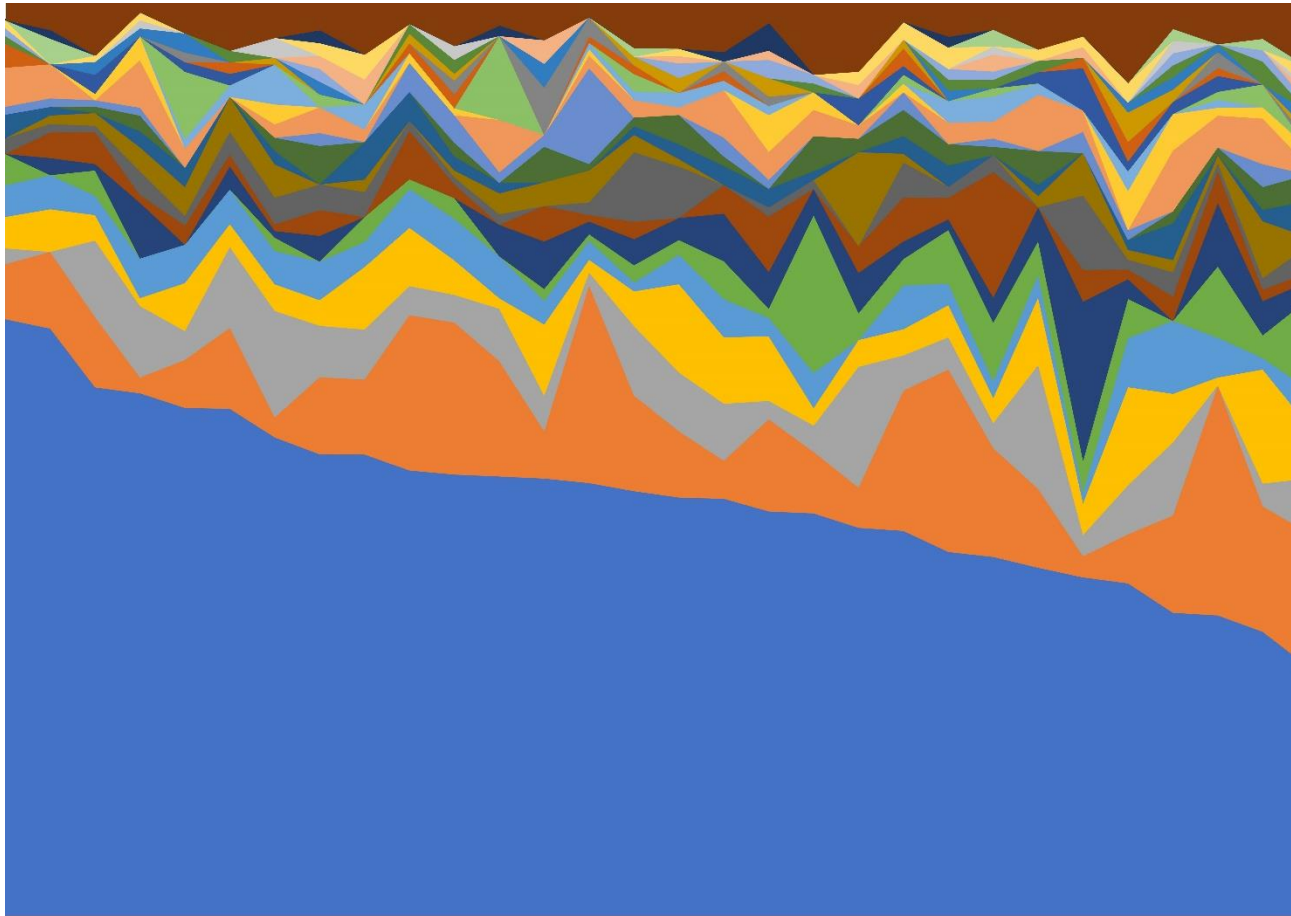
"Chinese sea" at Ars Electronica 2020 by Dario Rodighiero (Harvard Metalab, https://github.com/rodighiero/COVID-19), Eveline Wandl-Vogt (Austrian Academy of Science) and Elian Carsenat (NamSor)

NamSor CORE taxonomies

- NamSor API is available and already supports robust, fine-grained taxonomies for
  - Gender
  - US 'Race'/Ethnicity
  - Country/Origin
  - Diaspora

- We have other taxonomies available for academic projects.

.

Taxonomy

Gender

Classes

Male   Female

Gender classification model infers the likely gender, with probability :

Field	Example	Description
id	ref12315	The input identifier
firstName	John	The input given name / firstName
lastName	Smith	The input family name / surname / lastName
likelyGender	male	The likely gender : male or female
probabilityCalibrated	0.99	The calibrated probability : 0.5 is Unknown, +1 is sure
genderScale	-0.99	The scale is -1..0..+1 and is based on the probability (Probability = 0.5 -> Scale = 0; Gender = Male & Probabilty = 1 -> Scale = -1; Gender = Female & Probability = 1 -> Scale = +1)
score	41	A non calibrated Score (use Probability instead) : score = Math.log(getProbaFirst() / getProbaNotFirst()) maxed to 100

Taxonomy

US Census 'Race'/Ethnicity

4 Classes or 6 classes*

| W_NL (White) | B_NL (Black) | HL (Hispano-Latino) | A (Asian) |

US 'Race'/Ethnicity classifies names by race/ethnicity according to US 'Census' taxonomy, along with probabilities.

Field	Example	Description
id	ref12315	The input identifier
firstName	Mary	The input first name / given name
lastName	Cao	The input last name / surname
countryIso2	US	The country of residence, the host country (ex. US, CA, NZ, GB)
raceEthnicity	A	The likely 'race'/ethnicity : W_NL (white, non latino), HL (hispano latino), A (asian, non latino), B_NL (black, non latino)
raceEthnicityAlt	W_NL	The best alternative 'race'/ethnicity
raceEthnicitiesTop	A, W_NL, ...	The likely 'race'/ethnicities
probabilityCalibrated	0.91	The calibrated probability of having guessed right the 'race'/ethnicity as A (Asian)
probabilityCalibratedAlt	0.95	The calibrated probability of having guessed right the 'race'/ethnicity as either A or W_NL (White Non Latino)

*add header X-OPTION-USRACEETHNICITY-TAXONOMY: USRACEETHNICITY-6CLASSES for two additional classes, AI_AN (American Indian or Alaskan Native) and PI (Pacific Islander).
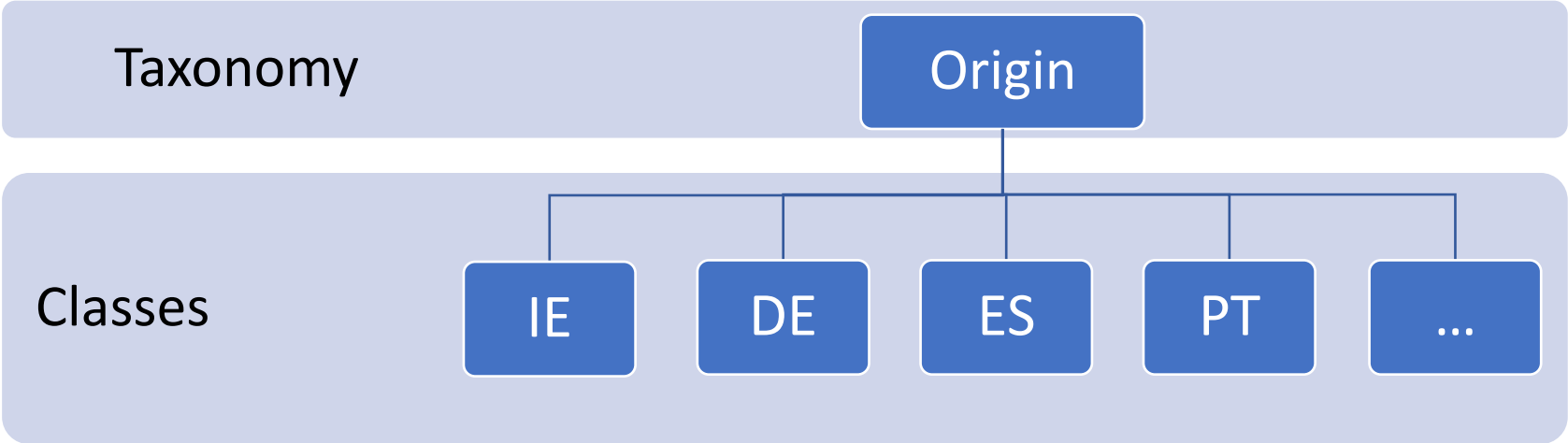
Taxonomy — Country

Classes — IE | DE | ES | MX | …

Country classifies names to ~250 countries with valid ISO2 codes, from Ireland (IE) to Spain (ES) or Mexico (MX) including all African and Asian countries.

id	ref12315	The input identifier
name	Jing Cao	The input full name
country	CN	The likely residence country ISO2 code, which CAN include melting-pot countries
countryAlt	TW	The best alternative residence country
region	Asia	An arbitrary grouping of countries by topRegion/Region/subRegion
topRegion	Asia	An arbitrary grouping of countries by topRegion/Region/subRegion
subRegion	Eastern Asia	An arbitrary grouping of countries by topRegion/Region/subRegion
countriesTop	CN, TW, HK…	The top 10 likely residence country ISO2 codes
probabilityCalibrated	.89	The calibrated probability of having guessed right the country of residence (CN)
probabilityCalibratedAlt	0.92	The calibrated probability of having guessed right the country of residence as either CN or TW.

Taxonomy → Origin

Classes → IE | DE | ES | PT | ...

Origin infers the likely country of origin from a name, based on naming patterns among ~130 countries with strong name identity (IE, DE, ES, PT etc.)

id	ref12315	The input identifier
name	Jing Cao	The input full name
country	CN	The likely residence country ISO2 code, which CAN include melting-pot countries
countryAlt	TW	The best alternative residence country
region	Asia	An arbitrary grouping of countries by topRegion/Region/subRegion
topRegion	Asia	An arbitrary grouping of countries by topRegion/Region/subRegion
subRegion	Eastern Asia	An arbitrary grouping of countries by topRegion/Region/subRegion
countriesTop	CN, TW, HK...	The top 10 likely residence country ISO2 codes
probabilityCalibrated	.89	The calibrated probability of having guessed right the country of residence (CN)
probabilityCalibratedAlt	0.92	The calibrated probability of having guessed right the country of residence as either CN or TW.

Taxonomy

Diaspora

Classes

Irish | German | Hispanic | Chinese | ...

Diaspora infers the likely ethnicity, diaspora or country of origin from a name, given a geographic context (ex. US, CA, ...) with ~130 ethnicities (Irish, Chinese, etc)

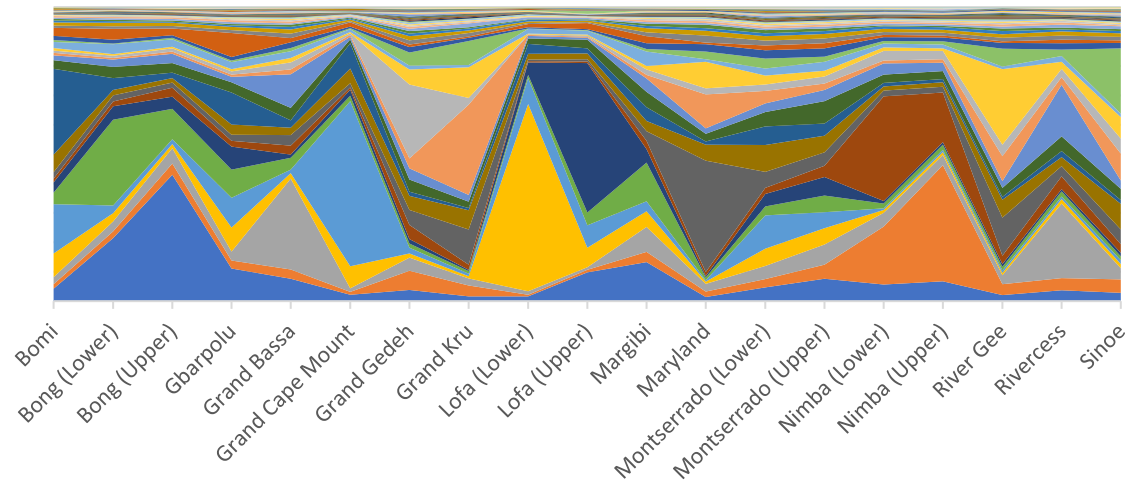| Field	Example	Description
id	ref12315	The input identifier
firstName	Mary	The input first name / given name
lastName	Cao	The input last name / surname
countryIso2	US	The country of residence, the host country (ex. US, CA, NZ, GB)
ethnicity	Chinese	The likely ethnicity
ethnicityAlt	Vietnamese	The best alternative ethnicity
ethnicitiesTop	Chinese, Vietnamese, Korean ...	The top 10 likely ethnicities
score	25	A non calibrated Score : score = Math.log(getProbaFirst() / getProbaNotFirst()) maxed to 100 ; NB: diaspora doesn't have calibrated probabilities YET

Limitations to such taxonomies

- Human societies are fractal in their diversity :
  - A coarse-grained classification model may not fit all markets (ex. 'African-American/Black vs. White vs. African / Black : how does North-African fit?)
  - A fine-grained classification model may be too fine-grained or controversial in specific regions
  - For example, IN/Indian is one class among 130 classes in our Origin/Diaspora taxonomy, but there are ~30 states in India with many ethnic/clan/caste system sub-groups

Example of complex regional or ethnic identities in Africa : Liberia.

This visualization utilizes unsupervised name classification algorithm, to recognize subgroups in different regions of Liberia.

Liberia - a regional onomastics 'mille-feuille'

Bomi, Bong (Lower), Bong (Upper), Gbarpolu, Grand Bassa, Grand Cape Mount, Grand Gedeh, Grand Kru, Lofa (Lower), Lofa (Upper), Margibi, Maryland, Montserrado (Lower), Montserrado (Upper), Nimba (Lower), Nimba (Upper), River Gee, Rivercess, Sinoe

- Privacy and self-identification : how can people 'override' the classification ?

Thank you !

Elian CARSENAT,

elian.carsenat@namsor.com

Phone : +33 6 52 77 99 07