# Picking on the Same Person:
Does Algorithmic Monoculture lead to Outcome Homogenization?

Rishi Bommasani, Kathleen A. Creel, Ananya Kumar
Dan Jurafsky, Percy Liang

Framing

Modern AI centers on sharing

ImageNet, PyTorch, BERT, Adam, …

What are the harms?

especially individual-centric harms

Systemic Failure:

all models fail for same person

Homogenization:

rate of systemic failure is high

Sharing training data
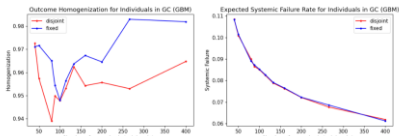
Fixed (more sharing) means the different companies use exactly the same training data.
Disjoint (less sharing) means the different companies use disjoint samples from the same distribution.
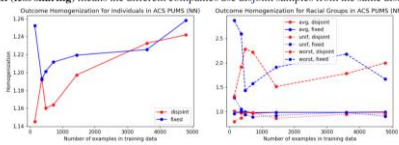
Left is homogenization at individual-level. **right** is for the 9 racial groups tracked in US census.
Compare left to right (**avg, unif**): homogenization is higher for individuals than racial groups.

Sharing foundation models
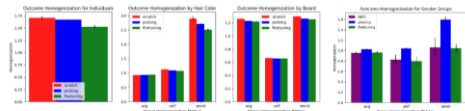
Effect of no foundation model (**scratch**) vs. different ways to use that are more/less expensive.
Vision: scratch is the most homogeneous, then **probing**, then **finetuning**.
Language: probing is the most homogeneous; **finetuning** and **BitFit** are similarly homogeneous.

Takeaway

Sharing data homogenizes
Sharing models is more complex

Need to think about distribution shift

Narrative
Decision-makers deploy systems that **share** components.

Sharing can be reinterpreted as *monoculture*!

Due to monoculture, will some individuals **exclusively** receive negative outcomes?

Experimental Findings
For training data, sharing homogenizes outcomes.
For foundation models, its more complicated.

TL;DR
Monoculture is ubiquitous & growing:
we need to actively study homogenization!

Metric

Companies **i** in {1, …, **k**}
Individuals **j** in {1, …, **N**}
I is indicator RV of failure

Per-company failure rate:

$$\mathtt{fail}(h^i) \triangleq \mathop{\mathbb{E}}_{x^i \sim D^i} I^i(x^i) = \mathop{\mathrm{Pr}}_{x^i \sim D^i}\left[I^i(x^i)=1\right]$$

Obs. systemic failure rate:

$$S \triangleq \mathop{\mathbb{E}}_{j}\left[\prod_i I^i(x_j^i)\right] = \mathop{\mathrm{Pr}}_{j}\left[I^1(x_j^1)=1 \wedge \cdots \wedge I^k(x_j^k)=1\right]$$

Homogenization metric:

$$H^{\mathbf{individual}}(h^1,\ldots,h^k) \triangleq \frac{S}{\prod_{i=1}^{k}\mathtt{fail}(h^i)} = \frac{\mathbb{E}_j\left[\prod_i I^i(x_j^i)\right]}{\prod_i \mathbb{E}_j\left[I^i(x_j^i)\right]}$$

Challenges for society

Diagnosis:
Opacity (unaware of monoculture)
Measurement:
Privacy and linking individuals
Rectification:
Incentives-compatibility

Why this matters?

Harms for *individuals*
Look beyond a single model
Systemic exclusion/hierarchy
Relational equality