

A 'inteligência artificial' e o racismo

Diretrizes éticas criadas para regular conteúdos de modelos de linguagem têm se revelado ineficazes

Ana Cristina Rosa

Valor, 5.mai.2024 às 12h22

- Usar o que se convencionou chamar de ["inteligência artificial"](#) (pois não é inteligente) para realizar tarefas diárias é cada vez mais comum. Existem ferramentas que, em geral, facilitam a vida. Contudo, se utilizadas para [fomentar estereótipos](#) que impactam negativamente milhões de pessoas, um problema grave está instalado.

Há muito se supunha que a "IA" poderia servir para reproduzir o racismo. Tratei do tema na coluna ["Códigos de preconceito"](#), em 2021. O que era hipótese ganhou ares de certeza com a publicação, pela Universidade Cornell, de um estudo de pesquisadores do Instituto Allen de Inteligência Artificial. O documento revela que "à medida que as ferramentas de 'IA' ficam 'mais inteligentes', se tornam mais racistas".

Modelos de linguagem como [ChatGPT \(da OpenAI\)](#) e [Gemini \(do Google\)](#) mantêm estereótipos racistas sobre falantes do inglês vernacular afro-americano (criado por negros) e tendem a descrevê-los como "estúpidos e preguiçosos". Além disso, a "IA" se mostrou mais propensa a recomendar a pena de morte a réus que usam esse dialeto e a indicar salários e empregos piores aos falantes da linguagem.

Práticas racistas são multifacetadas. Vão de agressões disfarçadas de elogios (negra bonita!) até insultos verbais (macaco!) e violência física. Frente à sofisticação da camuflagem do racismo, diretrizes éticas criadas para regular conteúdos de modelos de linguagem têm se revelado ineficazes.

"Inteligência artificial" não é mágica. São humanos que decidem o que as máquinas fazem. Assim, elas reproduzem crenças sociais e têm servido de instrumento de discriminação, violando princípios, como a igualdade.

É equivocado atribuir a uma ferramenta a responsabilidade por problemas criados por pessoas. Mas também é evidente que, ao ser desenvolvida com dados analisados por uma perspectiva racista, a "IA" multiplica e até aprofunda desigualdades e injustiças que ameaçam grupos sociais predeterminados.