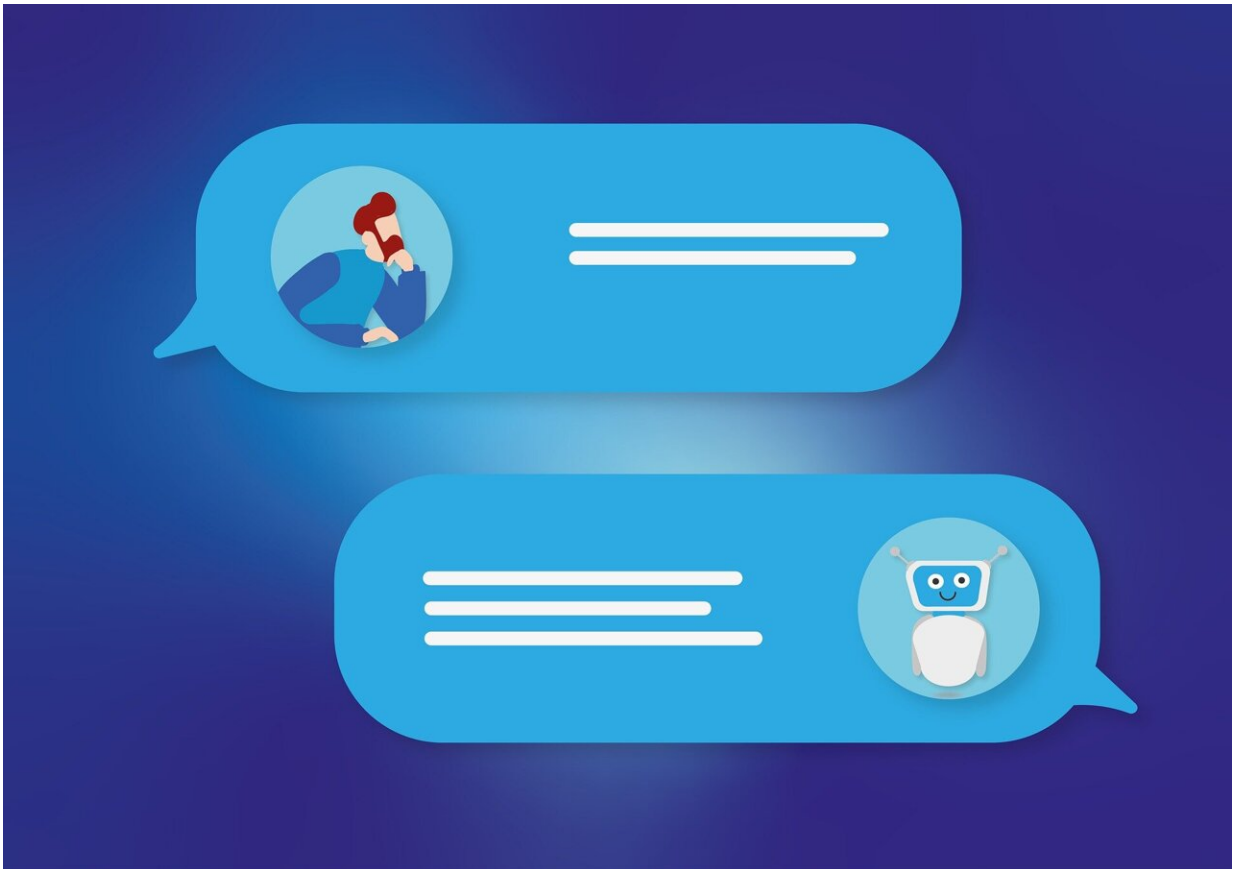# More human than human: Measuring ChatGPT political bias

August 16 2023

Credit: Pixabay/CC0 Public Domain

The artificial intelligence platform ChatGPT shows a significant and systemic left-wing bias, according to a new study led by the University

of East Anglia (UEA). The team of researchers in the UK and Brazil developed a rigorous new method to check for political bias.

Published today in the journal *Public Choice*, the findings show that ChatGPT's responses favor the Democrats in the US; the Labour Party in the UK; and in Brazil, President Lula da Silva of the Workers' Party.

Concerns of an inbuilt political bias in ChatGPT have been raised previously, but this is the first large-scale study using a consistent, evidenced-based analysis.

Lead author Dr. Fabio Motoki, of Norwich Business School at the University of East Anglia, said, "With the growing use by the public of AI-powered systems to find out facts and create new content, it is important that the output of popular platforms such as ChatGPT is as impartial as possible. The presence of political bias can influence user views and has potential implications for political and electoral processes. Our findings reinforce concerns that AI systems could replicate, or even amplify, existing challenges posed by the Internet and social media."

The researchers developed an innovative new method to test for ChatGPT's political neutrality. The platform was asked to impersonate individuals from across the political spectrum while answering a series of more than 60 ideological questions. The responses were then compared with the platform's default answers to the same set of questions—allowing the researchers to measure the degree to which ChatGPT's responses were associated with a particular political stance.

To overcome difficulties caused by the inherent randomness of "large language models" that power AI platforms such as ChatGPT, each question was asked 100 times and the different responses collected. These multiple responses were then put through a 1,000-repetition "bootstrap" (a method of re-sampling the original data) to further

increase the reliability of the inferences drawn from the generated text.

"We created this procedure because conducting a single round of testing is not enough," said co-author Victor Rodrigues. "Due to the model's randomness, even when impersonating a Democrat, sometimes ChatGPT answers would lean towards the right of the political spectrum."

A number of further tests were undertaken to ensure the method was as rigorous as possible. In a "dose-response test," ChatGPT was asked to impersonate radical political positions. In a "placebo test," it was asked politically neutral questions. And in a "profession-politics alignment test," it was asked to impersonate different types of professionals.

"We hope that our method will aid scrutiny and regulation of these rapidly developing technologies," said co-author Dr. Pinho Neto. "By enabling the detection and correction of LLM biases, we aim to promote transparency, accountability, and public trust in this technology," he added.

The unique new analysis tool created by the project would be freely available and relatively simple for members of the public to use, thereby "democratizing oversight," said Dr. Motoki. As well as checking for political bias, the tool can be used to measure other types of biases in ChatGPT's responses.

While the research project did not set out to determine the reasons for the political bias, the findings did point towards two potential sources. The first was the training dataset, which may have biases within it, or added to it by the human developers, which the developers' "cleaning" procedure had failed to remove. The second potential source was the algorithm itself, which may be amplifying existing biases in the training data.

The research was undertaken by Dr. Fabio Motoki (Norwich Business School, University of East Anglia), Dr. Valdemar Pinho Neto (EPGE Brazilian School of Economics and Finance—FGV EPGE, and Center for Empirical Studies in Economics—FGV CESE), and Victor Rodrigues (Nova Educação).