



Covering your bases with IP Indemnity

Deep Dive: Defining Open Source AI

Oct 2023

 **Deepak Kumar**
@deepakdk3478 · Follow

 **Tawanda Nyahuye**
@towernter · Follow



Blown away by the response of [@sourcegraph](#) Cody 🚀

I'm playing around with the [@sourcegraph](#) AI tool Cody and my mind is blown. I wanted to know what my test code snippet does and to check if it would notice code vulnerability, and it did just that.

🌟 Asked for a simple request to Shopify API:

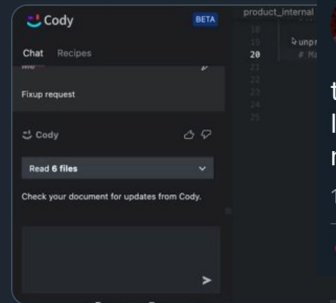
- 1 Not just added the API request code
- 2 Correctly saved the data to my models
- 3 Tossed in error handling

Cut my coding time in half 🕒
Just need to fine-tune and voila!

 **Blaise d'Estais**
@BlaisedEstais

First time I use a #AI that is really FAR "smarter" than me!
I use AI a lot for automating small tasks and saving time.

But [@SourcegraphCody](#) is a glimpse at the future, where [#GPT5](#) and [@LangChainAI](#) may help us save not just minutes but years!

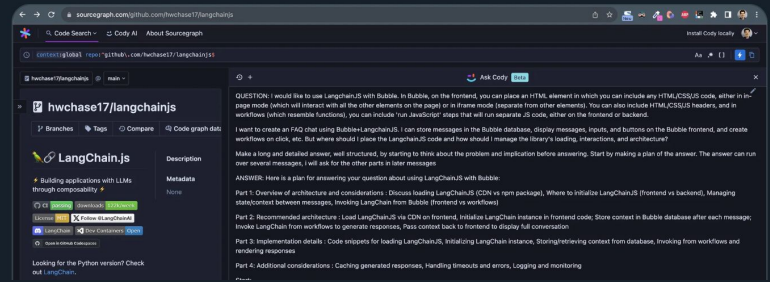
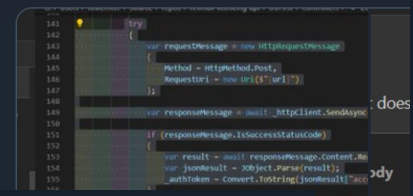


 **Waigoko Kimengemenge**
@jrobah · Follow

today I asked [@sourcegraph](#) 's Cody to explain some code I highlighted, it did and also gave suggestion on how to make the runtime faster

12:23 AM · Jun 15, 2023

👍 12 💬 Reply 🔗 Copy link




1:15 AM · Jun 19, 2023

Benchmarking Mojo code with Cody vs Julia

 **v1c7**
@vaib25vicky

Wow 🤯

 **PriNova** OP 08/24/2023 6:04 AM
Today, I made a benchmark comparison, with the help of Cody in VS Code, between the famous Mojo x35000 speedup

So I'm auditing this codebase using Cody AI of sourcegraph and I must say its the best assistant. It eases my code audit pain. I'm blown away by how convenient it is to read code now.

10:18 PM · Jul 1, 2023 · 169 Views

Thanks for this great Cody AI tool



Overview

1

Code AI platform
Justin Dorfman

What is Cody?
How Cody works.
The Copyleft question.

2

Protecting Customers
Tammy Zhu

Low copyright risk.
Litigation landscape.
Other mitigation.

3

IP Indemnity
Samantha Mandell

Cover your bases.
Reduce copyright risk.
Negotiating.

Code AI Platform

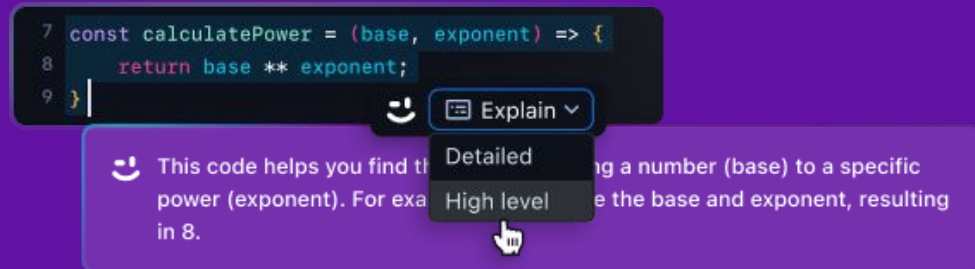
Justin Dorfman



Code AI Platform

What is Cody?

- AI coding assistant
- Writes, understands, and fixes code
- Web, VS Code, & JetBrains/Intellij
- Open source under Apache 2.0 license





Code AI Platform

What is Cody?



Momil



Deepak



Catalin



Szymon



Tino



Code AI Platform

How Cody Works

- Prompt
- LLM
- Context





Hugging Face presents

Training Cluster

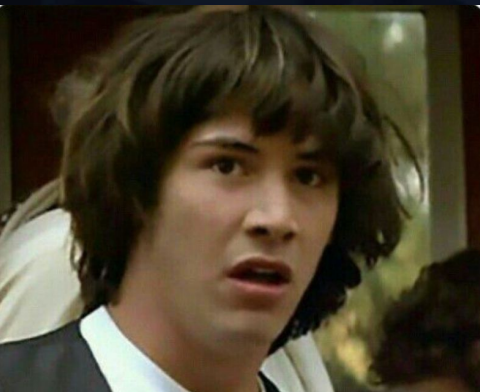
As a service

Train your LLM at scale on our infrastructure

I want to train a **70B** parameters **multimodal** model on a **7T tokens** dataset, on **1000** **H100 GPUs**

Estimate: 184 days on 1000 GPUs

\$18,461,354





Code AI Platform

The Copyleft problem

- “Does Cody use GPL code?”
- Unable to get a straight answer
- Sourcegraph Legal to the rescue

Don't worry,
I got this!



Tammy Zhu

Protecting Customers

Tammy Zhu



Protecting customers: overview

1

Copyright risk

2

Litigation
landscape

3

Other mitigation



What's the risk?

Copyright risks

- Copyleft virality
- Breach of OSS license
- Infringement







For AI coding users, copyright risk is low

Why?

In no case does copyright protection for an original work of authorship extend to any

- procedure
- process
- system
- method of operation

17 USC 102(b)



Litigation landscape: AI coding providers

No lawsuits to date against *users* of AI coding providers

No lawsuits to date against AI coding providers alleging copyright infringement

- *Doe v. GitHub* alleges the following

- DMCA violation
- Breach of contract
- Interference with economic relations
- Unjust enrichment
- Unfair competition
- Negligence

- No copyright claim

```
Copyright 2022 Sourcegraph, Inc.
```

```
Licensed under the Apache License, Version 2.0 (the "License");  
you may not use this file except in compliance with the License.  
You may obtain a copy of the License at
```

```
http://www.apache.org/licenses/LICENSE-2.0
```



Other mitigation

- Large training sets
- Permissively licensed training sets
- Filters and attribution
- OSS scans
- IP indemnity

IP Indemnity

Samantha Mandell



Indemnity Flows to the End User

What is Indemnity?

- A contractual agreement between two parties in which one party agrees to pay for potential losses or damages claimed by a third party as a result of their relationship
- Software: if Party A is sued because they use Party B's software as intended, Party B will cover Party A's losses
- Often, if there are multiple parties involved in a relationship or transaction, there might be multiple indemnification agreements



What AI tool users should look for with Indemnity

A comprehensive IP indemnity clause is:

- Uncapped
- Covers all plausible third parties (ie. individual copyright holders and open source licensors)
- Applies even if a user makes changes to the Output



Landscape of Indemnity among AI Coding Assistants

Cody (Sourcegraph)	Sample Tool #1	Sample Tool #2	Sample Tool #3
Uncapped	Caps dependent on agreement with provider for other services	None, <i>and</i> user required to indemnify provider for copyright infringement	None (but users not required to indemnify provider)
Covers all plausible 3rd parties	Covers all plausible 3rd parties		
Applies even if a user makes changes to the output	Does not apply if user makes changes to the output		
	Requires use of filtering		

Closing

Justin Dorfman