nature

CAREER COLUMN │ 14 Augest 2024

Chatbots in science: What can ChatGPT do for you?

Milton Pividori spent a year and a half studying how best to use ChatGPT in research. Here, he highlights three key lessons.

By Milton Pividori

Milton Pividori has explored how best to use AI chatbots to improve science.Credit: Kerkhoff Photography & Design

ChatGPT stunned the world on its launch in November 2022. Powered by a large language model (LLM) and trained on much of the text published on the Internet, the artificial intelligence (AI) chatbot, created by OpenAI in San Francisco, California, makes the latest advances in natural-language processing broadly accessible by providing a dialogue-based interface capable of answering complex questions, composing sophisticated essays and generating source code. One obvious question was: how could this tool improve science?

Over the past 18 months, with funding from the non-profit organizations Alfred P. Sloan Foundation in New York City and the Chan Zuckerberg Initiative in Redwood City, California, my laboratory has been exploring ways to incorporate the technology into daily tasks, such as conducting literature reviews, revising and writing scholarly text and programming code. Our goal is to assess how we can safely use this technology to produce better science and increase productivity. Here, we highlight some key lessons.

Engineer your prompt

To use a chatbot effectively, you need a good prompt. That might sound obvious, but some of my colleagues still get frustrated and give up when the tool fails to answer a poorly articulated question. This is understandable: the public has been bombarded with the idea that these models are 'intelligent', so it makes sense to think that they should understand whatever you ask. But this isn't true, which is why prompt engineering has become a fast-growing discipline in the field.

There are a lot of nuances to good prompt design, but the basic principles are simple:

• Be clear about what you want the model to do (use commands such as 'Summarize' or 'Explain').

• Ask the model to adopt a role or persona ('You are a professional copy editor').

• Provide examples of real input and output, potentially covering tricky 'corner' cases, that show the model what you want it to do.

• Specify how the model should answer ('Explain it to someone who has a basic understanding of epigenetics') or even the exact output format (for instance, as an

analysis-friendly JSON or CSV file).

• Optionally, specify a word limit, whether the text should use the active or passive voice, and any other requirements. Check out the 'Prompt Engineering Cheat Sheet' for more tips.

Here is a prompt that we use to revise manuscript abstracts, which we crafted on the basis of guidelines1 published in 2017.

You are a professional copy editor with ample experience handling scientific texts. Revise the following abstract from a manuscript so that it follows a context–content–conclusion scheme. (1) The context portion communicates to the reader the gap that the paper will fill. The first sentence orients the reader by introducing the broader field. Then, the context is narrowed until it lands on the open question that the research answers. A successful context section distinguishes the research's contributions from the current state of the art, communicating what is missing in the literature (that is, the specific gap) and why that matters (that is, the connection between the specific gap and the broader context). (2) The content portion (for example, 'here, we …') first describes the new method or approach that was used to fill the gap, then presents an executive summary of results. (3) The conclusion portion interprets the results to answer the question that was posed at the end of the context portion. There might be a second part to the conclusion portion that highlights how this conclusion moves the broader field forward (for example, 'broader significance').

Find the right tasks

When considering potential applications, ask yourself how much creativity the task requires, and what could happen if the model steers you wrong. What are the aspects of the task that only a person could contribute, and which are more mechanical — and usually, boring?

Take the literature-review stage of a research project, for instance. The goal of this iterative process is to produce a refined list of articles with a summary of their main ideas. This sounds like the perfect task for a chatbot assistant, and it is — but not at first. Defining a research question involves creative thinking; you need to read papers carefully, identify research gaps, develop a hypothesis and start thinking about how

you could address the problem experimentally. You probably want to understand as much as you can about each paper, including its figures, tables and supplementary materials. A chatbot might omit key information and, importantly, could prevent you from drawing creative and logical connections.

Milton Pividori and his collaborators developed a tool that integrates ChatGPT edits into the collaborative-writing tool Manubot.Credit: Milton Pividori and Casey S. Greene

Later in the process, however, your goals will be different. At this point, you might want to quickly 'read' (that is, summarize) articles that are less directly related to your work. In this case, using a chatbot assistant is less risky.

Our team has found some success using specialized tools (such as SciSpace) to search for articles, assess their relevance and 'chat' with the text. But general tools, such as ChatGPT, have been less useful. Whichever platform you choose, use standard search engines as well, to maximize the number of relevant papers that you find.

Write more, read less

In my experience, using a chatbot to write is less risky than using it to read. To have an LLM 'read' a paper, you have to trust it to accurately extract the most important points, because you might not read the article yourself. But when using it to write, you have complete control over the output and can catch 'hallucinations' — text that is nonsensical or inaccurate — when they occur.

When I start writing a manuscript, I already know what I want to say, but I often need help with crafting the text. In this case, it's useful to feed the chatbot the rules for structuring a scientific manuscript in your discipline1. As an alternative, you can write without help at first, and then use a chatbot to revise the text (for example, to apply the context–content–conclusion structure to a paragraph), review its suggestions and implement the good ones. When I was a postdoc, my colleagues and I developed an AI editor for the collaborative-writing framework Manubot. The editor takes a human-centric approach to automating the writing process2: a person first writes the text, the LLM revises their work and then the author reviews the changes. The tool uses the version-control service GitHub to keep track of which portions of the text were contributed by the user and which by the model — this can be important to document, considering that at least one contributor to Nature has been falsely accused of using a chatbot to write their manuscript.

When writing source code with a chatbot, you can take a similar approach: ask the LLM for code either to solve a problem or fix existing, buggy code. If you know what you want your code to do (the creative part), you'll need to write a prompt to instruct the model which language and libraries to use (the mechanical part). Then, you run the code to see whether it works. The worst that will happen is that the code produces the wrong outcome or conveys the wrong conclusion. You need to check the code carefully, even if you get what seems to be the right answer — and for that, you need to understand it.

This is a key point, especially for trainees: if you don't know how to do something, I strongly discourage you from using a chatbot to do it for you.

As LLMs become increasingly capable, they can help scientists to focus on the creative and challenging aspects of their work and offload the less intellectually stimulating

parts. The challenge is identifying those tasks that only humans can do — and recognizing the limitations that LLMs still pose.
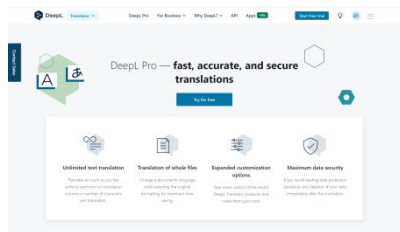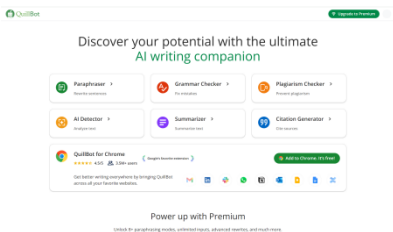
doi: https://doi.org/10.1038/d41586-024-02630-z

This is an article from the Nature Careers Community, a place for Nature readers to share their professional experiences and advice. Guest posts are encouraged.

Recommended AI tools

Latest on:

<u>AI Tools</u>   Machine learning   Authorship

QuillBot:

your complete write solution in English

Deepl:

fast、accurate、 and secure transations

VersaBot:

academic writing assistant inspire new academic ideas

AITools COLUMN |

09 APR 24

AITools FEATURE |

08 APR 24

AITools COLUMN |

08 APR 24