Anthropic's Claude Cowork Is an AI Agent That Actually Works

Cowork is a user-friendly version of Anthropic's Claude Code AI-powered tool that's built for file management and basic computing tasks. Here's what it's like to use it.

Reece Rogers   Jan 15, 2026 12:40 PM

Courtesy of Anthropic/Claude

As a software reporter at WIRED, I've tested a lot of shitty agents over the past couple of years. These experiences expose a consistent pattern of generative AI startups overpromising and underdelivering when it comes to these "agentic" helpers—programs designed to take control of your computer, performing chores and digital errands to free up your time for more important things. But the bots I installed on my laptop would struggle to complete even basic tasks. They just didn't work.

This poor track record makes Anthropic's latest agent,

Claude Cowork, a nice surprise. When I tested it by running it through some basic and intermediate demos the company suggested in addition to my own commands, it worked fairly well—especially for software that's still in beta. It can do things like organize files into folders, convert file types, generate reports, and even take over the browser to search the web or tidy up a Gmail inbox. When it comes to file management and computer interfaces, this tool feels like the start of a pleasant user experience evolution.

Last year, Anthropic nurtured a cult following for its Claude Code tool among developers who loved its ability to understand codebases and run commands, with tech staffers across San Francisco using it for their work seemingly all the time. But most people aren't members of some buzzy startup's technical staff.

"We tried a bunch of different ideas to see what form factor would make sense for a less technical audience that doesn't want to use a terminal," says Boris Cherny, Anthropic's head of Claude Code. For the past two months, Cherny has written all of his code with AI. Cowork was built using AI tools.

Released by Anthropic earlier this week as a research preview, Cowork takes the abilities available in the company's coding focused tool and makes the user experience more approachable. This tool is designed for the

wider group of nontechnical users, who may want to experiment with a new way of controlling their computers but get freaked out by a [command line](#).

Getting Started

Reece Rogers

Right now, Cowork is available only as part of a research preview to subscribers of Anthropic's $100-a-month plan, which is a [common release strategy](#) for generative AI companies soft-launching new features to early adopters.

Felix Rieseberg, a member of technical staff at Anthropic who focuses on Cowork, says he uses it to file expense reports and do file conversions. "If this PDF is too big, make it smaller," he says. "Turn these 20 JPEGs into one PDF. Make me a report about all of these things." Rieseberg is excited by how [more advanced users](#) are already experimenting with complex applications but sees the most straightforward, file-focused applications as his "favorite" uses of the research preview.

This early release is limited to the [Claude on Mac](#), with a wider rollout potentially down the line. And even though you can use it to interact with files on your computer, an internet connection is required for Cowork to run. The Cowork tab appears next to the Chat and Code tabs in the Claude app

for [macOS](). User sessions are labeled as "tasks" rather than "chats."

What About the Security Risks?

The biggest reason for not trying out Cowork is the ongoing [security]() risk inherent in these kinds of agents. Like most agents, Cowork is susceptible to prompt injection attacks, secret messages hidden online that try to trick AI tools and deviate them from tasks. You shouldn't expose sensitive data to a tool that can be compromised in this way.

"Since Claude can read, write, and permanently delete these files, be cautious about granting access to sensitive information like financial documents, credentials, or personal records," reads Anthropic's [online support page](). It suggests saving backups of critical files and creating a dedicated folder filled with nonsensitive information you want Claude to be able to access.

"We use a virtual machine under the hood," Cherny says. "This means you have to say which folders Claude has access to. And if you don't give it access to a folder, Claude literally cannot see that folder."

Cowork can also click around in your browser, if you allow it, to find information or help with your email inbox. The bot [asks for your permission]() for most steps it takes along the

way, like reading a website. This browser option for Cowork also includes an explicit disclaimer stating that hidden code in websites may "steal your data, inject malware into your systems, or take over your system" if you choose to use it.

Cherny says Anthropic designed Cowork with multiple safety mitigations, which include prompt-injection detection as well as focusing on keeping users in the loop on what the agent is doing and using "virtualization" to give the tool access to only the specifically requested files.

First Impressions

The first test I ran was to see whether [Claude](#) could organize the random assortment of screenshots, from corny memes to important invites, scattered across my desktop. To do this, I had to give Claude access to my entire desktop folder as well as the ability to adjust files, including permanently deleting things.

Before taking any actions, the chatbot asked my preferences for how it should sort the screenshots, with a recommendation to go with a separate folder for each month. Then, it spent around a minute processing the request and running commands. At the end of this endeavor, I revisited my desktop to see all the screenshots correctly sorted into three new folders, labeled by month. Neat! I hate doing that part.

Reece Rogers

Next, I wanted to ask it to do something a little more broad and involved, like organizing an email inbox. After I granted it access, Cowork asked what my goals in [Gmail] were as well as the types of clutter I wanted to focus on. The tool includes a few auto-generated answers, as optional buttons, under each question.

My first few tries at asking it to archive the [emails], instead of deleting them, ran into snags as the bot tripped up attempting to batch archive some promotional emails. As I watched it fail, I pivoted and asked Cowork to go ahead and delete a thousand of those unread messages. Then, the bot clicked around and deleted everything I asked it to … and nothing I didn't. Even so, the bot messing up is still a real possibility as Anthropic works through bugs and iterates on Cowork.

Finally, I connected Claude to a Google Calendar and asked Cowork to find me two tickets to an evening showing of *Marty Supreme* at a theater near me, then add that event to my calendar as a date night. Something a little more complicated, with financial stakes. The bot found a 9 pm showing at the Alamo Drafthouse but stopped short of buying the tickets, as a safety measure. After I took over and got the seats myself, Cowork did the calendar update.

Cowork isn't perfect in its current state, and its developers plan to keep making updates to the tool based on what users share in their feedback. Still, this is the first agent that has really clicked for me.