



UP  
TO **20%** reduction  
in bias in Foundational  
Multimodal Models.<sup>1</sup>

“By probing six models using data-intensive methods, the team mitigated biases by as much as 20 percent.”<sup>1</sup>

**Vasudev Lal, Principal Research Scientist of Cognitive AI, Intel Labs**

# Intel Labs Mitigates AI Bias in Foundational Multimodal Models by 20 Percent<sup>1</sup>

Intel Labs researchers developed an innovative approach to reduce bias in AI foundational models using social counterfactuals. They created a dataset of synthetic images that varied intersectional social attributes which allowed them to isolate and study the effect of each attribute. Using this method, they could probe six foundational models and reduce biases by up to 20%.<sup>1</sup> The team utilized large AI clusters equipped with 3rd Gen Intel® Xeon® Scalable processors and Intel® Gaudi® 2 AI accelerators to train foundational multimodal models and leverage results across data, text, images, and video. This work is part of Intel's commitment to Responsible AI, aiming to ensure AI models are accurate, grounded in authoritative sources, and free from harmful biases. Additionally, the researchers have open-sourced their dataset to help improve AI fairness across the industry.

## Products and Solutions

[3rd Gen Intel® Xeon® Scalable Processors](#)  
[Intel® Gaudi® 2 AI accelerators](#)

## Industry

Semiconductor  
Manufacturing

## Organization Size

10,001+

## Country

United States

## Learn more

[Case Study](#)  
[Video](#)