



AI Is Hitting Operational Limits as Companies Rush to Scale, Datadog Report Finds

April 21, 2026 at 9:15 AM EDT

Nearly 1 in 20 AI requests fail in production as capacity limits become the primary bottleneck to scaling AI reliably

NEW YORK, April 21, 2026 (GLOBE NEWSWIRE) -- As AI adoption accelerates, operational complexity – not model intelligence – is becoming the primary barrier to reliable AI at scale, according to new data from [Datadog](#), Inc. (NASDAQ: DDOG), the AI-powered observability and security platform.

Datadog's [State of AI Engineering 2026](#) report, based on real-world data from thousands of organizations running AI in production, highlights a compounding complexity challenge as AI systems scale. Nearly seven in ten companies (69%) now use three or more models alongside increasingly complex agent workflows. Around 5% of AI model requests fail in production, with nearly 60% of those failures caused by capacity limits – leading to slowdowns, errors, and broken experiences in AI-powered applications.

Additional key findings:

- Multi-model is now the norm: OpenAI remains the most widely used provider at 63% share, alongside rising adoption of Google Gemini and Anthropic Claude which grew by 20 and 23 percentage points, respectively.
- Agent framework adoption doubled year-over-year, accelerating development but also introducing more moving parts into production systems.
- The amount of data sent to AI models per request is also rising: the average number of tokens more than doubled for 'median use' teams (50th percentile of usage volume) and quadrupled for heavy users (90th percentile).

"AI is starting to look a lot like the early days of cloud," said Yanbing Li, Chief Product Officer at Datadog. "The cloud made systems programmable but much more complex to manage. AI is now doing the same thing to the application layer. The companies that win won't just build better models - they'll build operational control around them. In this new era, AI observability becomes as essential as cloud observability was a decade ago."

Speed Requires Control

Competitive pressure is accelerating AI deployment across startups and large enterprises alike. But as systems scale, speed without control creates risk. Failures are increasingly driven by system design, including fragmented workflows, excessive retries, and inefficient routing.

"The next wave of agent failures won't be about what agents can't do but what teams can't observe," said Guillermo Rauch, CEO at [Vercel](#), the company behind Next.js and a leading platform for building AI-powered web applications. "We built agentic infrastructure at Vercel because agents need the same production feedback loops as great software. Unlike traditional software, agents have control flow driven by the LLM itself, making observability not just useful, but essential."

"Innovation alone isn't enough," added Li. "To scale AI with confidence, organizations need real-time visibility across the entire stack – from GPU utilization to model behavior to agent workflows. Visibility and operational control are what allow teams to move fast without sacrificing reliability or governance. At scale, how you operate AI may matter more than the models you choose."

Read the full report - [The State of AI Engineering 2026](#) - and learn how Datadog is investing in AI observability to help teams operate and scale AI systems in production [here](#).

Report Methodology

Datadog analyzed anonymized usage data from thousands of customers using LLMs in production environments, with global coverage across industries and geographies.

About Datadog

Datadog is the AI-powered observability and security platform. Our SaaS platform integrates and automates infrastructure monitoring, application performance monitoring, log management, user experience monitoring, cloud security and many other capabilities to provide unified, real-time observability and security for our customers' entire technology stack. Datadog is used by organizations of all sizes and across a wide range of industries to enable digital transformation and cloud migration, drive collaboration among development, operations, security and business teams, accelerate time to market for applications, reduce time to problem resolution, secure applications and infrastructure, understand user behavior and track key business metrics.

Forward-Looking Statements

This press release may include certain "forward-looking statements" within the meaning of Section 27A of the Securities Act of 1933, as amended, or the Securities Act, and Section 21E of the Securities Exchange Act of 1934, as amended including statements on the benefits of new products and features. These forward-looking statements reflect our current views about our plans, intentions, expectations, strategies and prospects, which are based on the information currently available to us and on assumptions we have made. Actual results may differ materially from those described in the forward-looking statements and are subject to a variety of assumptions, uncertainties, risks and factors that are beyond our control, including those

risks detailed under the caption “Risk Factors” and elsewhere in our Securities and Exchange Commission filings and reports, including the Quarterly Report on Form 10-Q filed with the Securities and Exchange Commission on February 18, 2026, as well as future filings and reports by us. Except as required by law, we undertake no duty or obligation to update any forward-looking statements contained in this release as a result of new information, future events, changes in expectations or otherwise.

Contact:

press@datadoghq.com



Source: Datadog, Inc.