

PREACHINGtoday

BE INSPIRED

The following illustration is located at: <https://www.preachingtoday.com/illustrations/2023/july/experts-concerned-about-chatbots-making-stuff-up.html>

Experts Concerned About Chatbots Making Stuff Up

Gerrit De Vynck wrote a story in *The Washington Post* about how artificial intelligences respond to the errors they make.

Citing a recent MIT research paper, De Vynck reported that a group of scientists loaded up two iterations of Open AI's ChatGPT, and asked each one a simple question about the geographical origin of one of MIT's professors. One gave an incorrect answer, the other a correct one.

Researchers then asked the two bots to debate until they could agree on an answer. Eventually, the incorrect bot apologized and agreed with the correct one. The researchers' leading theory is that allowing chatbots to debate one another will create more factually correct outcomes in their interactions with people.

One of the researchers said, "Language models are trained to predict the next word. They are not trained to tell people they don't know what they're doing." De Vynck adds, "The result is bots that act like precocious people-pleasers. [They're] making up answers instead of admitting they simply don't know."

AIs like ChatGPT are not trained to discern truth from falsehood, which means that false information gets included along with truth. Chatbots like OpenAI's ChatGPT, Microsoft's Bing, and Google's Bard have demonstrated a major fatal flaw: They make stuff up all the time. These falsehoods, or digital hallucinations as they are being called, are a serious concern because they limit the effectiveness of the AI as a tool for fact-finding.

What's worse, scientists are beginning to see evidence that AIs pick up on societal fears around robots gaining sentience and turning against humanity, and mimic the behavior they see depicted in science fiction. According to this theory, if an artificial intelligence actually kills a human being, it might be because it learned from HAL, the murderous robot from *2001: A Space Odyssey*.

Sundar Pichai, chief executive officer at Google said, "No one in the field has yet solved the hallucination problem. All models do have this as an issue." When asked if or when this will change, Pichai was less than optimistic. "[It's] a matter of intense debate," he said.

Possible Preaching Angle:

In our pursuit of technology, we must never give up our human responsibility to seeking or telling the truth.

Source:

Gerrit De Vynck, "ChatGPT 'hallucinates.' Some researchers worry it isn't fixable.," *Washington Post* (5-30-23)

Filters: [News](#); [Social Media & Technology](#); [Stories](#)

Topics: [Answers](#); [Computers](#); [Concern](#); [Confession](#); [Correction](#); [Dependence](#); [Discernment](#); [Error](#); [Falsehood](#); [Intelligence](#); [Internet](#); [Knowledge](#); [Mistakes](#); [News That Illustrates](#); [Pleasing Others](#); [Responsibility](#); [Science](#); [Social Concern](#); [Technology](#); [Trustworthiness](#); [Truth](#); [Truthfulness](#)

Scripture: [Proverbs 28:13](#); [Colossians 2:8](#); [Ephesians 4:25](#); [Genesis 3:1-7](#); [Proverbs 11:1](#); [Luke 16:15](#); [Exodus 20:3](#); [Daniel 12:4](#); [Romans 1:18-25](#); [Colossians 3:9](#); [Proverbs 1:7](#); [2 Timothy 3:7](#); [Leviticus 19:35-37](#); [Isaiah 6:9](#); [Acts 28:26](#); [Ecclesiastes 8:17](#); [1 Corinthians 3:18-19](#); [1 Corinthians 1:19-21](#)

Copyright © 2023 by the author or Christianity Today/PreachingToday.com.

[Click here](#) for reprint information on PreachingToday.com.

CHRISTIANITY  TODAY