Text, logo

Description automatically generatedLogo

Description automatically generated with medium confidence

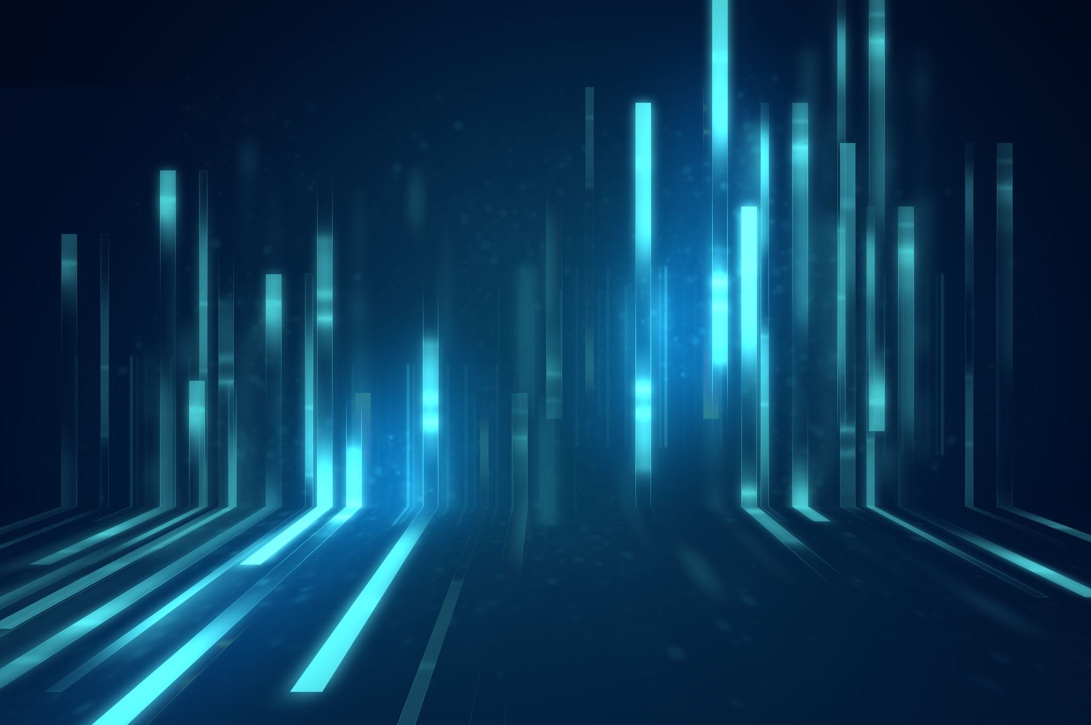
­­­

Fairness

Intermediate

Module 2

Bias Within Algorithms



# Introduction

In this module we will explain the relationship between bias and fairness; identify sources of bias and ways to address them; look at the social implications of bias within different domains; and finally apply learnings of fairness to different use-cases from one of Trilateral Research’s own projects.

­­­

# Reflection Sheet

1. If you had to make a list of intrinsic values, which ones would you choose, and why?
2. Do you think fairness is an intrinsic or instrumental value? Why?
3. You read 2 articles about using AI in policing. Is it possible to create fair and non-discriminatory AI for predictive policing? Why, or why not?

**Task**

Write a three-paragraph pitch to persuade a client to consider procuring an AI tool because your development team has mitigated bias. Show that you understand what fairness is, how it might conflict with other values, but how it can still be achieved.

Re-read the Amazon case study and answer:

1. How could this have been avoided?
2. Using the principles learned in this module, make an argument that Amazon is culpable for neglecting bias

­­­



**Self-assessment pass/fail questions**

1. Intrinsic value is...
   1. the value that a thing has “in itself,” or “for its own sake,” or “as such,” or “in its own right.”
   2. The value that a thing has for the purpose of achieving something else of value
   3. Hard to define
2. Can fairness conflict with other values?
   1. No
   2. Yes
   3. Yes, but fairness is always the most important ethical value
3. What did some philosophers think is intrinsically valuable?
   1. Pleasure/Happiness
   2. Sleeping a lot
   3. Nothing
4. Why is understanding fairness important to AI tool development?
   1. Because we are told so
   2. Because risks of unfairness occur in most societal contexts
   3. Because AI is inherently unfair

# Pre-reading

Intrinsic vs. Instrumental Value

<https://www.thoughtco.com/intrinsic-and-instrumental-value-2670651>

Why it’s so damn hard to make AI fair and unbiased

<https://www.vox.com/future-perfect/22916602/ai-bias-fairness-tradeoffs-artificial-intelligence>

AI Fairness Isn’t Just an Ethical Issue

<https://hbr.org/2020/10/ai-fairness-isnt-just-an-ethical-issue>

AI Bias - What Is It and How to Avoid It?

<https://levity.ai/blog/ai-bias-how-to-avoid>

Dutch scandal serves as a warning for Europe over risks of using algorithms

<https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>

Qs 1) a, 2) b, 3) a, 4) b Q

**Answers**