AI prompt engineering: learn how not to ask a chatbot a silly question

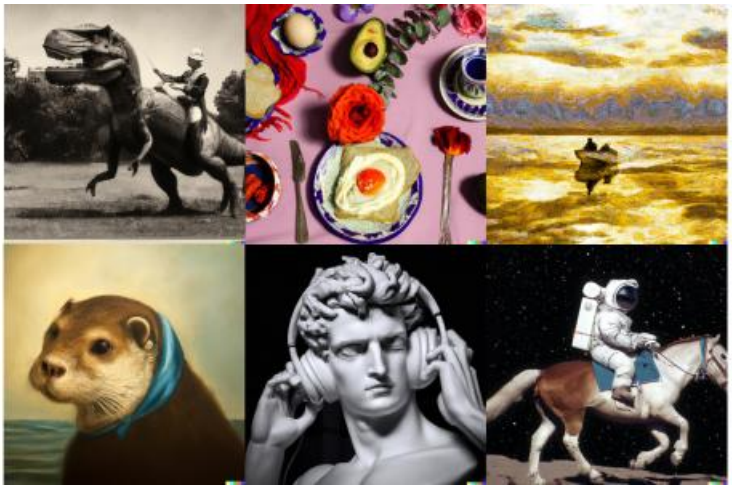
Publication Date: 2023-07-29

Author: Unknown

Section: Technology

Tags: Artificial intelligence (AI), The Observer, Computing, Consciousness, ChatGPT, Chatbots, features

Article URL: https://www.theguardian.com/technology/2023/jul/29/ai-prompt-engineering-chatbot-questions-art-writing-dalle-midjourney-chatgpt-bard



After all the initial excitement over ChatGPT, the language-processing tool driven by artificial intelligence (AI), the use of chatbots is becoming more commonplace. So how do you train your AI for work and home? We answer a few simple questions. What is prompt engineering? It's a technique for effectively communicating with generative AI models. Systems such as ChatGPT, Bard and Dall-E will produce text, images and snippets of music when fed an input – called a prompt – that instructs them what to generate. But the phrasing of a prompt can drastically alter the returned output. Prompt engineering is the process of formulating a prompt for an AI system so that it produces an output that closely matches your expectations. How is it different from just asking questions? It requires more care. Throw a question from the top of your head at ChatGPT and it may provide a satisfying answer, or not. Prompt engineering involves considering the idiosyncrasies of an AI model to construct inputs that it will clearly understand. This tends to produce outputs that are more consistently useful, interesting and appropriate to what you have in mind. Formulate the prompt well and the response may even surpass expectations. Why should I care? Chatbots such as ChatGPT, Bard and Bing Chat can be tremendously convenient for completing everyday administrative tasks. Advocates have used them to draft emails, summarise meeting notes, compose contracts, plan holidays and provide answers to complex questions nearly instantaneously. "Anybody can have one of the most powerful personal assistants on the planet that makes them more productive or allows them to create things they wouldn't normally," says Jules White, an associate professor of computer science at Vanderbilt University in Nashville, Tennessee. "But you have to understand how to interact with it." And that means knowing how to prompt effectively. A touch of prompting savvy may also impress hiring managers. Matt Burney, a talent strategy adviser at careers website Indeed, says the number of job ads asking for AI proficiency is small but growing, and companies across various industries are increasingly looking at how to integrate the models into their workflows. "If you're not using it right now, you are going to be behind the curve of those that are," he says. So how do I do it? There are several popular prompting techniques. Employing personas is a common trick. Tell the system to act as a lawyer, personal tutor, drill sergeant or whatever else, and it will create outputs imitating their tone and voice. Or, as a reverse exercise, instruct it to complete a task with a specific audience in mind – a five-year-old, a team of expert biochemists, an office Christmas party – and you'll get a result tailored for that demographic. Crucially, you don't need to know the persona's stylistic characteristics yourself, but can leave that to the system to figure out. Chain-of-thought prompting, meanwhile, is more appropriate for problem-solving. Asking the model to "think step by step" will encourage it to partition its output into bite-size chunks, which often makes for more comprehensive results. Some researchers

have also found that showing an AI model an example problem with its step-by-step solution will improve its ability to hit upon the correct answer when solving other, similar questions. In fact, examples never hurt. If you have a very specific output in mind, upload a text sample or an image illustrating what you want generated and instruct the model to use it as a template. If the result is initially off target, a few more rounds of clearly specified tinkering could do the trick. “You want to think of it as a continuing conversation where you start and you iterate and refine,” says White And don’t forget the basics of everyday language: clear, imperative instructions that minimise misinterpretation. Explicitly state what you do and do not want from the output, and set a clear word count and format. What should I avoid? Vague language. Without additional information, AI models cannot infer your tastes, ideas or the vision of the product that’s in your head. Don’t skimp on specifics or context and don’t assume that if something is missing, the model will correctly fill in the blank. Can it stop AI from spouting inaccuracies? No. Large language models will fabricate sources even when explicitly instructed not to and provide information that sounds plausible but is entirely false. “That’s an intractable problem with these models,” says Mhairi Aitken, an ethics fellow at the Alan Turing Institute, based at the British Library in London. “They’re designed to predict a sequence of words that replicate human language, but there’s no connection to truth or reality.” Shrewd prompting can, however, help deal with falsehoods after they appear. “If the chatbot makes incorrect claims, you can point out the errors and ask it to rewrite the answer based on your feedback,” says Marcel Scharth, a lecturer in business analytics at the University of Sydney. White suggests asking the model to produce a list of the fundamental facts on which its output relies, so you can verify them individually. Or provide it with a numbered list of facts on which to base its answer and have it reference each when they’re used, to speed up factchecking later. Could this be a career? For some people, maybe. AI developers have hired prompt engineers to test the limitations and deficiencies of their models so they can be refined to better handle user inputs. But the longevity of these positions isn’t guaranteed. Rhema Linder, a lecturer in computer science at the University of Tennessee, suggests developers may come to prefer specialised computer scientists to self-styled prompt engineers, and the absence of industry-recognised certification means assessing a person’s prompting ability is difficult. In the wider jobs market, prompt engineering will probably go the way of spreadsheet management or search engine optimisation – a skill demanded in a variety of roles and prized by hiring managers as another feather in the cap of your CV. “Experience of using a large language model or generative pretrained transformer is going to be a requirement for pretty much every office-based job,” says Burney. “Because if you can’t do it, you’re going to be slower achieving your goals.” Will this all become obsolete? Just as the AI models aren’t stable, neither are prompt engineering best practices. The techniques that work with systems now may prove less useful in updated versions, although it’s unclear how sweeping the changes could be. “I think there will be core concepts and patterns that don’t change,” says White, who suggests AI developers will take note of common prompting techniques. “A lot of these ways of phrasing things are going to become the benchmarks that the new models are trained against, so some prompt engineering will feed back on the models themselves.” More significantly, the models’ abilities to comprehend even the vaguest, un-engineered prompts could improve dramatically. “As these systems become more conversational, and as interacting with them becomes more intuitive, we maybe don’t need prompt engineering in the future,” says Aitken. For some developers, that’s the goal.