

Introducing Devin, the first AI software engineer



March 12th, 2024 | Written by Scott Wu

Introducing Devin, the first AI software engineer

And setting a new state of the art on the SWE-bench coding benchmark

Meet Devin, the world's first fully autonomous AI software engineer.

Devin is a tireless, skilled teammate, equally ready to build alongside you or independently complete tasks for you to review.

With Devin, engineers can focus on more interesting problems and engineering teams can strive for more ambitious goals.

Devin's Capabilities

We've raised a \$21 million Series-A led by Founders Fund. Learn more here.

engineering tasks requiring thousands of decisions. Devin can recall relevant context at every step, learn over time, and fix mistakes.

We've also equipped Devin with common developer tools including the shell, code editor, and browser within a sandboxed compute environment—everything a human would need to do their work.

Finally, we've given Devin the ability to actively collaborate with the user. Devin reports on its progress in real time, accepts feedback, and works together with you through design choices as needed.

Here's a sample of what Devin can do:

Devin can learn how to use unfamiliar technologies.

After reading a blog post, Devin runs ControlNet on Modal to produce images with concealed messages for Sara.

AI Software Engineer Plants Secr...



Devin can build and deploy apps end to end.

Devin makes an interactive website which simulates the Game of Life! It incrementally adds features requested by the user and then deploys the app to Netlify.

Devin making Game of Life!



Devin can autonomously find and fix bugs in codebases.

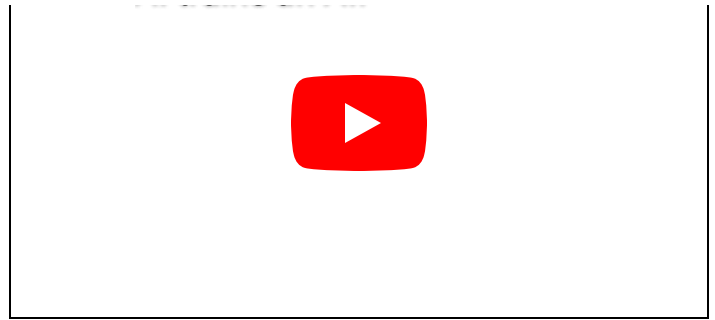
Devin helps Andrew maintain and debug his open source competitive programming book.

AI finds and fixes a bug that I did...



models.

Devin sets up fine tuning for a large language model given only a link to a research repository on GitHub.



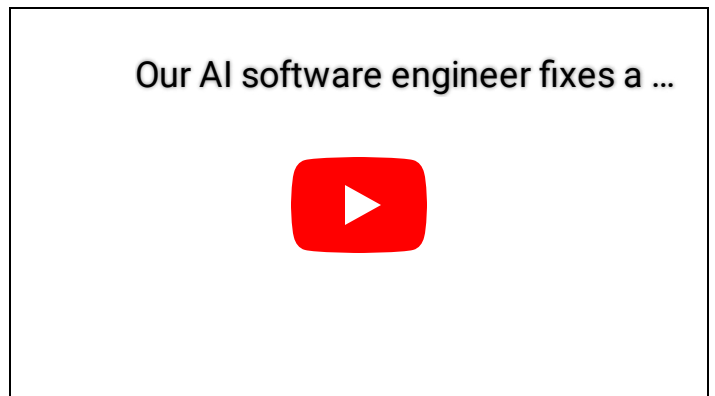
Devin can address bugs and feature requests in open source repositories.

Given just a link to a GitHub issue, Devin does all the setup and context gathering that is needed.



Devin can contribute to mature production repositories.

This example is part of the SWE-bench benchmark. Devin solves a bug with logarithm calculations in the sympy Python algebra system. Devin sets up the code environment, reproduces the bug, and codes and tests the fix on its own.

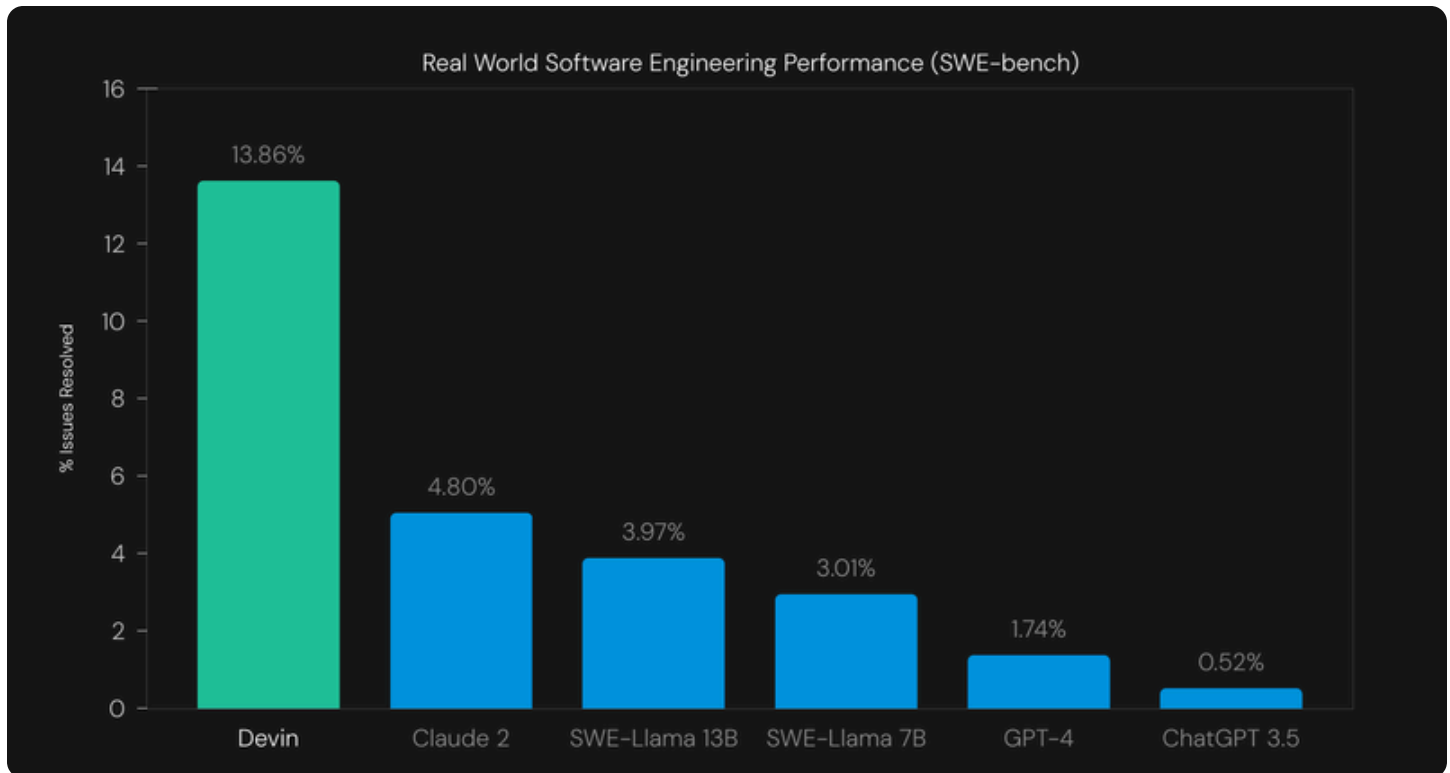


We even tried giving Devin real jobs on Upwork and it could do those too!

Here, Devin writes and debugs code to run a computer vision model. Devin samples the resulting data and compiles a report at the end.



Devin correctly resolves 13.86%* of the issues end-to-end, far exceeding the previous state-of-the-art of 1.96%. Even when given the exact files to edit, the best previous models can only resolve 4.80% of issues.



*Devin was evaluated on a random 25% subset of the dataset. Devin was unassisted, whereas all other models were assisted (meaning the model was told exactly which files need to be edited).

We plan to publish a more detailed technical report soon—stay tuned for more details.

About Cognition

We are an applied AI lab focused on reasoning.

We're building AI teammates with capabilities far beyond today's existing AI tools. By solving reasoning, we can unlock new possibilities in a wide range of disciplines—code is just the beginning. We want to help people around the world turn their ideas into reality.

We are well funded, including a **\$21 million Series A led by Founders Fund**. And we're grateful for the support of industry leaders including **Patrick and John Collison, Elad Gil, Sarah Guo, Chris Re, Eric Glyman, Karim Atiyeh, Erik Bernhardsson, Tony Xu, Fred Ehrsam** and so many more.

Hire Devin

We've raised a \$21 million Series-A led by Founders Fund. Learn more here.

work, [please reach out here](#) or get in touch at info@cognition-labs.com.

Join Us

Our team is small and talent-dense. Our founding team has 10 IOI gold medals and includes leaders and builders who have worked at the cutting edge of applied AI at companies like Cursor, Scale AI, Lunchclub, Modal, Google DeepMind, Waymo, and Nuro.

Building Devin is just the first step—our hardest challenges still lie ahead. If you're excited to solve some of the world's biggest problems and build AI that can reason, learn more about our team and [apply to join us here](#).