AI Problem Statement

Written By: Micah Austin, Morgan Wright, Cameron Seymore

Problem Statement

Artificial Intelligence trained on non-representative population data can have systematic biases that lead to unfair and inequitable outcomes across underrepresented groups.

Target Audience

Individuals who use or develop general-purpose AI systems, especially where needs and perspectives vary across users and must be taken into account. This includes differences in language use and interpretation, appreciation of artistic techniques and details, and diverse approaches to learning — understanding versus memorization when engaging with new information.

Issue

Many AI models are trained on datasets that do not accurately represent the full diversity of the general population. These datasets often overrepresent certain demographic groups while underrepresenting others, leading to unaccounted-for biases in model predictions and decision-making. As a result, AI systems may produce unfair or inaccurate outcomes for marginalized or underrepresented populations.

Where Is the Problem

The use of AI is becoming increasingly popular in many vital aspects of society such as jobs, healthcare, law enforcement, and education.

Significance

Without efforts to detect, measure, and mitigate training data bias, AI systems will continue to scale said inequitable outcomes across large populations.