

Machines Have Taken the Web— And Most of Them Don't Care About Search



The internet just quietly passed a milestone the marketing world had been bracing for: software, not people, now generates most of the traffic moving across the web. Fresh figures from Cloudflare's Radar platform, captured over the seven days ending June 5, 2026, put automated requests at 57.4% of all traffic hitting HTML content, leaving human visitors at 42.6%. The 50% line had been creeping closer for months, but the speed of the crossing caught even seasoned analysts off guard.

A majority that isn't built for search

The headline number is arresting on its own, but the composition underneath it tells the real story. Within the bot share, the systems crawling pages to feed AI training pipelines now make up roughly 50.6% of total traffic. Crawlers tied to actual search indexing account for a mere 10.7%. That gap is the crux of a growing standoff between the people who publish content and the companies building large AI models on top of it.



The distinction is not academic. A search crawler that visits a page at least opens the door, in theory, to a human clicking through from a results page later. A training crawler arrives, harvests the material to sharpen a model, and leaves nothing behind—no referral, no credit, no payment. For a publisher footing the hosting bill, the second kind of visit is pure cost with no return.

For more content like and follow me:



@bertblevins

INCGPT.COM

From a sliver to a flood in six months

Cloudflare is unusually well placed to measure this. Its network sits in front of millions of sites worldwide, which makes its sampling one of the closest things the industry has to a representative view of global traffic. Late in 2025, the company's own figures had AI bots at just 4.2% of HTML requests, with humans around 43.5% and other automated systems near 47.9%. The leap from that baseline to today's numbers in roughly half a year signals a genuine structural shift, not a statistical blip.



The pattern is consistent with earlier readings. By September 2025, training-oriented crawling had climbed to nearly 80% of all AI bot activity, up from about 72% a year before, while search crawling slid to 18% and user-triggered fetches sat at a token 2%. The current Cloudflare slice measures things differently—against all traffic rather than within AI activity alone—but it points in exactly the same direction.

The heavy hitters

Individual operators illustrate the surge. OpenAI's GPTBot expanded its slice of AI crawling from 4.7% in mid-2024 to 11.7% a year later. An analysis spanning more than seven billion server log entries found OpenAI's crawlers roughly tripled their activity after GPT-5 shipped in August 2025—even though usage signals from the same window hinted that the number of people directly using ChatGPT might actually be slipping. ByteDance's Bytespider has likewise become one of the most demanding bots by sheer request volume.



Commentators have not been gentle about the economics. One former enterprise SEO lead framed the situation as a tax levied on site owners: the AI firms monetize the data, while the publisher keeps paying to host the very content being scraped. SparkToro's co-founder rendered it as a darkly comic exchange—site owners essentially agreeing to bankroll the republishing of their own work. A separate analyst floated a feedback-loop theory: more people leaning on chatbots drives more queries, which in turn dispatches more crawlers.

Breaking the plumbing

The training-heavy crawl pattern does more than rack up requests; it actively undermines the caching systems sites depend on. Researchers from Cloudflare and ETH Zurich identified three behaviors that make AI crawlers especially disruptive. They fire off large volumes of parallel requests. Instead of clustering on popular pages—the workload caches are tuned for—they sweep methodically through obscure, rarely visited content. And because each crawler instance launches without shared session state, the same content request can register repeatedly as a fresh visitor.



That combination defeats the “least recently used” eviction logic most caches rely on, where the stalest content gets dropped first when storage tightens. When crawlers systematically rake through low-traffic pages, they shove genuinely popular material out of the cache, pushing load and cost back onto origin servers. The net effect: publishers pay for hosting, pay again for the extra strain training crawls create, and receive nothing in exchange.



Cloudflare's answer has been a string of tools. It rolled out crawler-blocking controls in mid-2024, then introduced a pay-per-crawl mechanism in private beta in 2025 so creators could charge AI companies for access rather than simply shutting the door. The company's chief executive has gone so far as to sketch out scenarios for the web's future that include the outright extinction of content creators as one possibility.



Does an AI mention actually pay?

Running parallel to the traffic debate is a sharper argument about value. Vendors have spent two years insisting that showing up in an AI-generated answer—click or no click—carries worth because it builds brand awareness. A prominent analytics specialist pushed back hard, arguing that AI brand mentions deliver no measurable commercial value, regardless of what vendor-funded studies claim.

01

The reasoning is empirical rather than theoretical. If AI visibility truly offset the organic traffic that features like Google's AI Overviews have eaten, then leads, sales, and form fills should hold steady even as raw visits decline. That is not what most operators are seeing. Sites suffering heavy organic losses are, by and large, watching conversions fall in step. The indirect halo of AI exposure simply isn't large enough to cover the volume gap for the typical business.

02

This cuts against a thick layer of vendor research suggesting AI-sourced visitors convert better or that AI mentions prime future branded searches. Supporting the skeptical view, one industry report showed zero-click searches at a record high in early 2026 while AI tools still drove under 2% of desktop visits—meaning the lost clicks are not quietly reappearing as AI engagement.

The numbers behind the anxiety

The wider data makes the stakes plain. Small publishers reportedly shed about 60% of their search traffic through early 2026, with Google Search dropping to roughly 5.8% of global page views by February—about half its share from the start of 2024. AI bots were found crawling retail sites nearly 200 times more per visit than Google, and a striking 80% of websites were exposed to agent spoofing, where bots disguise their identity to slip past controls.



Meanwhile, one major marketing platform disclosed a 27% year-over-year drop in customer organic traffic, zero-click searches on Google climbed from 56% toward 69% after AI Overviews launched, and Google's network ad revenue line slipped 4% to just under \$7 billion in the first quarter of 2026. Click-through research found that an AI Overview at the top of a results page can roughly halve the click rate for the leading organic result.

What it means for marketers

For teams juggling paid and organic budgets, the implications are concrete. Bot traffic inflates request counts without producing impressions, conversions, or any signal a campaign manager can optimize against, which muddies engagement metrics and session data. Tools that flag automated activity—now appearing inside mainstream analytics products—are becoming essential just to read the numbers honestly.



The deeper lesson concerns strategy. Content plans built on the assumption that strong rankings reliably convert into AI citations, and citations into traffic, deserve a hard look. The Cloudflare data is blunt: the largest stream of automated visitors arriving at most sites has no search intent at all. Generative engine optimization investments should be stress-tested against real conversion and revenue data—drawn from first-party sources rather than vendor-commissioned studies—because a bot that visits only to train a model leaves no commercial footprint whatsoever.

The bottom line

Two findings now sit side by side. Machines have become the web's majority audience, and most of them are reading to train models rather than to send anyone a visitor. At the same time, the consolation prize publishers were promised—visibility inside AI answers—does not appear to pay the bills for the businesses losing the most traffic. Together, they raise an uncomfortable question about who the open web is actually being built for, and who is footing the bill to keep it running.

