Perplexity Says Cloudflare Is Blocking Legitimate AI Assistants via @sejournal, @martinibuster

Perplexity published a response to Cloudflare's claims that it disrespects robots.txt and engages in stealth crawling. Perplexity argues that Cloudflare is mischaracterizing AI Assistants as web crawlers, saying that they should not be subject to the same restrictions since they are user-initiated assistants.

Perplexity AI Assistants Fetch On Demand

According to Perplexity, its system does not store or index content ahead of time. Instead, it fetches webpages only in response to specific user questions. For example, when a user asks for recent restaurant reviews, the assistant retrieves and summarizes relevant content on demand. This, the company says, contrasts with how traditional crawlers operate, systematically indexing vast portions of the web without regard to immediate user intent.

Perplexity compared this on-demand fetching to Google's user-triggered fetches. Although that is not an apples-to-apples comparison because Google's user-triggered fetches are in the service of reading text aloud or site verification, it's still an example of user-triggered fetching that bypasses robots.txt restrictions.

In the same way, Perplexity argues that its AI operates as an extension of a user's request, not as an autonomous bot crawling indiscriminately. The company states that it does not retain or use the fetched content for training its models.

Criticizes Cloudflare's Infrastructure

Perplexity also criticized Cloudflare's infrastructure for failing to distinguish between malicious scraping and legitimate, user-initiated traffic, suggesting that Cloudflare's approach to bot management risks overblocking services that are acting responsibly. Perplexity argues that a platform's inability to differentiate between helpful AI assistants and harmful bots causes misclassification of legitimate web traffic.

Perplexity makes a strong case for the claim that Cloudflare is blocking legitimate bot traffic and says that Cloudflare's decision to block its traffic was based on a misunderstanding of how its technology works.

Read Perplexity's response:

Agents or Bots? Making Sense of AI on the Open Web

https://www.searchenginejournal.com/perplexity-says-cloudflare-is-blocking-legitimate-ai-assistants/552927/