Reddit’s upcoming API changes will make AI companies pony up

Reddit announced new API changes today that will eventually pinch its content pipeline from being used to train artificial intelligence tools, including the models that power OpenAI’s ChatGPT, Google’s Bard, and Microsoft’s Bing AI. AI chatbots’ abilities to provide powerful answers have data resources like Reddit to thank — and now Reddit is planning to put that robot food behind a paywall.

Social media resources, including Reddit, are some of the sources used to train large language models (LLM) that can provide cogent responses to human prompts. Some of this data can be scraped in an unstructured manner, but Reddit’s API has helped companies make it easy to directly find and package useful data.

Reddit’s API, which has been available since 2008, has previously been fairly open for developers to do almost anything. That includes building tools that help moderate subreddits, creating Reddit browsing clients, and making the site easier to search. Reddit plans to keep the API free for some use cases, like for those who build moderation tools or use Reddit in educational and research environments.

Commercial usage will require a separate agreement with an undisclosed cost

Reddit’s new terms apply to developers who use its APIs in ways that require “broader usage rights” and won’t grant automatic licenses for anyone needing to modify user content, as published in its new Data API terms. This means commercial usage, like training LLMs, will not be granted a developer license and will instead require parties “to enter into a

separate agreement with Reddit.” Reddit has yet to detail how much it plans to charge companies using its data commercially.

Reddit hasn’t gone into finer detail on how API changes will directly affect third-party Reddit clients like Apollo, Rif, and Relay. It does mention in the Data API terms that it can enforce limits on how many API requests are made — which could be pretty high for clients since they need to use OAuth tokens for Reddit user authentication. Apollo’s sole developer, Christian Selig, asked Reddit how “enforcing rate limits” will affect apps like his. A Reddit admin replied vaguely, saying it depends on the volume of API usage and whether it’s “compliant with our terms.”

These API changes come as Reddit plans an initial public offering for later this year. Much of the company’s monetization comes in the form of advertising (which has its own API) and digital goods. But as more AI platforms emerge, Reddit wants to build on the value of its user-generated content. “The Reddit corpus of data is really valuable,” Reddit CEO Steve Huffman states in an interview with *The New York Times*. “We don’t need to give all of that value to some of the largest companies in the world for free.” The changes also follow a far broader lockdown of Twitter’s API under Elon Musk’s ownership — one that could hit commercial and noncommercial users alike.

The new Reddit terms will go into effect “following a 60-day notice period” after developers and third parties receive an official email notification. Reddit will also be releasing new in-house moderator tools that work with its official iOS and Android apps.

https://www.theverge.com/2023/4/18/23688463/reddit-developer-api-terms-change-monetization-ai