



AI Bias Mimics Human Bias

Rowland Chen, Creator, Framework Foundry

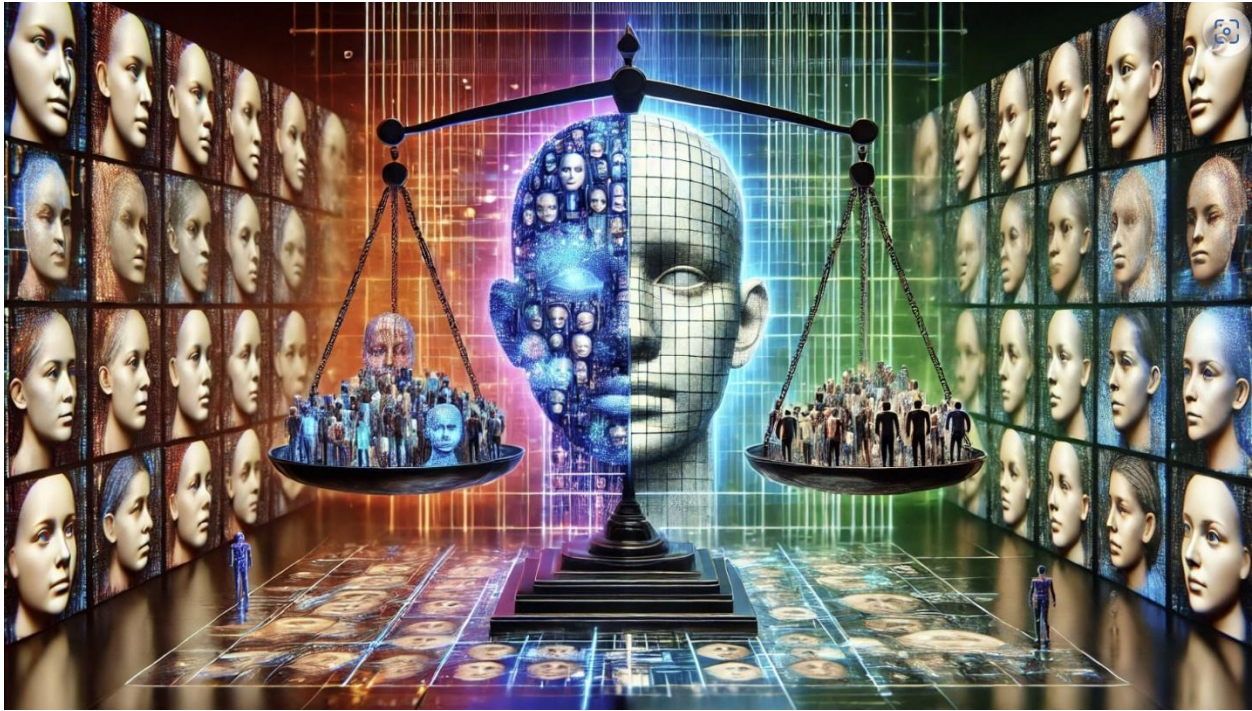


Figure 1. Rendering of the prompt “Create an image of preventing bias in AI” by ChatGPT. Note the biases in the image.

Regulating bias to eliminate it from artificial intelligence (AI) will not work. AI is a human invention, and its development will always have a human factor, at least in the foreseeable future.

Just as the 18th Amendment to the U.S. Constitution (alcohol prohibition) attempted to regulate alcohol consumption among its citizens, unintended consequences emerged. The most visible of those consequences was the creation of a criminal industry, illegal distribution of alcohol by outlaws, birthing organized crime.

The [European Union Artificial Intelligence Act of 2024](#) (the Act) is an attempt to regulate human behavior through broad-reaching regulation. The Act, among other goals, provides a framework for member states to follow with the intent of preventing bias from occurring in the development of AI systems, deployment, and usage in the EU member states and AI systems from “third countries”.

The Act is a noble attempt to modify the behaviors of AI developers, developers being organizations, groups, and individuals. Where the Act falls short is in providing guidelines on how to modify those behaviors, other than the implied emotional return on investment of supporting



the common good of the EU. What is the motivation for compliance other than the risk of fines and imprisonment?

Artificial intelligence is treated in the Act as a single technology while ignoring the multiple technologies comprising AI. Unpacking AI brings forward a level of detail that enables action, or inaction, as the case might be. Technologies comprising AI are decomposed into five categories, which in turn further decompose into a total of twenty varieties of AI.

Additionally, the design, development, and deployment of AI is represented as a series of stages, or a lifecycle. The process follows this sequence: basic theory, architecture, data, algorithm, and usage.

Understanding that AI can be unpacked by variety and by stage of lifecycle is an absolute requirement for beginning a dialogue concerning bias. The Act treats AI as a whole, without regard to branch or variety, and focuses only on data, algorithm development, and usage throughout the lifecycle. A gap exists in the first two stages: basic theory and architecture.

Basic theory in AI and architecture design are the two most important aspects of AI in preventing bias. The conceptualization and design of AI are the lifecycle stages with the greatest leverage for control, management, and prevention of bias.

Efforts to block or remove bias by focusing on clean data, coding, and AI usage are laudable. But these stages come too late to effectively prevent bias. Once bias enters an AI system, the level of effort required to remove it is akin to the Greek myth about Tantalus and the grapes that are continually out of his reach. Forever.

Influencing the exclusion of bias by legislative acts and regulations boils down to influencing and modifying human behaviors. People are biased even with their best intentions. Gender, race, culture, age, and a myriad of other biases will never go away completely.

As seen in Figure 1, an image of “preventing bias in AI” by ChatGPT demonstrates biased results.