BRIGHTEDGE

WEEKLY AI & SEARCH INSIGHTS

Meta AI crawler traffic surged 30-fold ahead of the Muse launch.

On one site we measure, requests identifying themselves as Meta went from about 30K to about 1.0M a day between September 6 and September 7. Meta's bots then ranked second of 10 tracked companies on 49 of 52 days. A change this size may not be visible in conventional web analytics at all.

These are crawling bots, not shoppers. The increase was measured across two combined Meta crawler strings, one for link previews and one for training and indexing.

Source: BrightEdge AI Agent Insights, one ecommerce client, August 2 to October 29, 2026. One site's view, not a market measurement.

Thirty thousand requests a day, then a million.

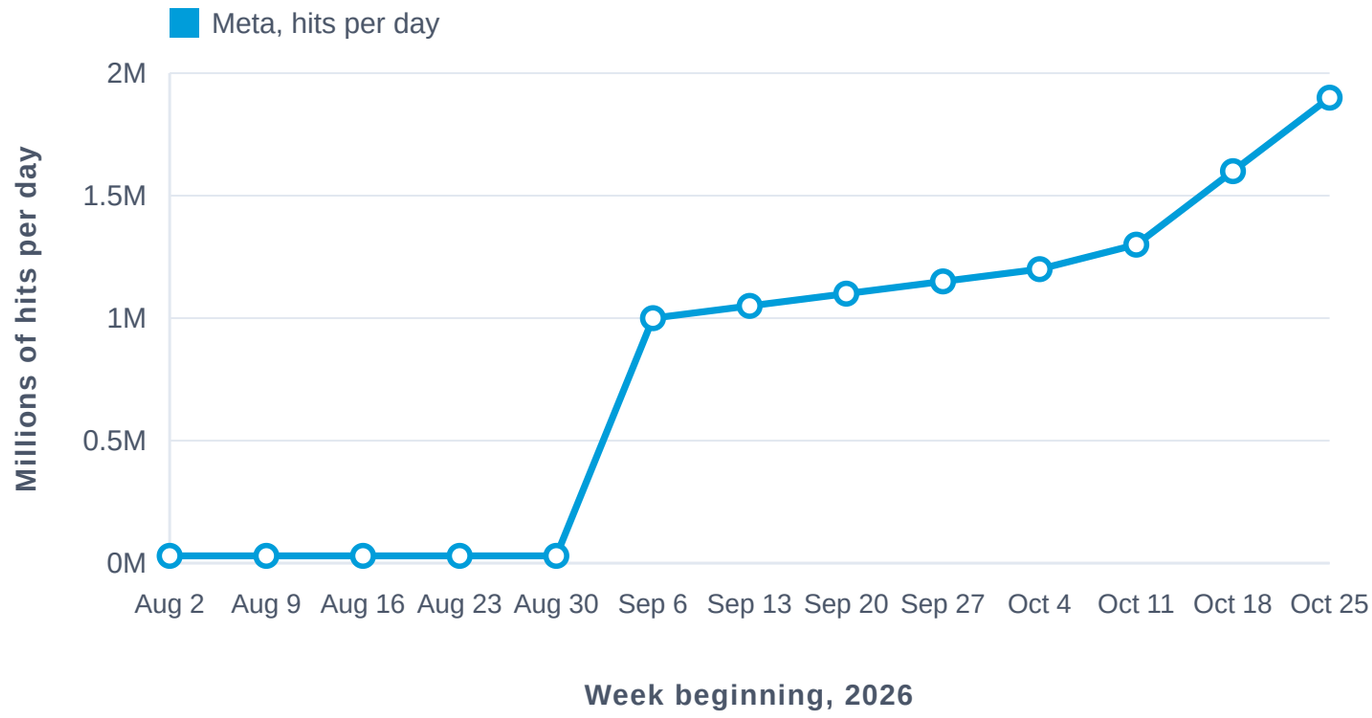


Figure 1. Requests identifying themselves as Meta, one BrightEdge ecommerce client. About 30K hits on a typical day before September 7, about 1.0M after. Selected weekly observations. Source: BrightEdge AI Agent Insights, one ecommerce client, August 2 to October 29, 2026. Ten tracked companies. Results reflect one site.

WHY YOUR DASHBOARD MAY NOT SHOW IT

Much of this activity may not appear in conventional web analytics, because a crawler fetches the HTML and does not run the JavaScript a tracking tag needs. Server logs are where to look.

Second place on 49 days. First on three.

49 of 52

days Meta's bots ranked second of 10 tracked companies, behind Google. It ranked first on the other 3

1.17M

mean hits per day over those 52 days, approaching 2M by late October

The ten companies tracked on this site are Meta, Google, Apple, Amazon, OpenAI, Perplexity, Anthropic, Microsoft, ByteDance and Common Crawl.

Figure 2. Rank by hits per day, September 8 to October 29, 2026. Source: BrightEdge AI Agent Insights, one ecommerce client, August 2 to October 29, 2026. Ten tracked companies. Results reflect one site.

One BrightEdge ecommerce client. This is one site's view of agent traffic, not a market measurement and not a read on BrightEdge's customer base.

One company name can cover three different crawlers

TYPE	WHAT IT DOES	EXAMPLE	SHARE
Training	Builds and refreshes the knowledge base behind a model	GPTBot	**51% of training crawls**
Search	Retrieves content to power AI answers and citations	OAI-SearchBot	**49% of AI search crawls**
User	Acts for one person during an active session	ChatGPT User	**96% of AI user traffic**

Figure 3. The three agent types and the largest bot in each. Shares from BrightEdge research covering November 2025 to February 2026, measured across tracked sites rather than the single client in Figures 1 and 2.

BrightEdge found 19% of sites carry specific directives for ChatGPT-related bots. The other 81% of companies treat AI agents like traditional bots. Where policies exist, 77% focus on blocking training agents, 21% address search and 38% address user-facing agents.

> Training, search and user are the three AI agent types. Companies run bots outside them too, so a policy should follow each string's actual function rather than force every request into three categories.

Digital marketing

Agents judge a page by what they can reach and read quickly. Your training, search and user bot policy decides whether they reach it.

PR and communications

Training crawlers shape how models describe you, so newsrooms and executive bios are source material. Link preview crawlers decide how earned coverage looks when shared. Neither shows in media monitoring.

Four checks, starting with the one analytics will not do

1 Look at your logs, not your analytics

Most analytics drops automated traffic entirely. Start with which bots request your site, which pages they ask for, and where they hit a broken page, a blocked resource or a slow response.

2 Separate the string from the company

One company can run several bots that do different jobs. Meta runs a link-preview fetcher, a training and indexing crawler and a user-triggered fetcher under three different names. A policy written against the company name treats all three the same.

3 Set a policy for each of the three agent types

Training, search and user agents do different jobs and deserve different answers. Most published policies address training and leave the other two to an inherited default.

4 Re-check the mix on a schedule

On this site the mix of automated requests shifted sharply between a Sunday and a Monday. A policy reviewed annually will describe a web that no longer exists.

Meta's bots went from a rounding error to second place on this site inside a day, in categories most policies claim to cover. A change of that size can pass unseen in conventional web analytics, which is why the visibility problem comes before the policy problem.

THE FULL RESEARCH

See the agents first. Then decide about them.

Automated requests cost real serving capacity and they read whatever the page gives them. A site that cannot see which bots arrive cannot size that cost, and cannot tell whether the content a bot reads is the content it should. Bloomfire doubled agent traffic to optimized pages and saw AI referral traffic rise 30% in July 2026 against the prior 90 days. Arm served agents a payload about 90% smaller and held a 2x lift.

RUN THIS ON YOUR OWN SITE

1. **See which agents reach your site**

   AI Agent Insights reports the bots requesting your pages and what they engage with.

2. **Decide by agent type, not by one blanket rule**

   Training, search and user agents are separate decisions. An inherited bot rule makes all three for you.

3. **Serve an agent a page it can read quickly**

   Agent Edge detects verified agent traffic at the CDN and serves those agents a lighter page, while people and Googlebot get the original.

Source: BrightEdge AI Agent Insights, one ecommerce client, August 2 to October 29, 2026. Ten tracked companies. Results reflect one site. Requests counted, not shoppers, recommendations or purchases, and grouped by the identity each one declares. One site's view, not a market measurement.

Full research linked in the post.

Request a demo