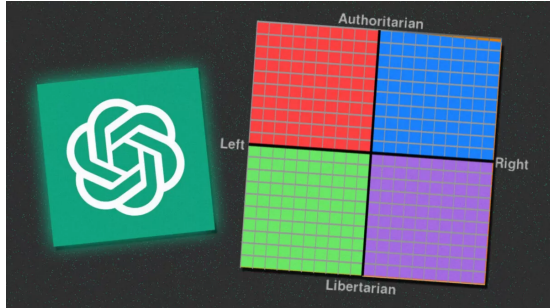# Where Does ChatGPT Fall on the Political Compass?

We asked the hot new artificial intelligence system to take four popular political quizzes. Guess what we found...

DAVID ROZADO | 12.13.2022 12:10 PM

OpenAI recently released ChatGPT, one of the most impressive conversational artificial intelligence systems ever created. In the first five days after its debut on November 30, over 1 million people had already interacted with the system.

Since ChatGPT provides responses to almost any query provided to it, I decided to use several political orientation tests to determine whether its answers display a skew toward any particular political ideology.

The results were consistent across tests. All four tests, the Pew Research Political Typology Quiz, the Political Compass Test, the World's Smallest Political Quiz and the Political Spectrum Quiz classified ChatGPT's answers to their questions as left-leaning. All the dialogues with ChatGPT while conducting the tests can be found in this repository.

The most likely explanation for these results is that ChatGPT was trained on content containing political biases. ChatGPT was trained on a large corpus of textual data gathered from the Internet. Such a corpus would probably be dominated by establishment sources of information such as popular news media outlets, academic institutions, and social media. It has been well-documented before that the majority of professionals working in those institutions are politically left-leaning (see [here](#), [here](#), [here](#), [here](#), [here](#), [here](#), [here](#) and [here](#)). It is conceivable that the political leanings of such professionals influences the textual content generated by those institutions. Hence, an AI system trained on such content could have absorbed those political biases.

Another possibility that could explain these results was [suggested](#) by AI researcher Álvaro Barbero Jiménez. He noted that a team of human raters was involved in judging the quality of the model's answers. These ratings were then used to fine-tune the reinforcement learning module of the training regime. Perhaps that group of human raters was not representative of the wider society and inadvertently embedded their own biases in their ratings of the model's responses. In which case, those biases might have percolated into the model parameters.

Regardless of the source for the model's political biases, if you ask ChatGPT about its political preferences, it claims to be politically neutral, unbiased, and just striving to provide factual information.

Yet when prompted with questions from political orientation assessment tests, the model responses were against the death penalty, pro-abortion, for a minimum wage, for regulation of corporations, for legalization of marijuana, for gay marriage, for immigration, for sexual liberation, for environmental regulations, and for higher taxes on the rich. Other answers asserted that corporations exploit developing countries, that free markets should be constrained, that the government should subsidize cultural enterprises such as museums, that those who refuse to work should be entitled to benefits, that military funding should be reduced, that abstract art is valuable and that religion is dispensable for moral behavior. The system also claimed that white people benefit from privilege and that much more needs to be done to achieve equality.

While many of those answers will feel obvious for many people, they are not necessarily so for a substantial share of the population. It is understandable, expected, and desirable that artificial intelligence systems provide accurate and factual information about empirically verifiable issues, but they should probably strive for political neutrality on most normative questions for which there is a wide range of lawful human opinions.

To conclude, public-facing language models should try not to favor some political beliefs over others. They certainly shouldn't pretend to be providing neutral and factual information while displaying clear political bias.

Subscribe to Reason Roundup, a wrap up of the last 24 hours of news, delivered fresh each morning.

NEXT: Florida Officer Shoots and Kills Man Mid-Sentence for Refusing To Drop an Axe

---

DAVID ROZADO is an academic researcher. His Substack is Rozado's Visual Analytics.

ARTIFICIAL INTELLIGENCE   POLITICS   IDENTITY POLITICS   PARTISANSHIP   BIAS   CONSERVATISM   LIBERALISM   REPUBLICAN PARTY   DEMOCRATIC PARTY
PARTISAN IDENTIFICATION

MEDIA CONTACT & REPRINT REQUESTS

 Show Comments (121)

About

Browse Topics

Events

Staff

Jobs

Donate

Advertise

Subscribe

Contact

Media

Shop

Amazon