

She's in the Attic

The Complete Guide to AI Traffic, Attribution & Assisted Conversion in GA4

How to track where AI platforms are sending you traffic, understand what it's worth, and make the case for channels that do the work but don't get the credit.

Updated May 2026 · shesintheattic.com

Part of the Science & Art series · See also: 009 — Who Tells the Robots What To Do

What this guide covers

This is a single end-to-end reference for understanding AI traffic in GA4 and using that data to make a more complete case for organic and AI channel value. It covers setup, measurement, the structural limitations of the data, and the attribution and funnel analysis that turns raw numbers into a defensible argument.

The guide is structured in four parts:

- Part One: Setting up AI traffic tracking in GA4 — the channel group, the regex, the configuration
- Part Two: Why the numbers will be lower than you expect — referrer stripping, consent gaps, server logs, the GSC inflation issue
- Part Three: Attribution modelling — model comparison, first-click vs last-click, what the delta tells you
- Part Four: Path exploration and assisted conversions — real journey sequences, funnel analysis, custom reports, presenting findings

Prerequisites: A GA4 property with conversion events configured. Editor or Administrator access. If you are working with an e-commerce site, enhanced measurement and purchase events set up. Nothing here requires GA4 360.

Why AI traffic matters now

AI is developing at a pace that makes most digital marketing planning feel like it was written for a different era. Search engines still dominate, but the direction of travel is clear: AI platforms are becoming a meaningful source of referral traffic, and that traffic does not behave like anything that came before it.

ChatGPT has already announced paid advertising for 2026. There are credible reports of a native Shopify checkout being built into ChatGPT. Perplexity is running its own shopping and answer experiences with cited sources. These platforms are not sitting alongside the web — they are increasingly mediating access to it.

The question is not whether AI platforms will divert meaningful traffic and revenue. That is already happening. The question is whether you will be able to see it when it does, understand what it is worth, and make decisions accordingly.

Most GA4 setups are not configured to answer that question. This guide fixes that.

Setting up AI traffic tracking in GA4

Part One

The channel group, the regex, the configuration

Why the default GA4 channels miss AI traffic

Out of the box, GA4 has a default channel group that includes Organic Search, Direct, Referral, Paid Search and others. AI-referred traffic — visits that arrive after a user clicks a link in ChatGPT, Perplexity, Claude, Gemini or similar — gets dropped into Referral by default, if it gets attributed at all.

This means AI traffic is buried inside a channel that also contains link clicks from forums, directories, press coverage and partner sites. There is no way to isolate it, track its trend, or evaluate its quality without creating a dedicated channel.

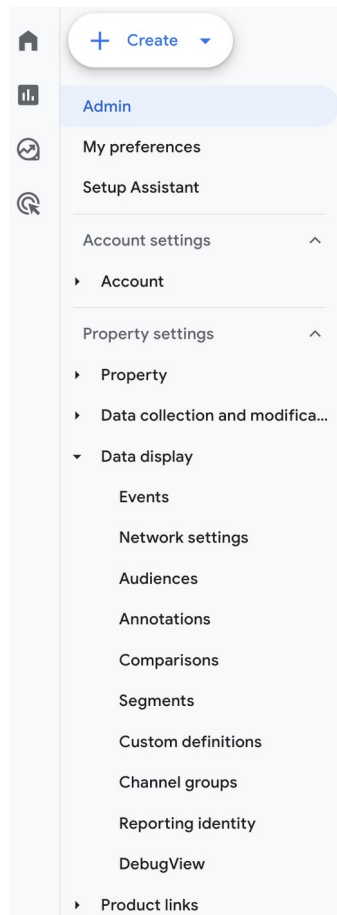
Step 1: Create a new Channel Group

In GA4, go to Admin (the cog icon, bottom left), then navigate to:

- Data Display
- Channel Groups

You will see your existing channel groups. Click the three dots next to Default channel group and select Copy to create new.

Name the new group 'Channel Group AI' and add a short description so anyone else with access to the property understands what it is.



Step 2: Add the AI Traffic channel

Channel groups

| Name | Date set to primary | Description |
|-----------------------|---------------------|---|
| Primary channel group | | |
| Default channel group | | Google Analytics predefined channel group |

Search Create new channel group

| Name | Last modified ↓ | Description |
|-----------------------|-----------------|---|
| Default channel group | | Google Analytics predefined channel group |

Copy to create new

Inside your new channel group, click Add new channel and configure:

- Channel name: AI Traffic
- Condition type: Source — Matches regex

Channel details

Channel name

AI Traffic

Channel conditions

Add conditions to define your custom channel

Match AT LEAST ONE rule in this group

Source

matches regex `^\\.openai\\.|\\.copilot\\.|\\.\\.ai\\b\\.|\\.chatgpt\\.|\\.gemini\\.|\\.gpt\\.|\\.neeva\\.|\\.writesonic\\.|\\.nimble\\.|\\.perplexity\\.|\\.google\\.bard\\.|\\.bard\\.google\\.|\\.edgeservices\\.|\\.bngpt\\.|\\.gemini\\.google\\.|\\.deepseek\\.|\\.meta\\.ai\\.|\\.grok\\.com\\.*$`

Paste the following regex into the condition field:

```
^\\.openai\\.|\\.copilot\\.|\\.\\.ai\\b\\.|\\.chatgpt\\.|\\.gemini\\.|\\.gpt\\.|\\.neeva\\.|\\.writesonic\\.|\\.nimble\\.|\\.perplexity\\.|\\.google\\.bard\\.|\\.bard\\.google\\.|\\.edgeservices\\.|\\.bngpt\\.|\\.gemini\\.google\\.|\\.deepseek\\.|\\.meta\\.ai\\.|\\.grok\\.com\\.*$
```

This regex has been updated from earlier versions to include Claude, DeepSeek, Meta AI and Grok. The `\\.ai\\b` pattern catches AI-named domains broadly. Review your sources periodically as new platforms emerge and add specific entries for any significant source not covered by the broad pattern.

Click Save.

Step 3: Re-order the channel group

This step is critical and frequently missed. GA4 assigns sessions to channels in order from top to bottom. If Referral sits above AI Traffic in the list, AI-referred sessions will be attributed to Referral instead of your new channel.



The screenshot shows the 'Reordering' interface in Google Analytics 4. At the top right, there are 'Cancel' and 'Apply' buttons. Below the title, there is a 'Channel name' header. A list of channel groups follows, each with a double-dash icon to its left. The groups are: Direct, Cross-network, Paid Shopping, Paid Search, Paid Social, Paid Video, Paid Other, Display, Organic Shopping, Organic Social, Organic Video, Organic Search, Email, Affiliates, AI Traffic, and Referral. The 'AI Traffic' and 'Referral' rows are enclosed in a red rectangular box, indicating they are the focus of the reordering step.

Click Reorder and move AI Traffic above Referral. Click Save group, then Apply.

Custom channel groups apply retroactively in GA4. Once saved, past sessions matching your regex will be reclassified in your reports. You do not need to wait for new data.

Step 4: Find the report

Navigate to Acquisition → Traffic Acquisition. Change the primary dimension from the default to your new channel group. Scroll to find AI Traffic in the channel list.

From here you can add secondary dimensions to slice further: landing page, device category, country, browser. For e-commerce properties, the Sessions and Revenue columns will show last-click attributed revenue for each channel.

Step 5: Extend to GSC and Search Console Insights

If you have Search Console linked to GA4, you can view organic search performance alongside AI traffic data in the same property. Go to Acquisition → Search Console → Queries to see which search terms are driving organic sessions alongside your AI channel data.

This is useful context for the attribution analysis in Part Three — understanding which queries bring organic traffic helps clarify where organic sits in the funnel relative to AI-referred visits.

Why the numbers will be lower than you expect

Part Two

Referrer stripping, consent gaps, server logs and the GSC inflation issue

Having the channel set up is only useful if you understand what the data is and is not telling you. AI traffic figures in GA4 will almost always be lower than the actual volume of AI-referred visits. Understanding why matters before you start drawing conclusions.

The referrer-stripping problem

GA4 can only attribute a session to an AI source if the referring platform passes a referrer header when the user clicks a link. Many do not.

Mobile apps and in-app browsers from ChatGPT, Claude and Perplexity frequently strip referrer headers entirely. When that happens, GA4 has no referrer to work with and classifies the visit as Direct traffic. This is not a GA4 failure — it is a structural limitation of how referrer data is passed between applications.

ChatGPT in particular behaves inconsistently depending on whether the user is on the web app, mobile app or API. Expect systematic undercounting of ChatGPT-referred traffic specifically.

The practical implication: if you see an unexplained rise in Direct traffic alongside growing AI channel data, a proportion of that Direct uplift is likely AI-referred sessions that lost their referrer in transit. The two signals should be read together.

AI citations do not always produce clicks

Being cited in a Perplexity or ChatGPT response does not guarantee a visit. Users frequently get the answer they need from the AI summary without clicking through to the source. This is the zero-click problem applied to AI: your content can be actively contributing to an AI platform's response while generating no measurable traffic at all.

This means GA4 session data understates your total AI exposure. Visibility and clicks are separate things. Only clicks show up in GA4.

Cookie-based tracking and consent gaps

GA4 relies on first-party cookies by default. Where users decline consent via a cookie banner, GA4 receives no data for those sessions. In markets with high decline rates — Germany and France particularly — this can represent a significant proportion of actual traffic.

GA4 uses modelled conversions to fill consent gaps, which is on by default. This creates a blurry line between observed data and inferred data. The model is not disclosed and cannot be audited.

Safari's Intelligent Tracking Prevention limits first-party cookie duration to seven days. A returning user after eight days looks like a new user in GA4. Multi-session journeys spanning more than a week are routinely broken for Safari users, which has direct implications for the path analysis covered in Part Four.

Cross-device journeys

Without User ID implementation — a persistent identifier passed when users are logged in — GA4 cannot connect sessions across devices. A user who first arrives via an AI citation on mobile and converts on desktop looks like two separate users with no connection between them. Most sites do not implement User ID. Most multi-device journeys are therefore invisible in GA4.

Sampling in Explorations

GA4's free tier samples data in the Explorations section above certain volume thresholds. Path Exploration — used in Part Four — sits in Explorations. For high-traffic properties, findings from path exploration are statistically representative rather than complete. GA4 360 removes sampling but carries significant cost. Treat path exploration findings from high-volume free-tier properties as directional.

What server-side logging adds

Server logs are recorded at the server before any client-side code runs. They are unsampled, consent-independent and unaffected by JavaScript blockers or browser privacy features. Every HTTP request is recorded regardless of user behaviour in the browser.

Server logs are particularly valuable for understanding AI crawler activity. GPTBot, ClaudeBot, PerplexityBot and other AI crawlers appear in server logs even when they generate no GA4 sessions. This gives a separate view of how AI platforms are indexing and interacting with your content, distinct from the referral clicks that GA4 tracks.

The limitation is that server logs record requests, not journeys. There is no session stitching, no conversion data, no engagement signals. You cannot build attribution models from raw server logs without significant additional processing. Referrer stripping affects server logs in the same way it affects GA4.

Server-side logging and GA4 are complementary rather than alternatives. Server stats validate what GA4 is and is not seeing. GA4 provides the session and conversion layer that server logs cannot. Neither source is complete on its own.

The GSC impression inflation issue

Google reported a Search Console logging issue that artificially inflated impression data between May 2025 and April 2026. If you are using GSC data alongside GA4 to benchmark or forecast organic performance, impression figures from this period should be treated with caution. Click data was less affected than impressions.

For forecasting work, this period should be flagged as a structural anomaly in any model that uses GSC impressions as an input.

She's in the Attic · Independent SEO Consultancy · shesintheattic.com

See also: 009 — [Who Tells the Robots What To Do](#)

This guide may be shared freely. Please attribute the source.