

Data is never neutral

Another important thing about probabilistic systems is that they are not representing all available information equally. They certainly cannot be considered a perfect representation of the world's population.

These systems can suffer from gender bias, racial bias, age bias, political bias, cultural bias and many other types of bias.

The problem starts with the data itself. If the data contains certain patterns, the model can learn those patterns. If some groups are overrepresented and others are underrepresented, the resulting system can reflect those differences.

Datasets are never completely neutral. They are created by people, from sources selected by people, and the decisions made while collecting and preparing that data can influence the final model.

A lot of the information available on the internet is also dominated by Western sources and by languages and cultures that have a much larger online presence. This means that asking an AI system about the "world" does not necessarily mean that we are getting an equal representation of the whole world.

The model only knows what is available to it through its training and other information sources.

This makes me think about a simple chain:

World → Data → Dataset → Model → Output

Bias can enter at almost every stage of this process.

Revision #1

Created 2026-09-13 10:18:13 UTC by Cherry

Updated 2026-09-13 10:20:18 UTC by Cherry