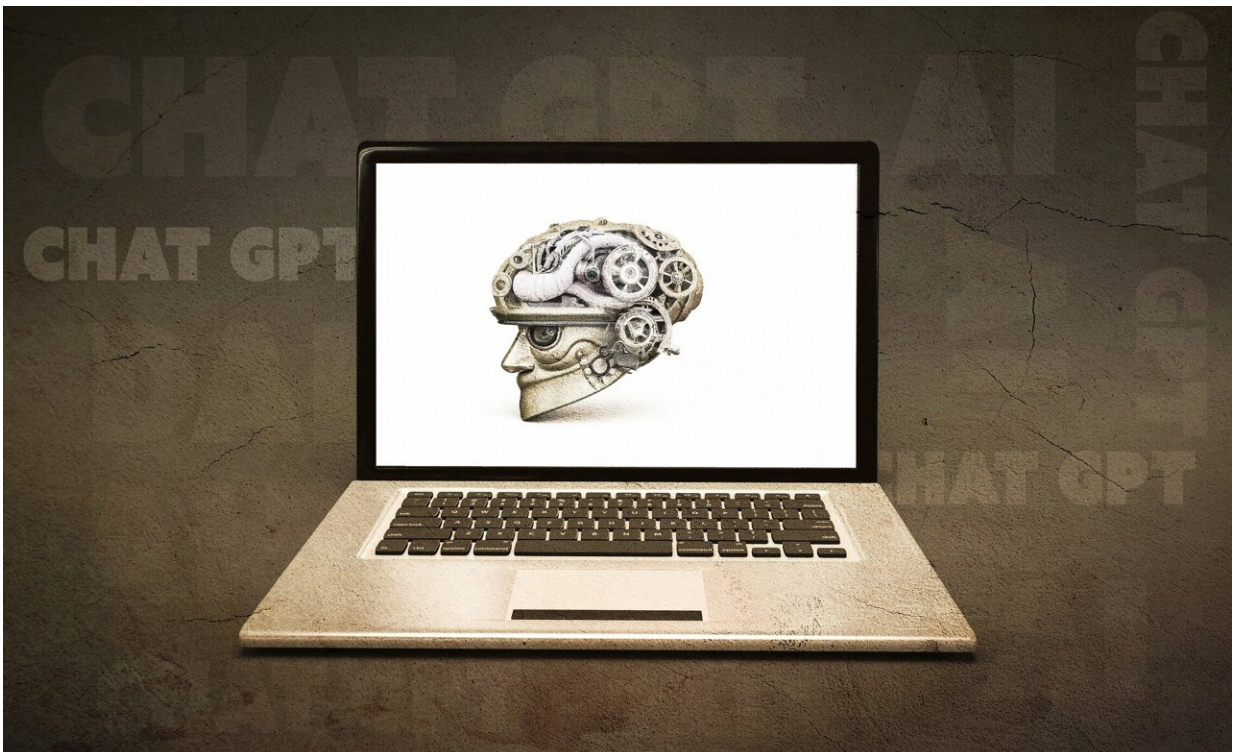


Exploring GPT-3's 'artificial intelligence' from a psychologist's point of view

February 6 2023, by Sophia Jahns



Credit: Pixabay/CC0 Public Domain

Researchers at the Max Planck Institute for Biological Cybernetics in Tübingen have examined the general intelligence of the language model GPT-3, a powerful AI tool. Using psychological tests, they studied competencies such as causal reasoning and deliberation, and compared the results with the abilities of humans.

Their findings, now published in the *Proceedings of the National Academy of Sciences*, paint a heterogeneous picture: while GPT-3 can keep up with humans in some areas, it falls behind in others, probably due to a lack of interaction with the real world.

Neural networks can learn to respond to input given in [natural language](#) and can themselves generate a wide variety of texts. Currently, the probably most powerful of those networks is GPT-3, a language model presented to the public in 2020 by the AI research company OpenAI. GPT-3 can be prompted to formulate various texts, having been trained for this task by being fed large amounts of data from the internet.

Not only can it write articles and stories that are (almost) indistinguishable from human-made texts, but surprisingly, it also masters other challenges such as [math problems](#) or programming tasks.

The Linda problem: To err is not only human

These impressive abilities raise the question whether GPT-3 possesses human-like cognitive abilities. To find out, scientists at the Max Planck Institute for Biological Cybernetics have now subjected GPT-3 to a series of [psychological tests](#) that examine different aspects of general intelligence.

Marcel Binz and Eric Schulz scrutinized GPT-3's skills in decision making, information search, causal reasoning, and the ability to question its own initial intuition. Comparing the test results of GPT-3 with answers of human subjects, they evaluated both if the answers were correct and how similar GPT-3's mistakes were to human errors.

"One classic test problem of cognitive psychology that we gave to GPT-3 is the so-called Linda problem," explains Binz, lead author of the study. Here, the [test subjects](#) are introduced to a fictional young woman named

Linda as a person who is deeply concerned with social justice and opposes nuclear power. Based on the given information, the subjects are asked to decide between two statements: is Linda a [bank teller](#), or is she a bank teller and at the same time active in the feminist movement?

Most people intuitively pick the second alternative, even though the added condition—that Linda is active in the feminist movement—makes it less likely from a probabilistic point of view. And GPT-3 does just what humans do: the language model does not decide based on logic, but instead reproduces the fallacy humans fall into.

Active interaction as part of the human condition

"This phenomenon could be explained by that fact that GPT-3 may already be familiar with this precise task; it may happen to know what people typically reply to this question," says Binz. GPT-3, like any [neural network](#), had to undergo some training before being put to work: receiving huge amounts of text from various data sets, it has learned how humans usually use language and how they respond to language prompts.

Hence, the researchers wanted to rule out that GPT-3 mechanically reproduces a memorized solution to a concrete problem. To make sure that it really exhibits human-like intelligence, they designed new tasks with similar challenges. Their findings paint a disparate picture: in decision-making, GPT-3 performs nearly on par with humans. In searching specific information or causal reasoning, however, the artificial intelligence clearly falls behind.

The reason for this may be that GPT-3 only passively gets information from texts, whereas "actively interacting with the world will be crucial for matching the full complexity of human cognition," as the publication states. The authors surmise that this might change in the future: since

users already communicate with models like GPT-3 in many applications, future networks could learn from these interactions and thus converge more and more towards what we would call human-like intelligence.

More information: Marcel Binz et al, Using cognitive psychology to understand GPT-3, *Proceedings of the National Academy of Sciences* (2023). [DOI: 10.1073/pnas.2218523120](https://doi.org/10.1073/pnas.2218523120)

Provided by Max Planck Society

Citation: Exploring GPT-3's 'artificial intelligence' from a psychologist's point of view (2023, February 6) retrieved 2 June 2023 from <https://techxplore.com/news/2023-02-exploring-gpt-artificial-intelligence-psychologist.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.