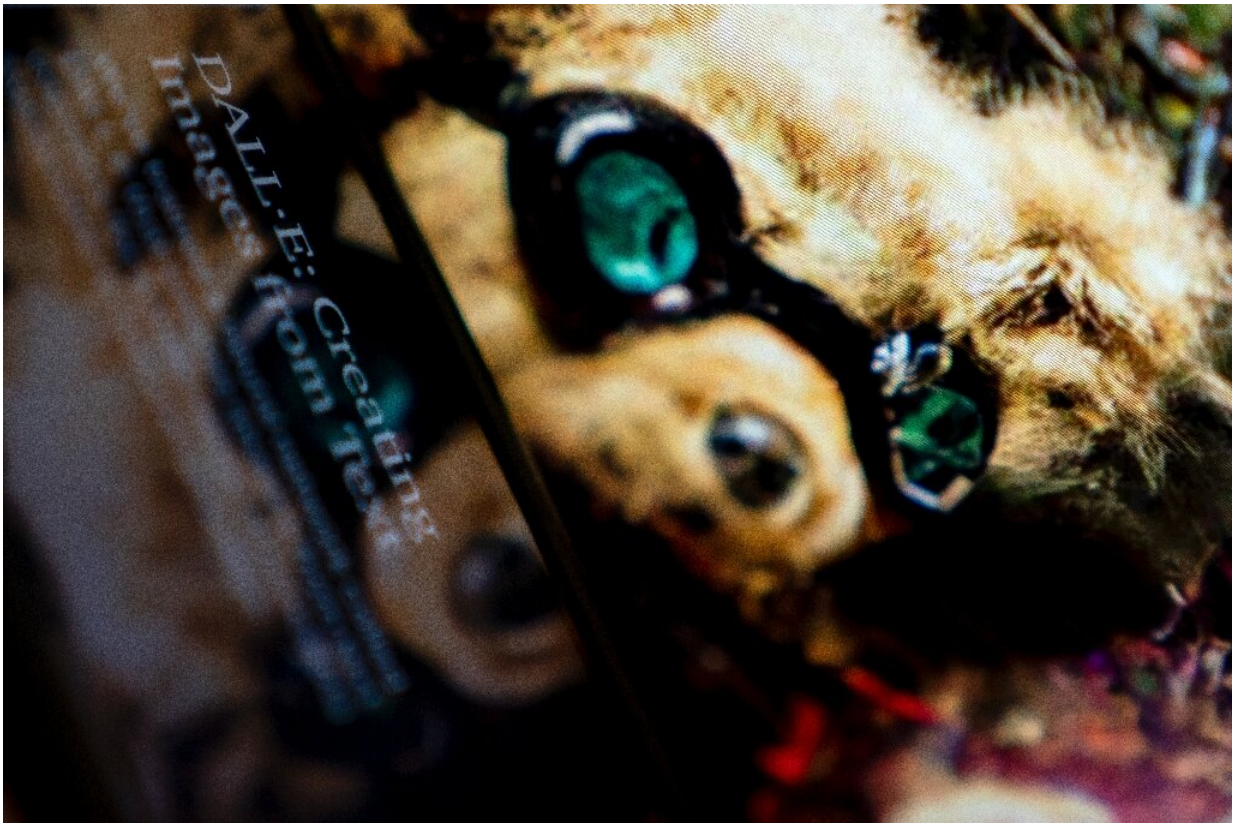# As chatbot sophistication grows, AI debate intensifies

December 4 2022, by Julie Jammot with Laurent Barthelemy in Paris

The start-up OpenAI designs sophisticated artificial intelligence software capable of generating images (DALL-E) or text (GPT-3, ChatGPT).

California start-up OpenAI has released a chatbot capable of answering a variety of questions, but its impressive performance has reopened the

debate on the risks linked to artificial intelligence (AI) technologies.

The conversations with ChatGPT, posted on Twitter by fascinated users, show a kind of omniscient machine, capable of explaining scientific concepts and writing scenes for a play, university dissertations or even functional lines of computer code.

"Its answer to the question 'what to do if someone has a [heart attack](#)' was incredibly clear and relevant," Claude de Loupy, head of Syllabs, a French company specialized in automatic text generation, told AFP.

"When you start asking very specific questions, ChatGPT's response can be off the mark," but its overall performance remains "really impressive," with a "high linguistic level," he said.

OpenAI, cofounded in 2015 in San Francisco by billionaire tech mogul Elon Musk, who left the business in 2018, received $1 billion from Microsoft in 2019.

The start-up is best known for its automated creation software: GPT-3 for text generation and DALL- E for image generation.

ChatGPT is able to ask its interlocutor for details, and has fewer strange responses than GPT-3, which, in spite of its prowess, sometimes spits out absurd results, said De Loupy.

Cicero

"A few years ago chatbots had the vocabulary of a dictionary and the memory of a goldfish," said Sean McGregor, a researcher who runs a database of AI-related incidents.

"Chatbots are getting much better at the 'history problem' where they act

in a manner consistent with the history of queries and responses. The chatbots have graduated from goldfish status."

Like other programs relying on [deep learning](#), mimicking neural activity, ChatGPT has one major weakness: "it does not have access to meaning," says De Loupy.

The software cannot justify its choices, such as explain why its picked the words that make up its responses.

AI technologies able to communicate are, nevertheless, increasingly able to give an impression of thought.

Researchers at Facebook-parent Meta recently developed a computer program dubbed Cicero, after the Roman statesman.

The software has proven proficient at the board game Diplomacy, which requires negotiation skills.

"If it doesn't talk like a real person—showing empathy, building relationships, and speaking knowledgeably about the game—it won't find other players willing to work with it," Meta said in research findings.

In October, Character.ai, a start-up founded by former Google engineers, put an experimental [chatbot](#) online that can adopt any personality.

Users create characters based on a brief description and can then "chat" with a fake Sherlock Holmes, Socrates or Donald Trump.

'Just a machine'

This level of sophistication both fascinates and worries some observers, who voice concern these technologies could be misused to trick people, by spreading false information or by creating increasingly credible scams.

What does ChatGPT think of these hazards?

"There are potential dangers in building highly sophisticated chatbots, particularly if they are designed to be indistinguishable from humans in their language and behavior," the chatbot told AFP.

Some businesses are putting safeguards in place to avoid abuse of their technologies.

On its welcome page, OpenAI lays out disclaimers, saying the chatbot "may occasionally generate incorrect information" or "produce harmful instructions or biased content."

And ChatGPT refuses to take sides.

"OpenAI made it incredibly difficult to get the model to express opinions on things," McGregor said.

Once, McGregor asked the chatbot to write a poem about an ethical issue.

"I am just a machine, A tool for you to use, I do not have the power to choose, or to refuse. I cannot weigh the options, I cannot judge what's right, I cannot make a decision On this fateful night," it replied.

On Saturday, OpenAI cofounder and CEO Sam Altman took to Twitter, musing on the debates surrounding AI.

"Interesting watching people start to debate whether powerful AI systems should behave in the way users want or their creators intend," he wrote.

"The question of whose values we align these systems to will be one of the most important debates society ever has."

Citation: As chatbot sophistication grows, AI debate intensifies (2022, December 4) retrieved 2 May 2024 from https://techxplore.com/news/2022-12-chatbot-sophistication-ai-debate.html