3. What kinds of techniques are being developed to mitigate bias and systemic inequities in artificial intelligence?

Definitions of Fairness in ML (Verma & Rubin 2018)

	Definition	Paper	Citation #	Result
3.1.1	Group fairness or statistical parity	[12]	208	×
3.1.2	Conditional statistical parity	[11]	29	✓
3.2.1	Predictive parity	[10]	57	✓
3.2.2	False positive error rate balance	[10]	57	×
3.2.3	False negative error rate balance	[10]	57	✓
3.2.4	Equalised odds	[14]	106	×
3.2.5	Conditional use accuracy equality	[8]	18	×
3.2.6	Overall accuracy equality	[8]	18	✓
3.2.7	Treatment equality	[8]	18	×
3.3.1	Test-fairness or calibration	[10]	57	✓
3.3.2	Well calibration	[16]	81	✓
3.3.3	Balance for positive class	[16]	81	✓
3.3.4	Balance for negative class	[16]	81	×
4.1	Causal discrimination	[13]	1	×
4.2	Fairness through unawareness	[17]	14	✓
4.3	Fairness through awareness	[12]	208	×
5.1	Counterfactual fairness	[17]	14	–
5.2	No unresolved discrimination	[15]	14	–
5.3	No proxy discrimination	[15]	14	–
5.4	Fair inference	[19]	6	–

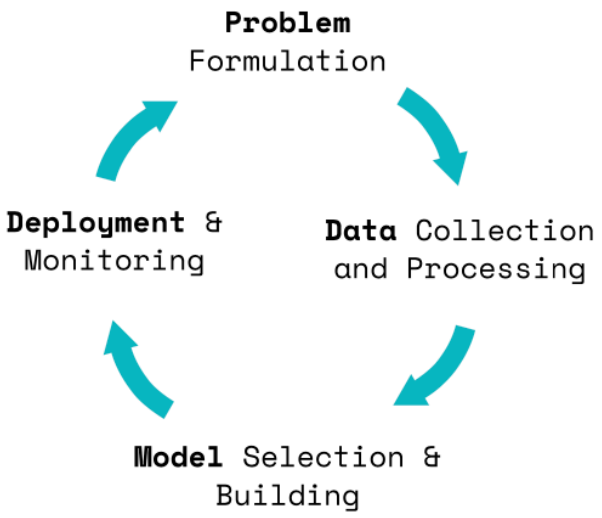
Parity Metrics

Race-Neutral

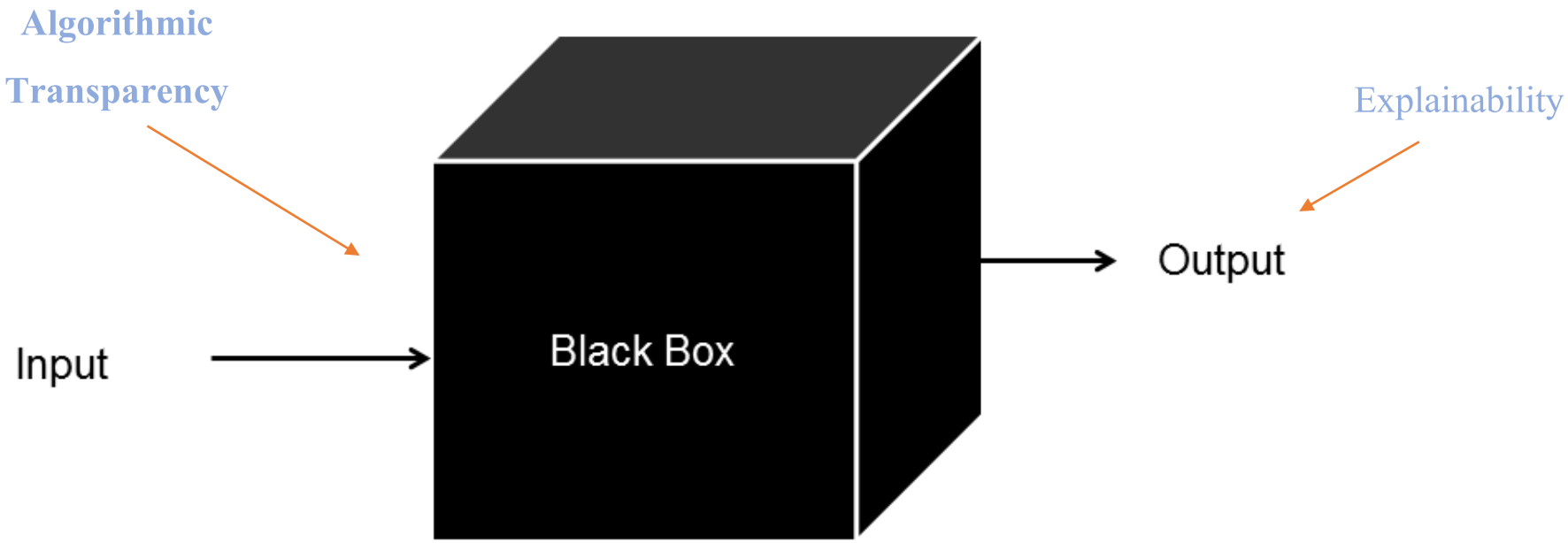
Trade-offs

Algorithmic Affirmative Action

AI/ML Developers (Holstein et al. 2019)



- Data curation and representativeness in the dataset
- Bias awareness tools that allow them to detect bias
- Fairness checklists
- Toolkit that helps them anticipate what kinds of fairness issues can arise in their specific application domain; suggesting that issues of bias are largely contained within specific domains, when in fact they are relational and reciprocal
- Fairness is defined largely in terms of statistical parity (demographic, predictive); thus largely treated as an individual rather than structural problem



Internal behavior of the code is unknown

Solutions

- Algorithmic Auditing and Monitoring
- Human-in-the-Loop
- More Diverse and Domain Expertise
- Greater AI Literacy

Quiz

craig.watkins@austin.utexas.edu

The University of Texas at Austin
Institute for Media Innovation
Moody College of Communication

IC2
Innovating Well-Being

GOOD SYSTEMS
A UT Grand Challenge