

The danger of relying on OpenAI's Deep Research

Economists are in raptures, but they should be careful

The Economist (Free Exchange), Feb 13th 2025

In early February OpenAI, the world's most famous artificial-intelligence firm, released Deep Research, which is "designed to perform in-depth, multi-step research". With a few strokes of a keyboard, the tool can produce a paper on any topic in minutes. Many academics love it. "Asking OpenAI's Deep Research about topics I am writing papers on has been incredibly fruitful," said Ethan Mollick of the University of Pennsylvania. Some economists go further. "I am *sure* for B-level journals, you can publish papers you 'wrote' in a day", said Kevin Bryan of the University of Toronto. "I think of the quality as comparable to having a good PhD-level research assistant, and sending that person away with a task for a week or two," said Tyler Cowen of George Mason University, an economist with cult-like status in Silicon Valley.

Should you shell out \$200 a month for Deep Research? Mr Cowen has hyped fads in the past, as he did with Web3 and Clubhouse, a once-popular social-media network. On the other hand, if Deep Research approximates a form of artificial superintelligence, as many believe, then \$2,400 a year is the greatest bargain in the history of the world. To help you decide, your columnist has kicked the tyres of the new model. How good a research assistant is Deep Research, for economists and others?

The obvious conclusions first. Deep Research is unable to conduct primary research, from organising polls in Peru to getting a feel for the body language of a chief executive whose company you might short. Nor can it brew a coffee, making it a poor substitute for a human assistant. Another complaint is that Deep Research's output is almost always leaden prose, even if you ask it to be more lively. Then again, most people were never good writers anyway, so will hardly care if their ai assistant is a bit dull.

Use Deep Research as an assistant for a while, though, and three more important issues emerge: "data creativity", the "tyranny of the majority" and "intellectual shortcuts". Begin with data creativity. OpenAI's model can handle straightforward questions—"what was France's unemployment rate in 2023?"—without breaking step. It can handle marginally more complex questions—"tell me the average unemployment rate in 2023 for France, Germany and Italy, weighted by population"—with ease.

When it comes to data questions requiring more creativity, however, the model struggles. It wrongly estimates the average amount of money that an American household headed by a 25- to 34-year-old spent on whisky in 2021, even though anyone familiar with the Bureau of Labour Statistics data can find the exact answer (\$20) in a few seconds. It cannot accurately tell you what share of British businesses currently use ai, even though the statistics office produces a regular estimate. The model has even greater difficulty with more complex questions, including those involving the analysis of source data produced by statistical agencies. For such questions, human assistants retain an edge.

The second issue is the tyranny of the majority. Deep Research is trained on an enormous range of public data. For many tasks, this is a plus. It is astonishingly good at producing detailed, sourced summaries. Mr Cowen asked it to produce a ten-page paper explaining David Ricardo's theory of rent. The output would be a respectable addition to any textbook.

Yet the sheer volume of content used to train the model creates an intellectual problem. Deep Research tends to draw on ideas that are frequently discussed or published, rather than the best stuff. Information volume tyrannises information quality. It happens with statistics: Deep Research is prone to consulting sources that are easily available (such as newspapers), rather than better data that may be behind a paywall or are harder to find.

Something similar happens with ideas. Consider the question—much discussed by economists—of whether American income inequality is rising. Unless prompted to do otherwise, the model blandly assumes that inequality has soared since the 1960s (as is the conventional wisdom) rather than remained flat or increased only a bit (the view of many experts). Or consider the true meaning of Adam Smith's "invisible hand", the foundational idea in economics. In a paper published in 1994, Emma Rothschild of Harvard University demolished the notion that Smith used the term to refer to the benefits of free markets. Deep Research is aware of Ms Rothschild's research but nonetheless repeats the popular misconception. In other words, those using Deep Research as an assistant risk learning about the consensus view, not that of the cognoscenti. That is a huge risk for anyone who makes their income through individual creativity and thought, from public intellectuals to investors.

The idiot trap

A third problem with employing Deep Research as an assistant is the most serious. It is not an issue with the model itself, but how it is used. Ineluctably, you find yourself taking intellectual shortcuts. Paul Graham, a Silicon Valley investor, has noted that AI models, by offering to do people's writing for them, risk making them stupid. "Writing is thinking," he has said. "In fact there's a kind of thinking that can only be done by writing." The same is true for research. For many jobs, researching is thinking: noticing contradictions and gaps in the conventional wisdom. The risk of outsourcing all your research to a supergenius assistant is that you reduce the number of opportunities to have your best ideas.

With time, OpenAI may iron out its technical issues. At some point, Deep Research may also be able to come up with amazing ideas, turning it from an assistant to the lead researcher. Until then, use Deep Research, even at \$200 a month. Just don't expect it to replace research assistants any time soon. And make sure it doesn't turn you stupid. ■