

OpenAI Introduces ChatGPT Agent That Can 'Use Virtual Computers' to Complete Complex Tasks

ChatGPT agent can now perform work using its own computer.

OpenAI has offered two distinct types of agents: Operators (which can browse the web and perform tasks independently) and Deep Researchers (which specialize in synthesizing large amounts of online information). Today, the company announced ChatGPT agents, a new AI that combines Operator's web browsing capabilities, Deep Research's research strengths, and ChatGPT's conversational skills into a single powerful agent.

ChatGPT agents can now perform tasks using their own virtual desktop. Based on user queries, they can navigate websites, filter results, prompt users to log in when required, run code, perform analytics, create spreadsheets and PowerPoints, and more.

ChatGPT agent will have access to the following tools to complete tasks assigned by the user:

1. A visual web browser that interacts with the web through a GUI
2. A text-based browser for simple web queries based on reasoning
3. A terminal
4. Direct API access
5. Ability to connect to ChatGPT connectors.

Since the ChatGPT agent does all of its work using its own virtual machine, it will have all the context it needs to complete the task. For example, the agent can access a website using a browser, download a file from the website, manipulate that same file by running a command in the terminal, and then view the output back in a visual browser.

OpenAI claims that the ChatGPT agent achieves state-of-the-art (SOTA) performance on various benchmarks that measure web browsing and real-world task completion. Here are some highlights:

1. Humanity's Last Exam: ChatGPT agent achieved a new SOTA pass@1 score of 41.6. When running up to 8 trials at once and choosing the trial with the highest self-reported confidence, the score increased to 44.4.
2. FrontierMath: ChatGPT agent achieves 27.4% accuracy.
3. OpenAI's internal benchmark (evaluating model performance on complex, economically valuable knowledge work tasks): ChatGPT agent output is equal to or better than humans in about half of the cases.
4. DSbench?: ChatGPT agent outperforms human performance by a significant margin on data science tasks.
5. SpreadsheetBench: ChatGPT agent scored 45.5%, compared to 20.0% for Copilot in Excel.
6. BrowseComp?: ChatGPT agent sets new SOTA at 68.9%.

7. WebArena: ChatGPT agent reached 65.4%.

ChatGPT agent is now available in the ChatGPT tools menu with a new 'agent mode'. While the agent is performing a task, users can find a narration on the screen; they can also interrupt and take control of the browser whenever needed.

ChatGPT agents will be available to all ChatGPT Pro users soon. ChatGPT Plus and Team users will get access in the coming days, while Enterprise and Education users will have to wait a few more weeks. ChatGPT Pro users will be limited to 400 messages per month with the agent, while other paid plans are limited to 40 messages per month. However, users can purchase additional agent usage with flexible credit-based options.

You finished reading the article "**OpenAI Introduces ChatGPT Agent That Can 'Use Virtual Computers' to Complete Complex Tasks**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.