

Bias & Fairness Testing

Bias Detection Areas



Demographic bias (age, sex, race)



Socioeconomic and geographic bias



Disability and accessibility bias



Language and cultural bias

Fairness Metrics

- **Demographic parity:** Equal outcomes across groups
- **Equalized odds:** Equal TPR and FPR across groups
- **Calibration:** Predictions equally accurate for all groups
- **Individual fairness:** Similar individuals treated similarly

Test across diverse populations, especially vulnerable and underserved groups