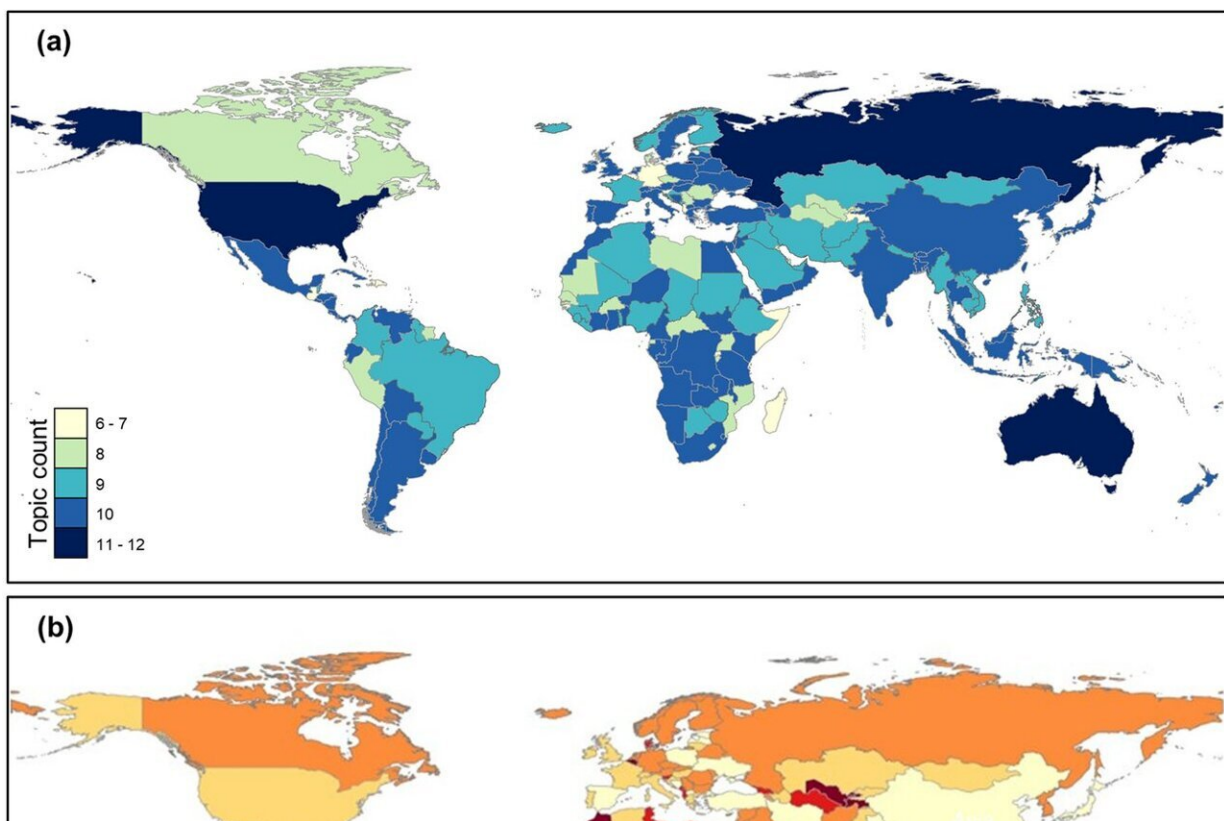# Should chatbots chime in on climate change? Study explore potential of AI platforms for climate literacy

April 30 2024, by Travis Williams

Topic count and accuracy results for GPT-4. a) Spatial variation of topic counts across the first GPT-4 iteration, where topic count increases from light green to dark blue. b) Map of the standard deviation of topic counts for GPT-4 across all ten iterations where standard deviation increases from light yellow to dark red. c) Accuracies for droughts (orange triangles and line), cyclones (yellow circles and line), and floods (green squares and line) for GPT-4. Maps and graphs were

Can chatbots provide accurate information about the dangers of climate change? Well, that depends on a variety of factors including the specific topic, location being considered, and how much the chatbot is paid, according to a group of Virginia Tech researchers.

"I think what we found is that it's OK to use artificial intelligence, you just have to be careful and you can't take it word-for-word," said Gina Girgente, who graduated with a bachelor's degree in geography last spring. "It's definitely not a foolproof method."

Girgente was part of an interdisciplinary research team that posed questions about three climate change-related hazards—tropical storms, floods, and droughts—in 191 countries to both free and paid versions of ChatGPT. Developed by OpenAI Inc., ChatGPT is a large-language model designed to understand questions and generate text responses based on requests from users.

The group then compared the chatbots' answers against hazard risk indices they generated using data from the Intergovernmental Panel on Climate Change, a United Nations body tasked with assessing science related to climate change.

"Overall, we found more agreement than not," said Carmen Atkins, lead author and second-year Ph.D. student in the Department of Geosciences. "The AI-generated outputs were accurate more than half the time, but there was more accuracy with tropical storms and less with droughts."

In an article published in *Communications Earth & Environment*, the

group reported ChatGPT4's responses aligned to the indices with the following degrees of accuracy:

- Tropical storms—80.6%
- Flooding—76.4%
- Droughts—69.1%

They also found the chatbot's accuracy to increase when asked about more developed countries and that the paid version of the platform, currently ChatGPT-4, was more accurate than the free version, which at the time was ChatGPT-3.5.

They found less consistency in answers about the same topic when asked about certain regions, especially many countries in Africa and the Middle East, that are considered low income or developing countries. And they found the paid version of the platform, currently ChatGPT-4, was more accurate than the free version, which at the time was ChatGPT-3.5.

"We're the first to look at this in a systematic way, as far as we're aware, and beyond that, it's really a call to action for more people to look into this issue," Atkins said.

According to the publication, the researchers selected the ChatGPT service because of its popularity, particularly among users aged 18–34, and its rate of use in developing nations.

Atkins and Girgente, who will be graduate student at the University of Denver in the fall, said the idea for the project stemmed from their time in the class of Junghwan Kim, an assistant professor in the Department of Geography. There, Kim shared some of his own research studying the geographic biases in ChatGPT when asked about environmental justice topics.

"It opened up this whole idea of the different things we could look at with ChatGPT," Atkins said. "And for me, personally, climate change and specifically climate change literacy, which is critical to combating misinformation, made me want to see if ChatGPT could be helpful or harmful with that issue."

Kim is also a co-author of this new study, as is Manoochehr Shirzaei, an associate professor of geophysics and remote sensing.

Kim said he felt the results would be especially important to share with students who might put too much faith in the chatbots and that they had also impacted his own use of the software.

While the group believes their findings demonstrate the clear need for a cautious approach to using generative artificial intelligence, they also believe this study is only the beginning of the needed research, which will need to be repeated as the different software continues to evolve.

"This is just the tip of the iceberg, and more than anything, we just want to draw attention to these issues," Atkins said. "Generative AI is part of our world now, but we need to use it in as well educated a way as we can."

Provided by Virginia Tech