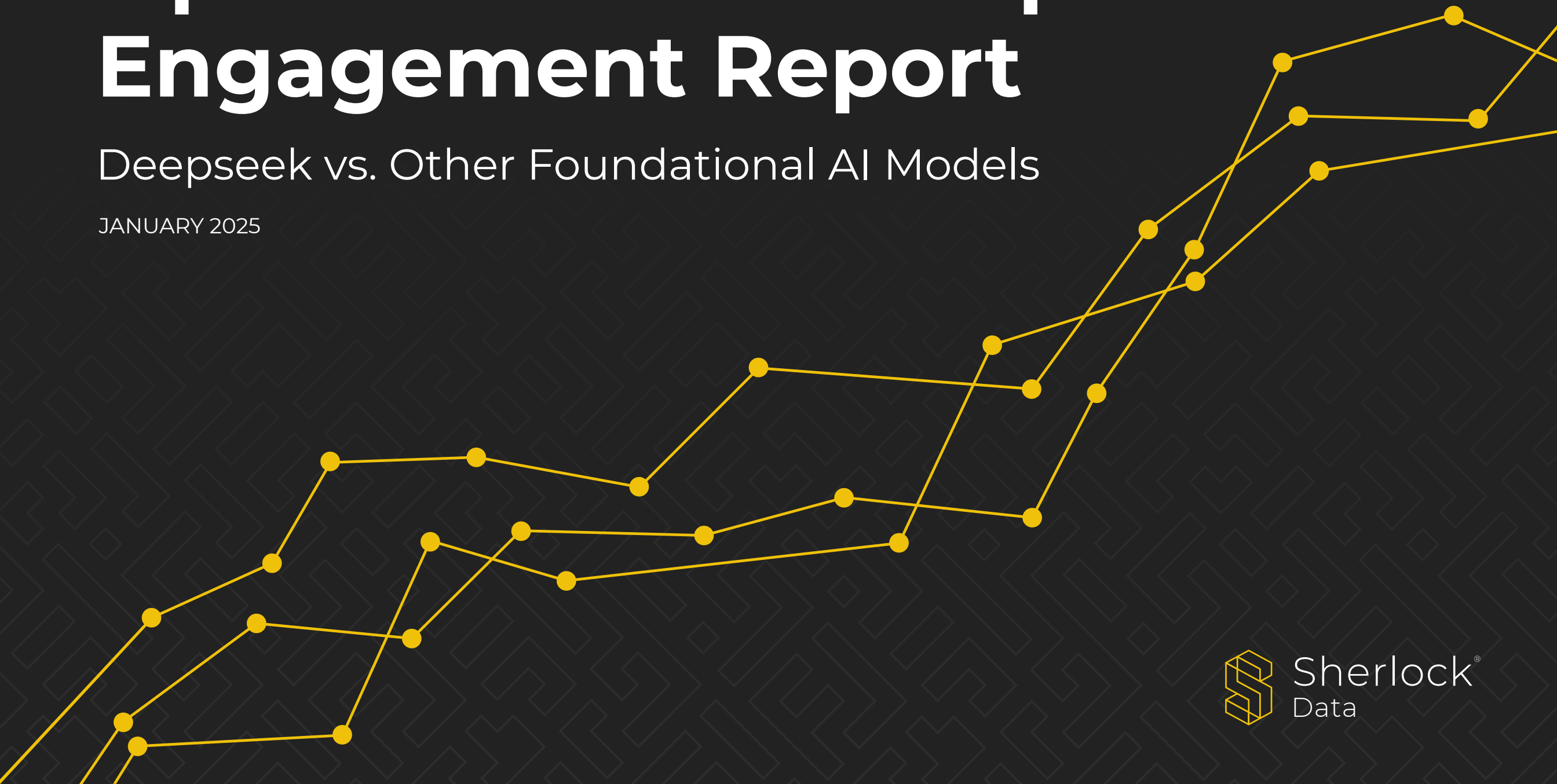


Open Source Developer Engagement Report

Deepseek vs. Other Foundational AI Models

JANUARY 2025



Sherlock[®]
Data

Are you an institutional investor looking for an edge in 2025? Check out **Sherlock Data**! Incubated at Fidelity Center for Applied Technology (FCAT®) and available for sample or purchase on [Snowflake Marketplace](#), Sherlock provides valuable insights into where software engineers are building and could support your predictions about the next big AI or blockchain company.

Why track **software engineering activity**?

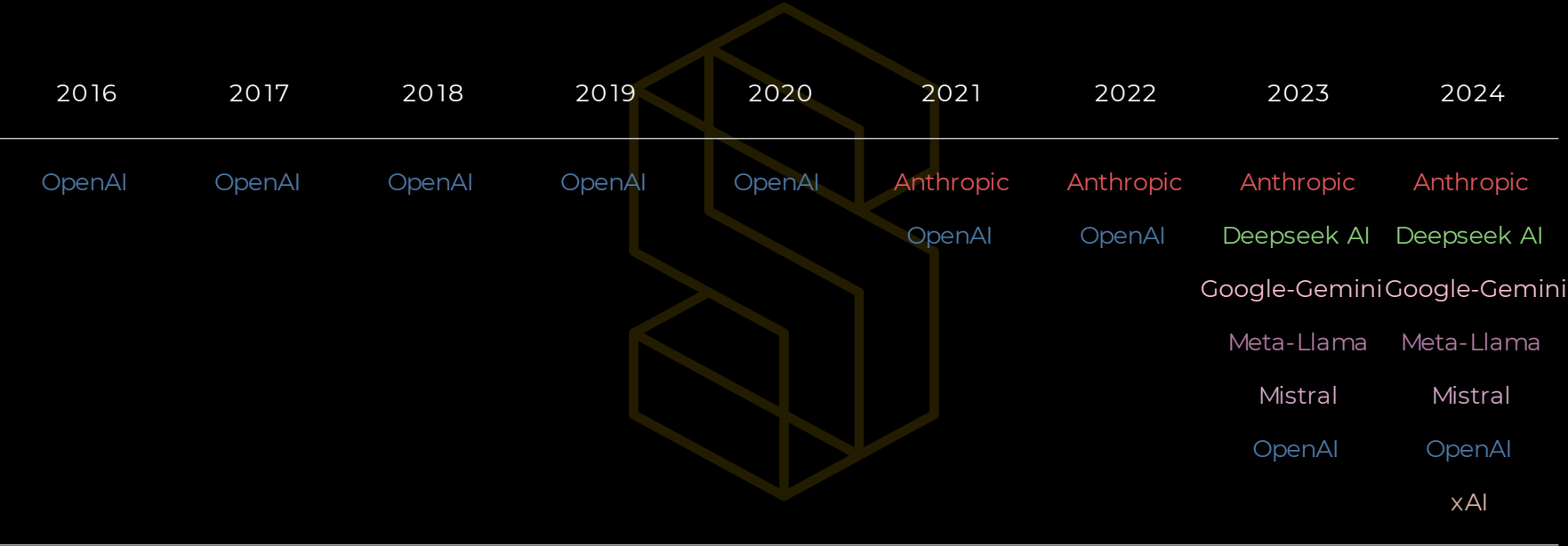
- **Monitor** AI model development efforts
- **Compare** activity across private and public companies
- **Witness** the rise or fall of a blockchain or dApp
- **Identify** thematic trends within a category
- **Manage** risk within your existing portfolio
- **Validate** an investment thesis
- **Generate** new ideas and strategies
- **Anticipate** strategic partnerships
- **Contrast** SDKs across a sector

To learn more, including how to sample or buy this data:

Visit our listing on Snowflake 

Sherlock Data only includes development activity found in certain public repositories and is not indicative of a company's entire public and private codebase or development portfolio.

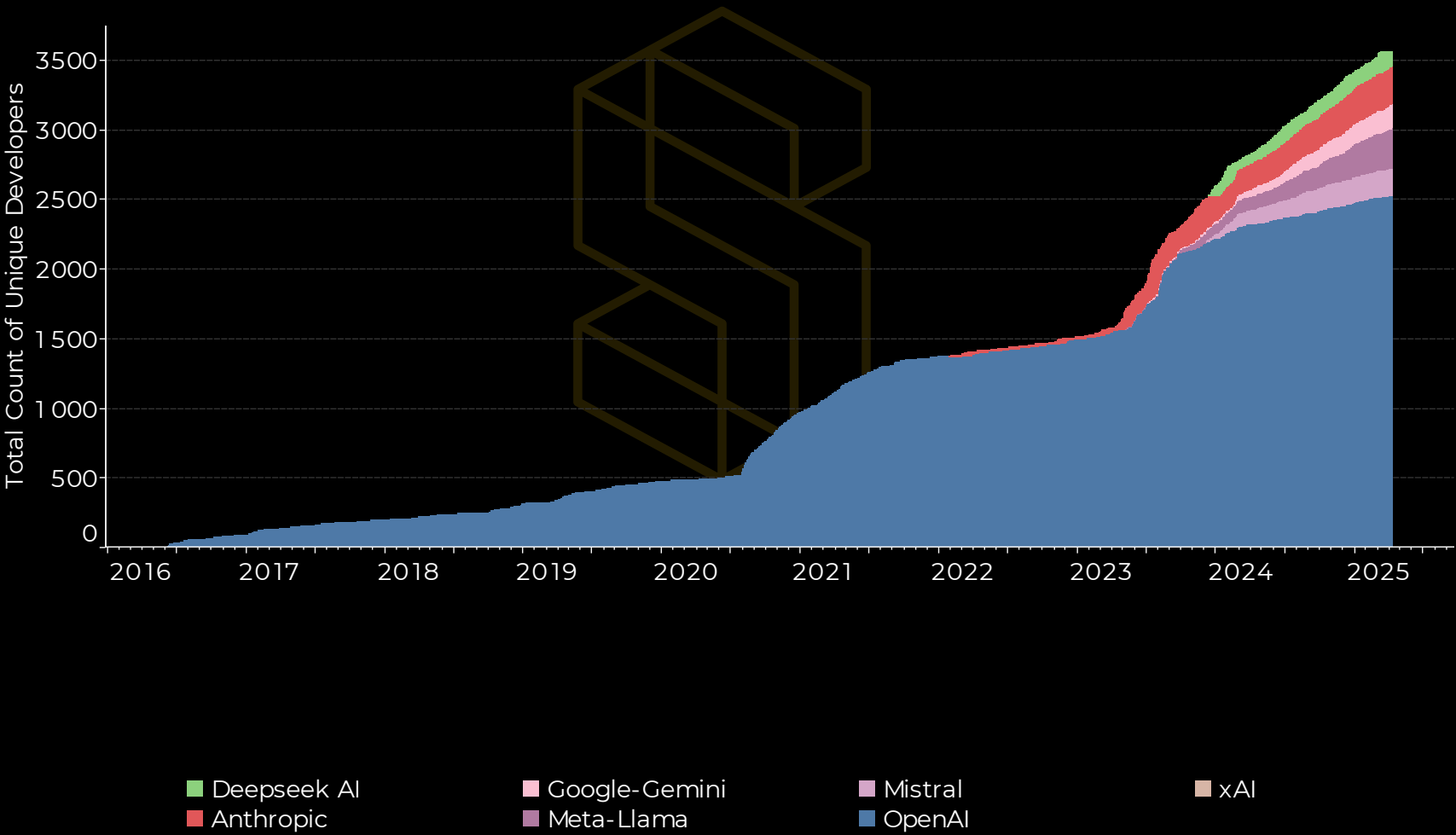
Activity in Public Repositories over Time



Deepseek first published code in a public repository during October of 2023, 7 years after OpenAI started publishing code. They released R1 in January 2025.

Activity is defined as publishing at least 1 code commit in the company's public repository during the specified time period.

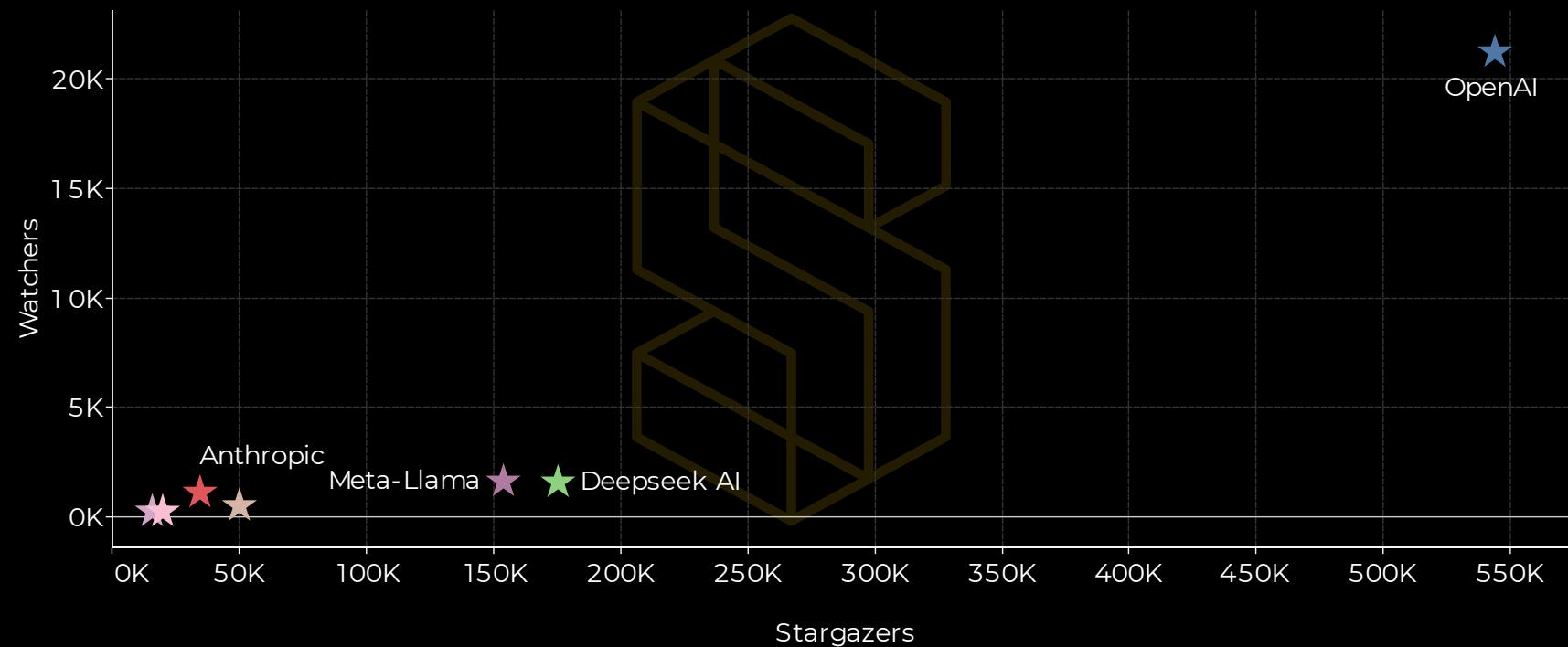
Cumulative Count of Developers Contributing to Company's Public Repositories



105 engineers have contributed to Deepseek's public repositories since inception, compared to 169 for Google Gemini and over 2,500 for OpenAI.

Unique Developer Counts are calculated as the cumulative number of unique developers that have committed at least 1 line of code to a project.

Community Engagement in Public Repositories

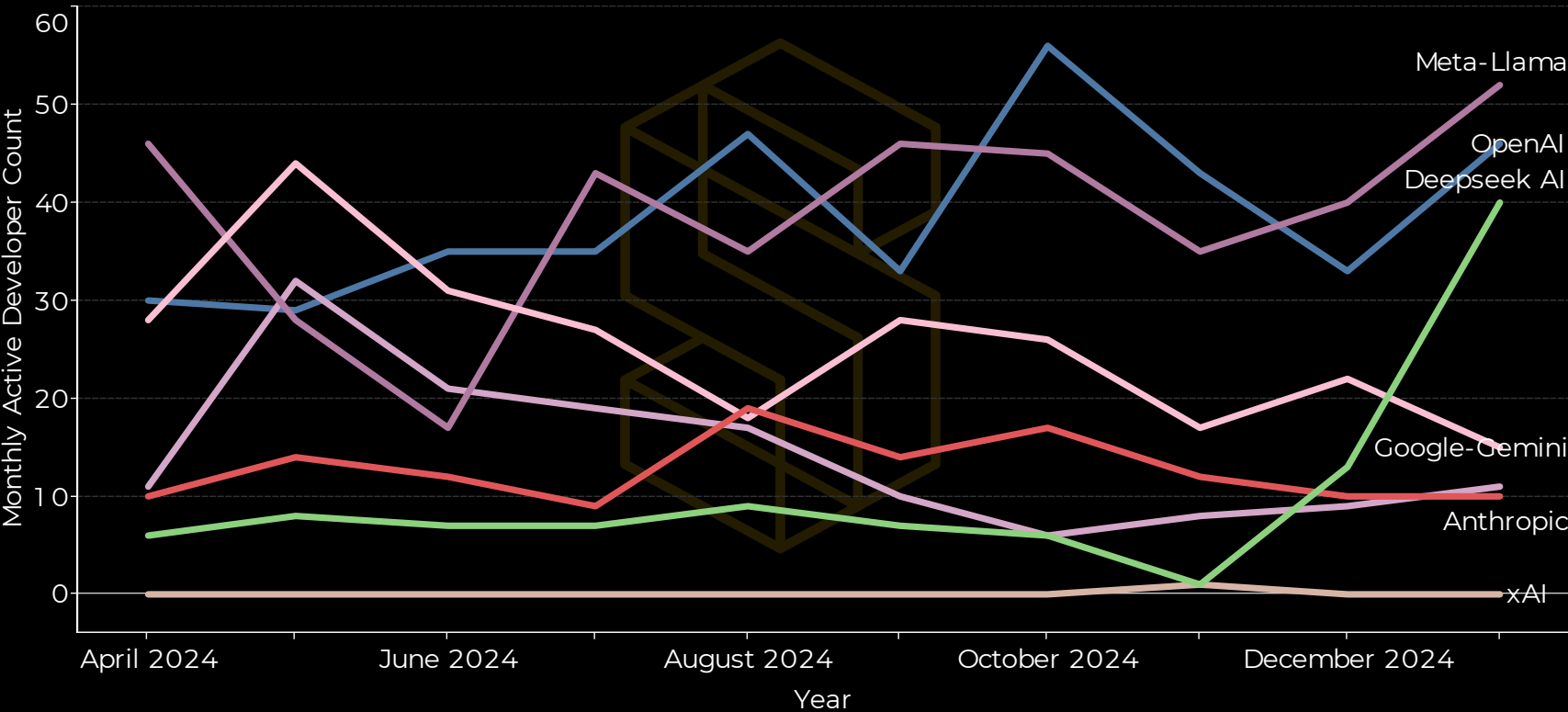


Despite its late start, Deepseek public repositories already have 175K stargazers and 2K watchers — beating out every reported model except OpenAI.

Stargazers are the count of users that have starred or "liked" a repository for the selected project.

Watchers are the count of users that receive a notification when a repository for the selected project is updated.

Monthly Active Core Developers in Public Repositories

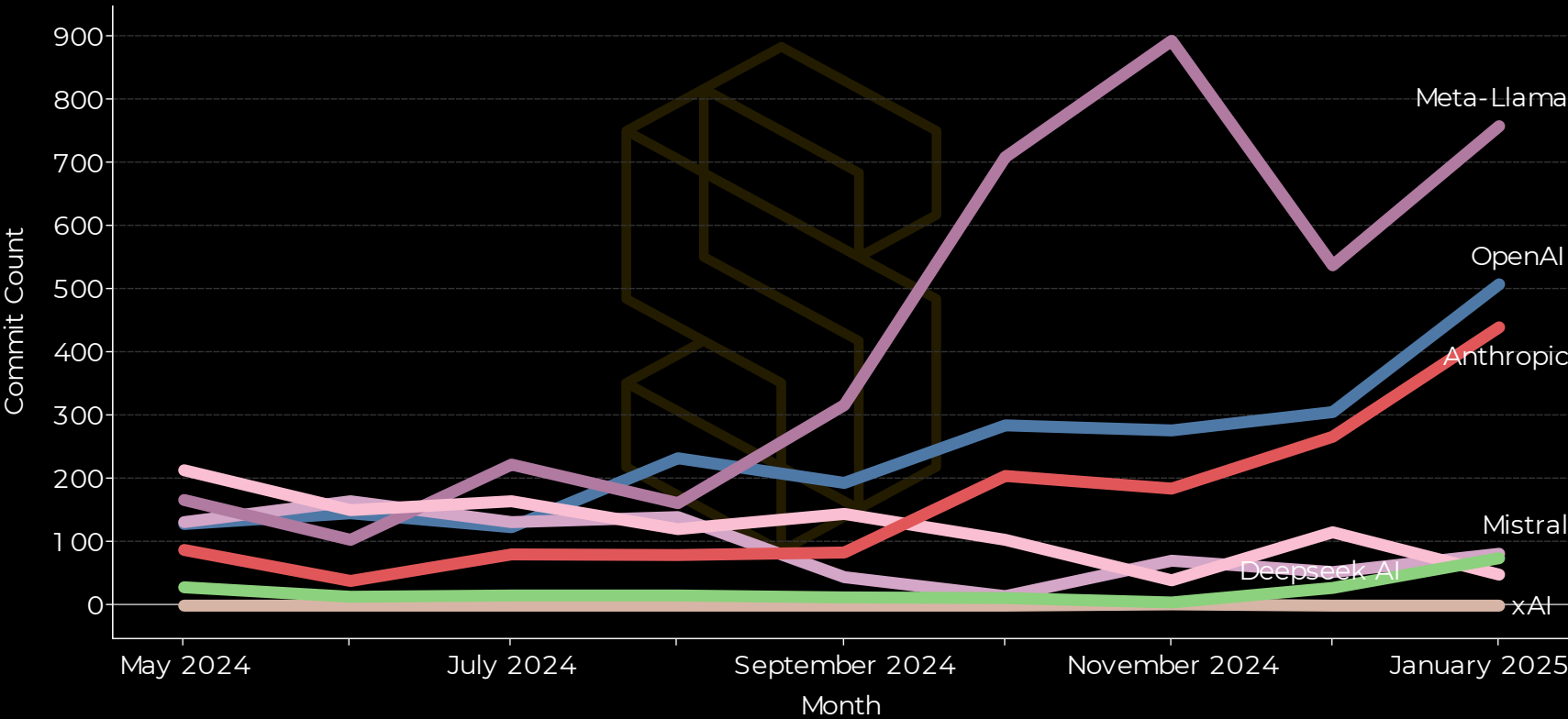


In the previous 9 months, Deepseek had roughly 10 active developers per month committing to public repositories. In January, that number rose 400%, likely to address demand for R1.

Monthly Active Developers are the unique developers that have made at least 1 code commit in the last 30 days.

Deepseek AI Anthropic Google-Gemini Meta-Llama Mistral OpenAI xAI

Monthly Code Commits in Public Repositories

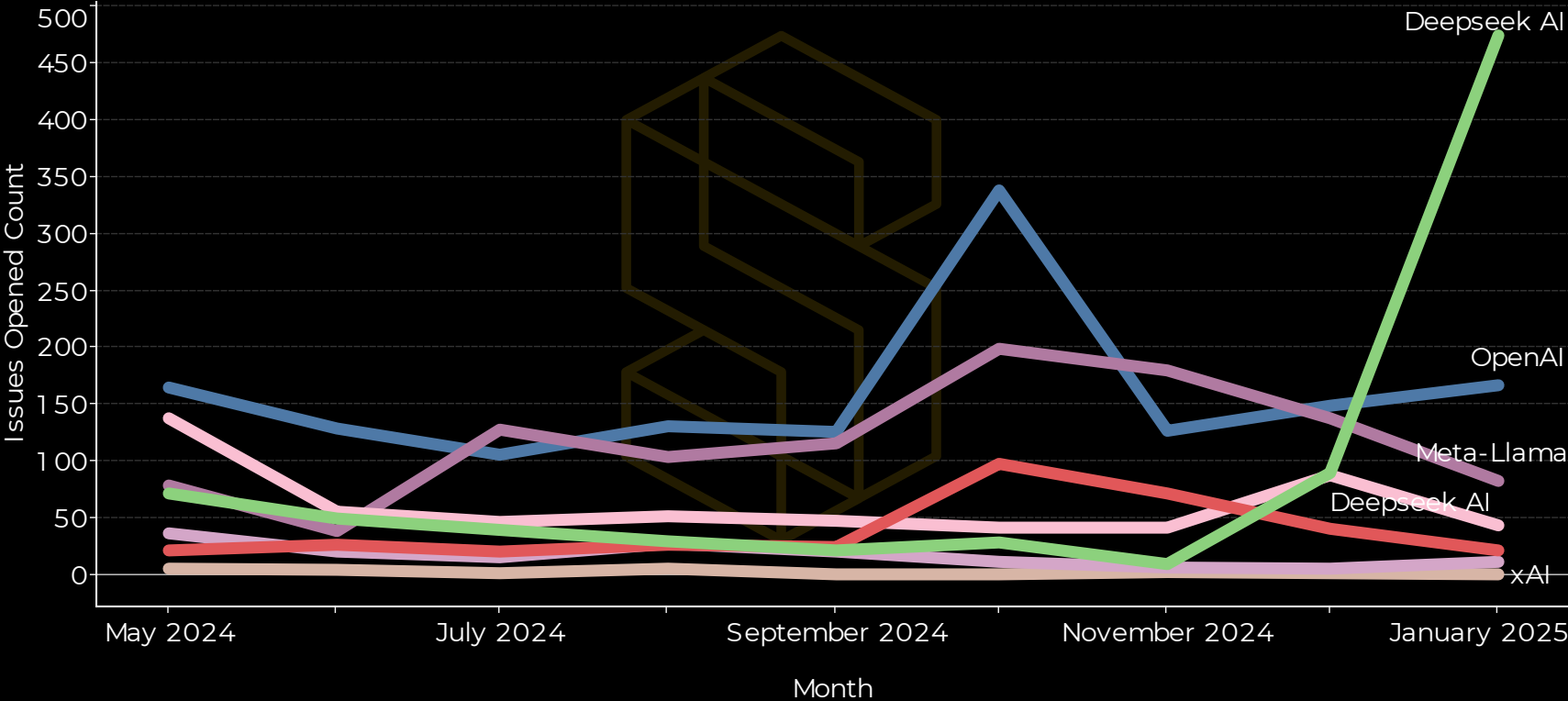


Deepseek developers were committing 10-20 times per month in the months leading up to the release of R1. In January, that number picked up to 75 commits — a fraction of the activity from Meta, Google, Anthropic, and OpenAI developers.

Commits are changes to a code base and its files within a selected repository (aggregated here at one month intervals).

Deepseek AI Google-Gemini Mistral xAI
Anthropic Meta-Llama OpenAI

Monthly Issues Opened in Public Repositories

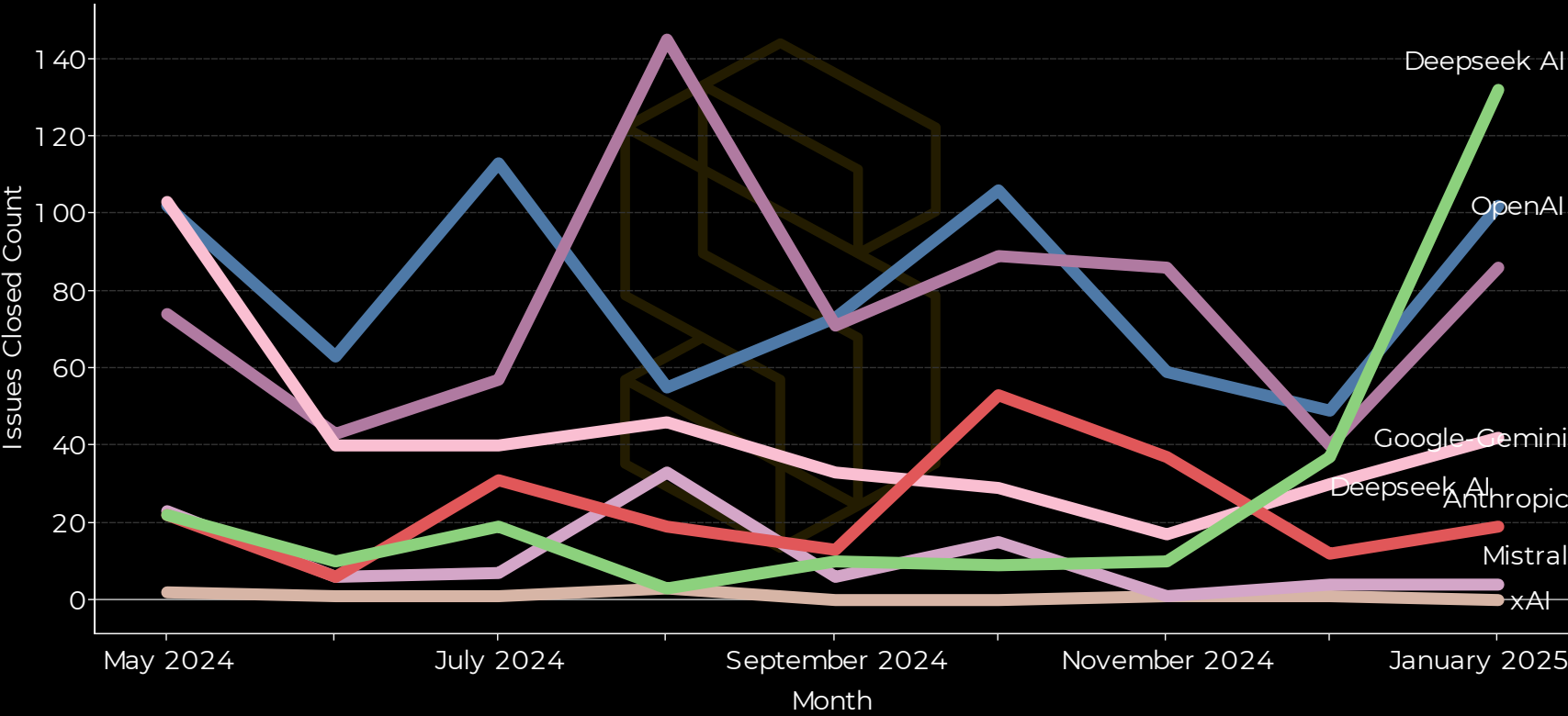


Deepseek opened 474 issues or bugs in January 2025, racing to improve their models while in the public spotlight.

Issue Opened Count is the number of suggested improvements, tasks, or questions related to a code base (aggregated here at one month intervals).

Deepseek AI Anthropic Google-Gemini Meta-Llama Mistral xAI OpenAI

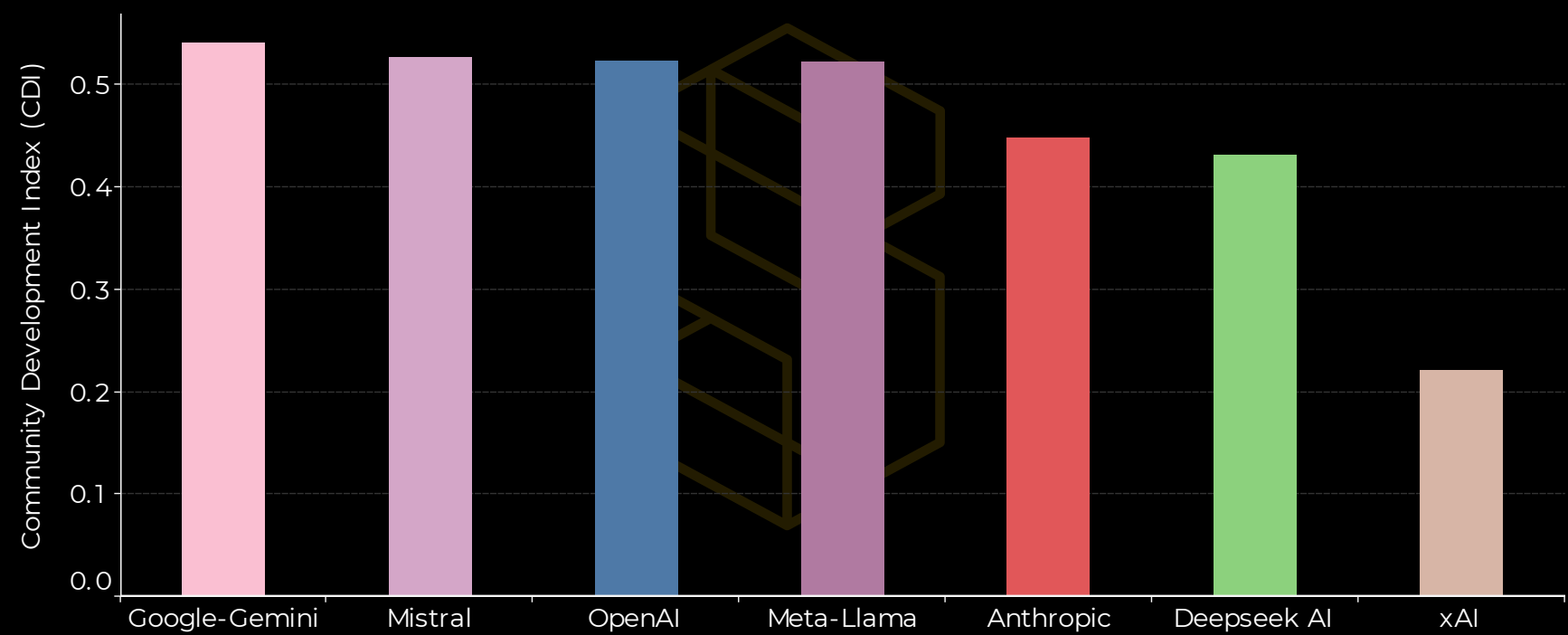
Monthly Issues Closed in Public Repositories



Deepseek closed 132 issues or bugs in January 2025. They are shipping improvements at a rapid pace.

Issues Closed Count is the number of opened issues that were addressed and resolved in a code base (aggregated here at one month intervals).

Community Development Index for Public Repositories (Past 90 Days)



Deepseek has a slightly lower Community Development Index (CDI) than its primary competitors, likely due to the total number of engineers working on the model.

CDI measures how dedicated a developer community is to building an open source project.

Scores can range from 0 to 5.

Learn more about CDI at <https://chaoss.community/?p=4455>

Public Repository Community Development Index (Past 90 Days)

	Meta-Llama	Google-Ge..	Mistral	OpenAI	Anthropic	Deepseek AI	xAI
Community Development Index	0.57	0.56	0.54	0.52	0.50	0.46	0.17
% of Pull Requests Linked to Issues	5%	3%	5%	4%	3%	10%	0%
% of Pull Requests Merged by Non-Author	31%	52%	65%	32%	23%	60%	0%
% of Code Commits Reviewed by Non-Author	30%	55%	52%	11%	8%	20%	0%
% of Code Commits Linked to a Pull Request	100%	100%	100%	100%	100%	100%	50%
% of Organization Repositories Modified	100%	100%	100%	100%	100%	100%	50%
Unique Code Contributors	37	18	10	29	24	6	2
Average Weekly Code Commits	23	5	3	7	7	1	0
Lines of Code Modified	16,456	3,689	2,547	11,450	5,062	246	3

Deepseek's low CDI score is a reflection of the size of the team and the amount of code they are modifying, not their commitment to following best practices in reviewing each other's code.

CDI measures how dedicated a developer community is to building an open source project.

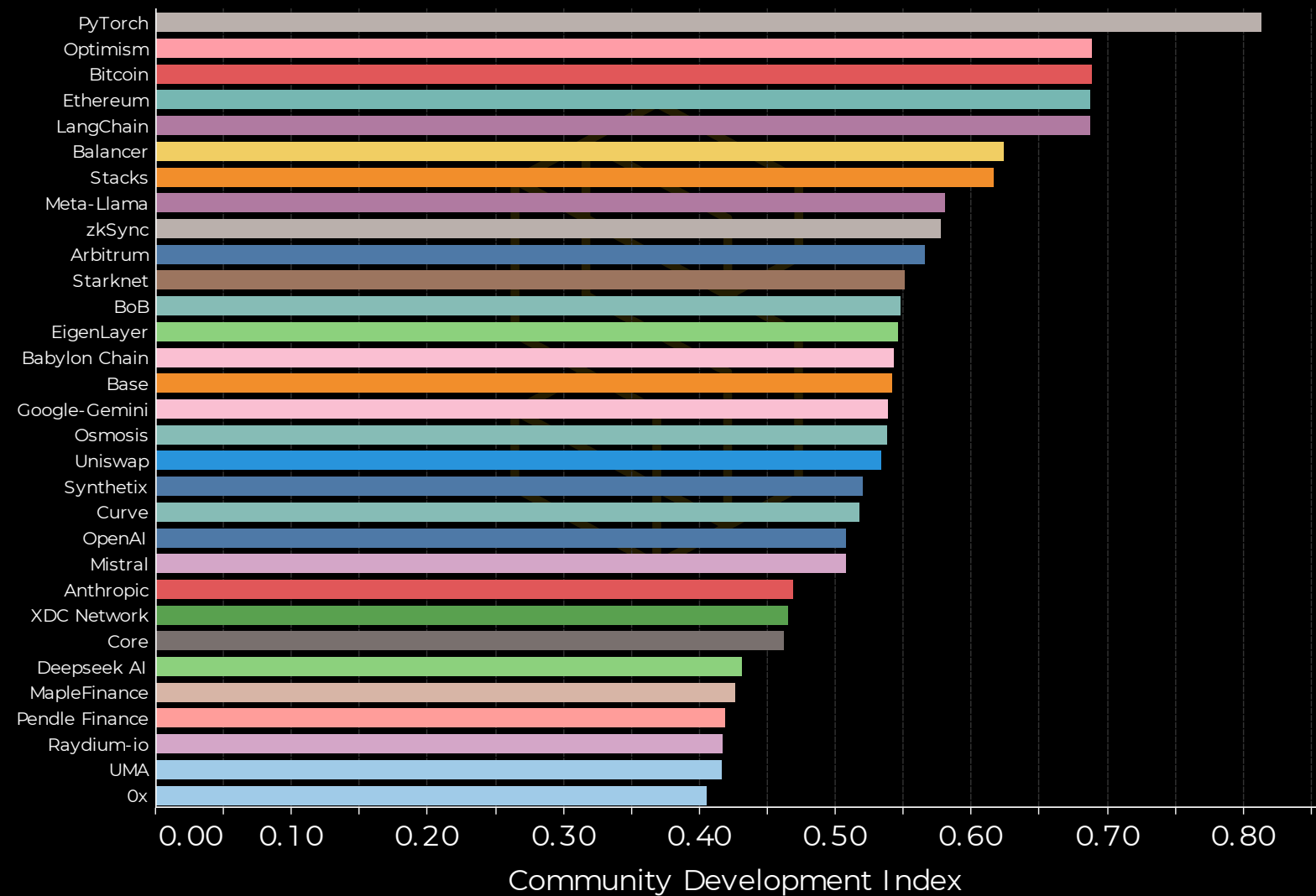
Scores can range from 0 to 5.

The metrics listed above are used to calculate the CDI aggregate score.

Learn more about CDI at <https://chaoss.community/?p=4455>



Public Repository Community Development Index (Past 90 Days)



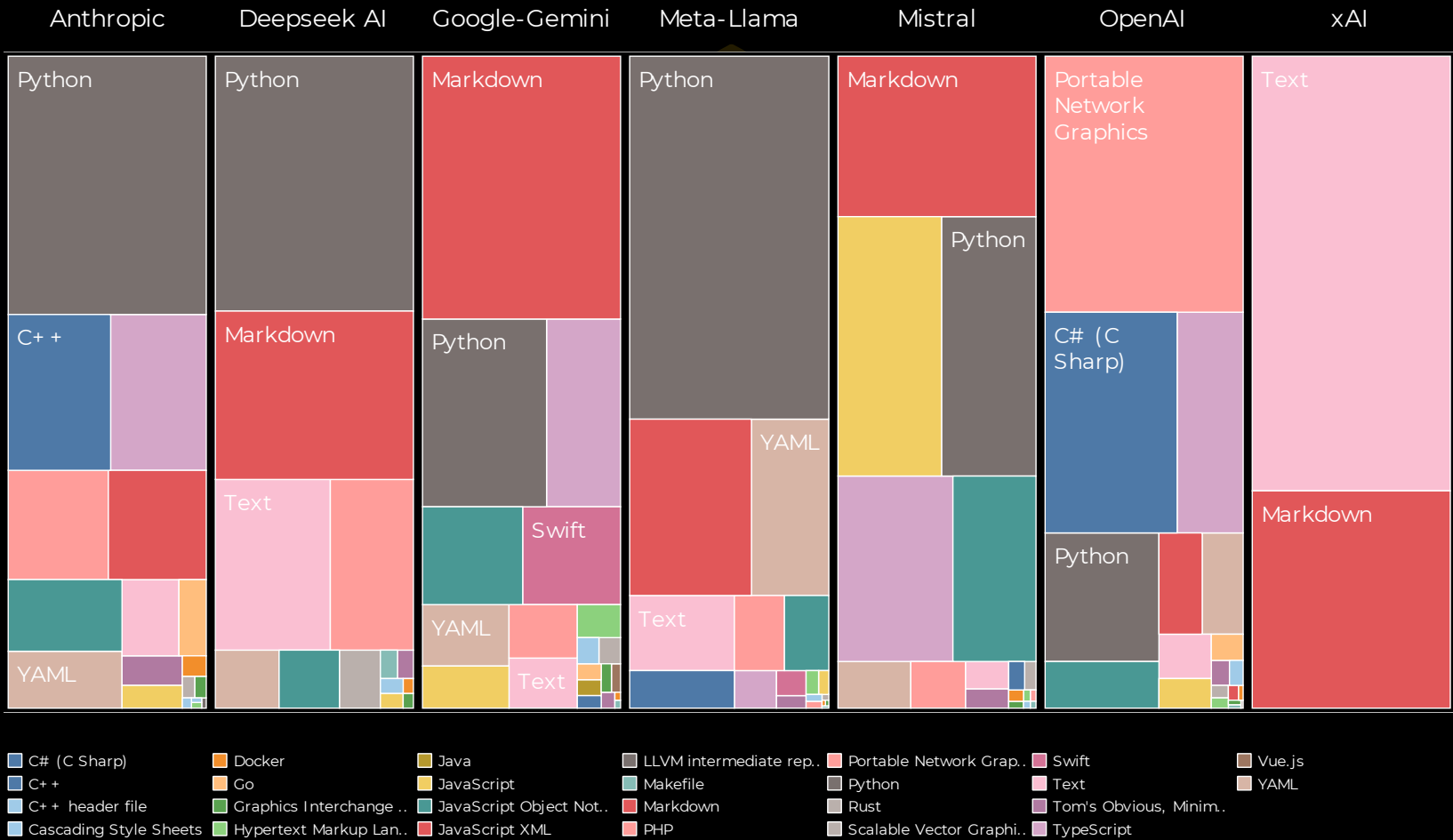
Zooming out, Deepseek has a low CDI relative to most other open source blockchain and AI companies. Among AI companies included in the report, PyTorch and Langchain have the strongest developer communities.

CDI measures how dedicated a developer community is to building an open source project.

Scores can range from 0 to 5.

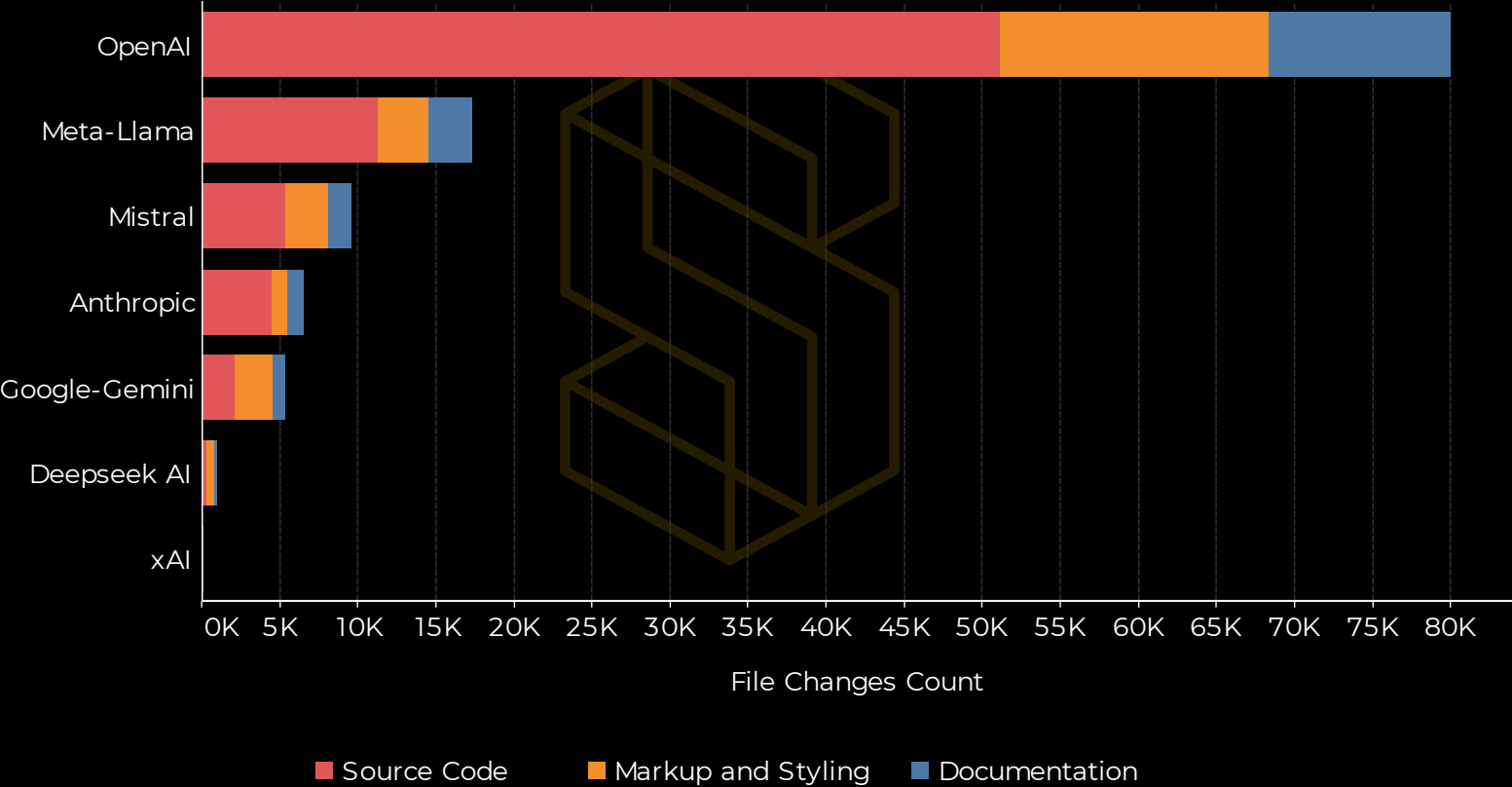
Learn more about CDI at <https://chaoss.community/?p=4455>

Public Repository File Types



Most of the code Deepseek engineers are updating is contained in Python and Markdown files.

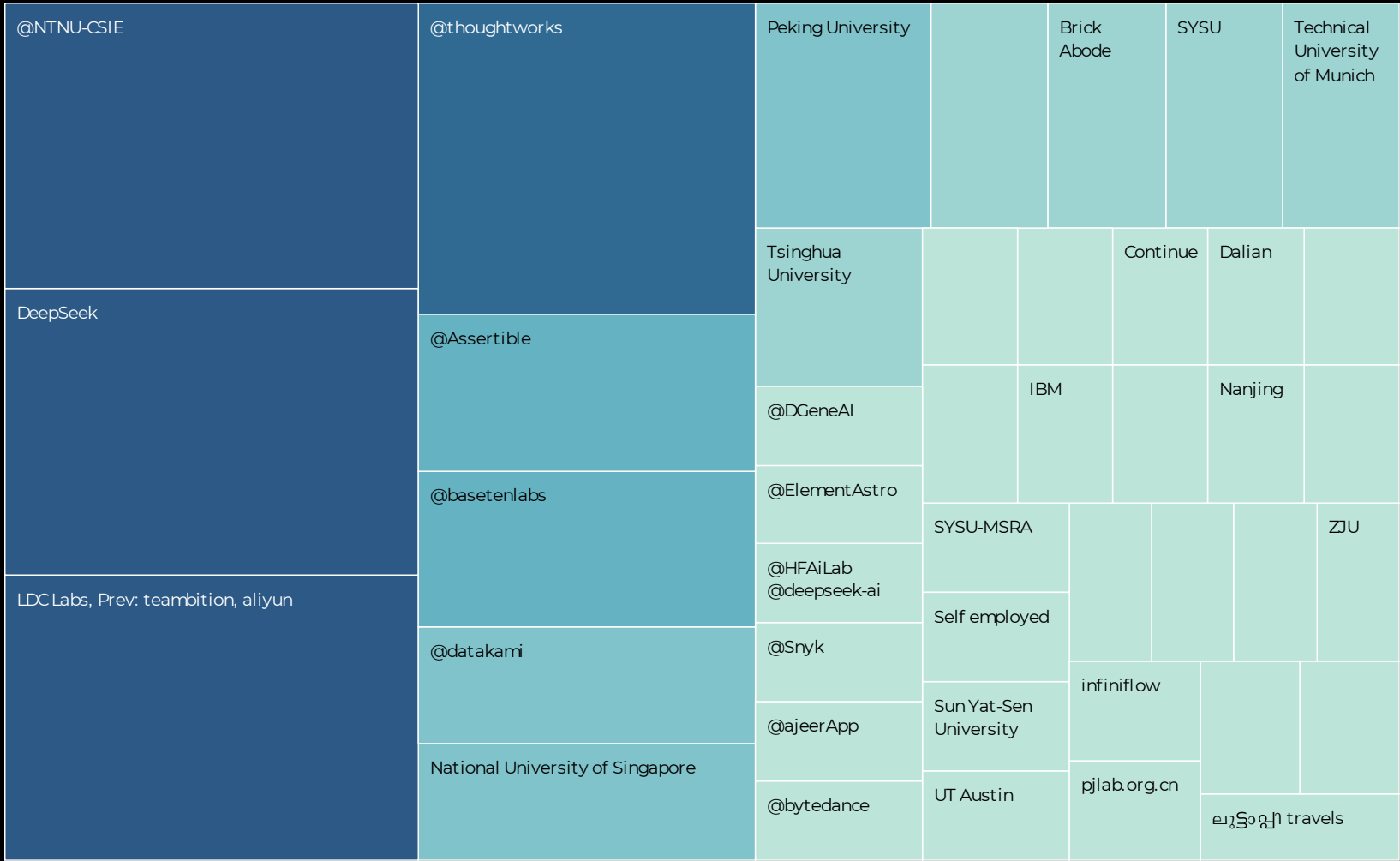
Public Repository File Changes (Past 12 Months)



Relative to its competitors, Deepseek has made minimal file changes in the last 12 months.

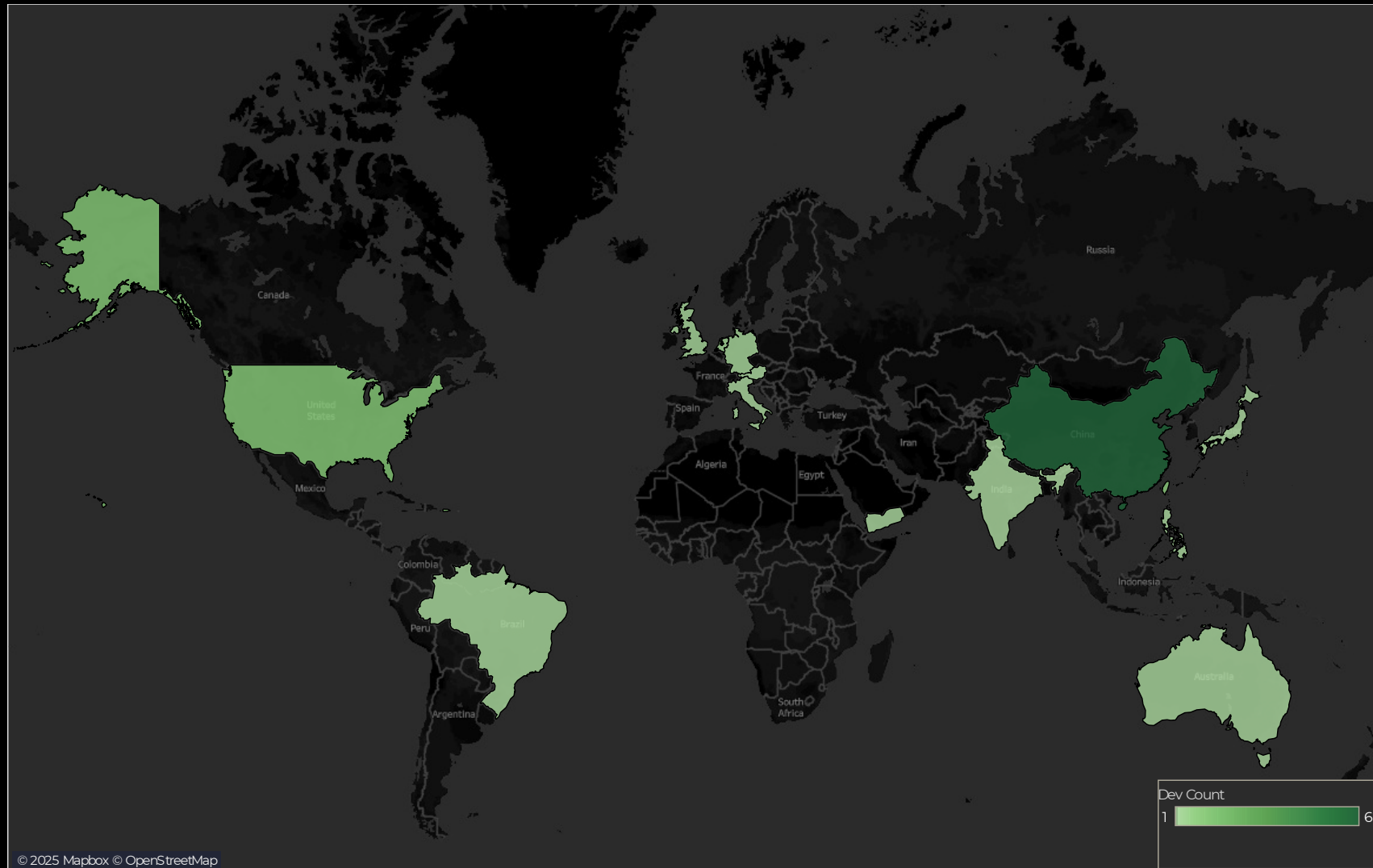
Code commits can impact multiple files within a repository, leading to additions, deletions, or modifications. Sherlock Data has classified these changes into three categories (Documentation, Markup & Styling, and Source Code) based on the properties of each file.

Deepseek Engineer Self Reported Company



While some Deepseek engineers list Deepseek as their employer in their developer profiles, many list the National Taiwan Normal University and other institutions.

Deepseek Engineer Self Reported Location



Deepseek contributors have listed a number of countries beyond China in their developer profiles. Locations named include Taiwan, Japan, and the United States.



Sherlock[®]
Data

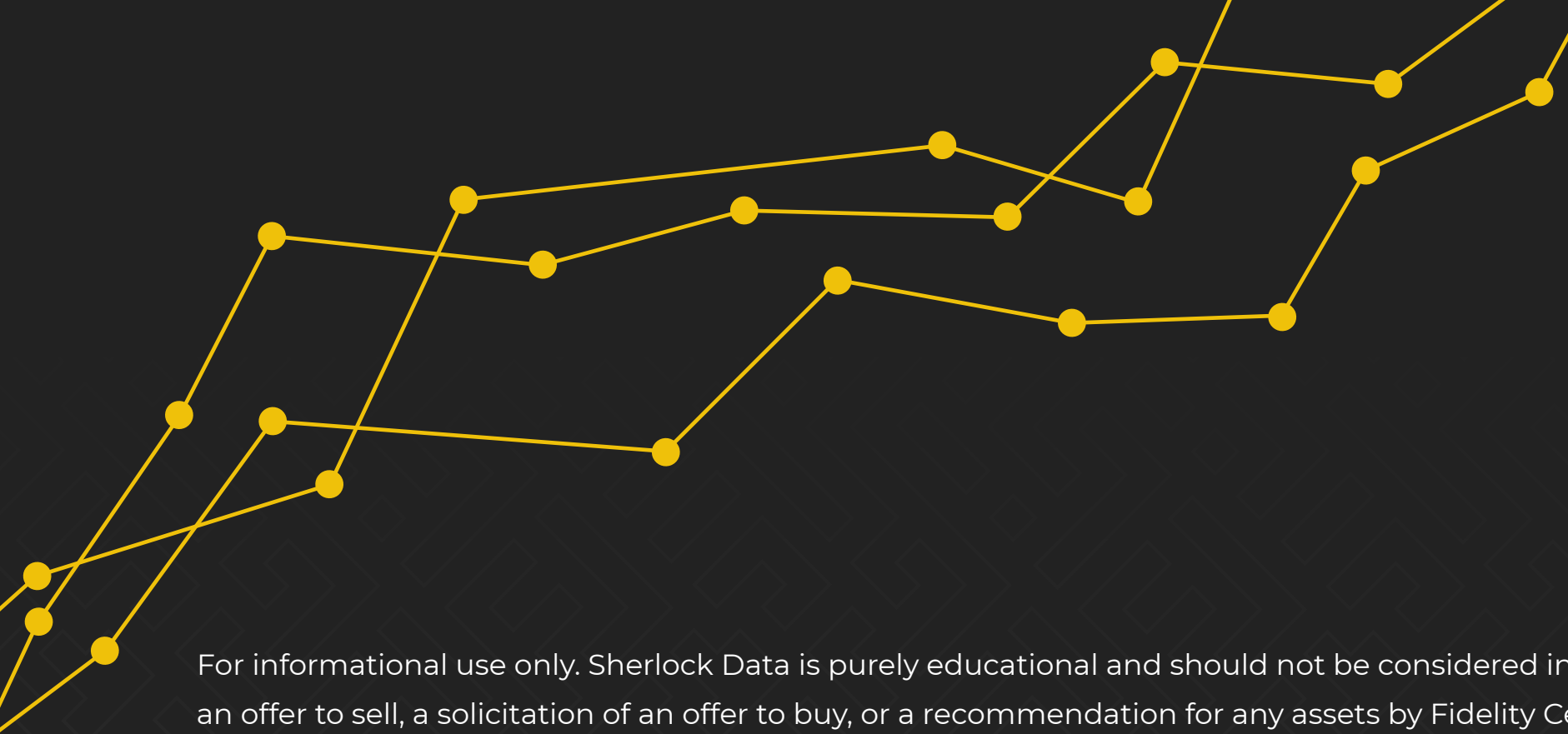
Empowering Investors with Actionable Insights

To learn more, including how to sample or buy this data:

Visit our listing on Snowflake 

Or send us an email at:
sherlockdata@sherlockanalytics.com

Powered by the Fidelity Center for Applied Technology (FCAT[®])



For informational use only. Sherlock Data is purely educational and should not be considered investment advice or construed as an offer to sell, a solicitation of an offer to buy, or a recommendation for any assets by Fidelity Center for Applied Technology LLC (FCAT®). FCAT is not providing tax, investment, legal or advice of any kind. Digital assets are speculative and highly volatile, can become illiquid at any time, and are for investors with a high-risk tolerance. Investors in digital assets could lose the entire value of their investment. Access to and use of this report is subject to the [Sherlock Terms of Use](#), including that you cannot copy, reproduce, transmit, publicly display, publicly perform, publish, adapt, edit, compile, or create derivative works from the report.

FCAT is not affiliated with any third parties referenced herein including third-party providers. These organizations are not Fidelity employees or agents. Fidelity does not guarantee the accuracy, completeness or timeliness of any content, including content provided by third-party data providers, and it should not be solely relied upon for making any investment decision.