NOTES FROM THE BOT LAB · EDITION OF NOVEMBER 28, 2026

We Opened the "Direct" Box: What 1.86M Log Lines Show About AI Shortlists

By Will Tygart · Tygart Media · Canonical edition at https://tygartmedia.com/we-opened-the-direct-traffic-box-ai-bots/

Direct Answer: Across 1,859,626 tygartmedia.com server log lines over 31.8 days, raw direct traffic is 94.2% unhydrated bots. Filtering to real human browsers reveals Claude-User fetching deep guides 0–5 minutes before human arrivals at 4.42x chance, 23.4% returning AI first-touch buyers in GA4, and 37.5x more typed homepage visits than branded clicks.

Two days after publishing Billboards, Bots, and Heading Toward the Mountain, we audited **1,859,626** in-window server log lines from **tygartmedia.com** (31.84 days, Oct 28–Nov 28, 2026, quarantining 3,895 automated scanner networks) alongside 989 verified 30-second-plus active reader sessions, 90 days of Google Analytics 4 first-touch cookies, 90 days of Google Search Console data, and Bing Webmaster logs to see what lives inside raw "Direct" traffic.

1. Four Layers Inside Raw "Direct" Traffic

Before interpreting Direct visits, two synthetic layers must be removed: **94.2% (82,464 / 87,520)** of raw no-referrer browser sessions loaded 0 CSS/JS/font assets (**95.2%** of GA4 desktop-new unengaged), and an automated Nov 25–26 4-second headless sweep inflated **3,701** single-page loads under 10 seconds. Filtering to **>=10s clean human Direct** leaves 876 deep-slug arrivals alongside 428 typed homepage ('/') and hub sessions browsing **9.23x deeper multi-page (19.6% vs. 2.6%; 20.8% on '/')**.

Layer Inside Raw "Direct"	31.8-Day Volume	Browser & Depth Signature	What Is Actually Happening
1. Stealth HTML-Only Scrapers	82,464 sessions (94.2% of raw no-referrer)	0 real CSS, JS, or font assets loaded; 95.2% of GA4 desktop new unengaged	Automated scrapers cloning pages without hydrating a real browser.
2. Automated Headless Bursts (<10s)	3,701 deep-slug sessions	Asset loads complete in under 10s with 0% multi-page depth (incl. Nov 25–26 4-slug burst)	Short headless sweeps (85–92 hits each across 4 slugs) masquerading as browsers.
3. Clean Deep Direct (>=10s)	876 origin / 187 active 30s+ sessions	2.6% multi-page overall; led by /claude-desktop-app/ (23) & /claude-edu/ (22)	Desktop apps and chat windows stripping Referer headers on deep-guide clicks.
4. Typed Homepage '/' & Hubs	428 origin sessions (375 '/' + 53 hubs)	19.6% multi-page (9.23x deeper) and 22.0% same-network repeat (2.57x higher)	Remembered brand visits returning days after shortlisting Tygart Media.

2. The Bot-to-Human Stopwatch & Zero-Search Direct Proof

Claude-User fires before humans arrive: Across 31.8 days, Claude-User pulled exact deep slugs 0–5 minutes *before* a Direct human landed **14 times vs. 3 times after (4.42x above matched time-of-day chance; 6.62x on clean >=10s Direct)** and **37 vs. 12 inside 30 minutes (3.08x)**. By contrast, our reactive search control **DuckDuckBot/1.0** ran backward (**5 Pre vs. 24 Post = 0.21x** within 30m of 1,474 DDG clicks), while **ChatGPT-User** ran steady background caching (**724 hits/month** on a single pricing guide, 68 of 160 visible AI-referrer clicks landing 1h–7d after live-agent pulls).

Across **2,189 deep URLs** controlling for Google/Bing impressions and word count, live AI agent fetches predict 30-second-plus active Direct readers at **5.53x odds (8.59% vs. 1.67%, 5-fold ROC-AUC = 0.8325)** and >=10s origin Direct at **3.98x odds**. Even among Direct pages earning **0 Google clicks and 0 Bing clicks**, **45 of 65 (69.2%)** had live AI agent pulls — led by `/claude-desktop-app/` (17 live pulls, 23 Direct).

3. Returning AI Buyers Masked by IP Churn & GA4 Last Non-Direct Click

At the origin network level, **209 of 4,459 (4.69%)** Direct-first networks returned Direct >=1h later (158 at 24h+, **median lag 54.62 hours = 2.3 days**; 141 switched deep slugs), while mobile/VPN IP rotation yielded 0 exact-IP matches across 148 AI-referrer networks. Switching to **90-day GA4 first-touch cookies** uncovers Core AI buyers returning at **23.4% (44/188) vs. 13.5% (800/5,930) Search (1.96x odds)** — with **41 of 44** overwritten back onto AI source credit under GA4 Last Non-Direct Click:

90-Day GA4 First-Touch Source	Returning Users %	Visible Under AI Source	Visible Under (direct)	How Last-Click Reporting Masks It
Claude.ai (78 cohort)	28.21% (22 returning)	19 returning sessions	3 returning sessions	86% of returning visits stayed inside Claude.ai session credit.
Microsoft Copilot (57 cohort)	22.81% (13 returning)	13 returning sessions	0 returning sessions	100% overwritten back onto Copilot by Last Non-Direct Click.
ChatGPT (53 cohort)	16.98% (9 returning)	9 returning sessions	0 returning sessions	100% overwritten back onto ChatGPT despite repeat buyer returns.
Search Baseline (Bing + DDG + Google)	13.5% (800 / 5,930)	Baseline control cohort	Small direct leakage (7)	Core AI first-touch buyers return at 1.96x higher odds (23.4% vs 13.5%).

4. Branded-Search Bypass & Macro Lag

Typed `tygartmedia.com/` visits (**375** in 31.8d; 52 GA4 engaged in 30d) outnumbered **90 days** of combined Google (6) + Bing (4) branded clicks (**10 total; 8 exact 'tygart media'**) by **37.5x (46.9x exact)**, and daily live AI pulls correlated positively with typed '/' + hub traffic **2 days later (Spearman rho = 0.3915, p = 0.0324)**.

Honest Measurement Boundaries & Takeaway: Origin logs cannot stitch a Day-1 zero-click AI mention across rotating mobile/VPN IP addresses without cookies, and automated 4-second headless sweeps slip past raw asset filters until you enforce a >=10s span gate. Build your homepage and deep technical specs as a unified landing strip for AI-shortlisted buyers arriving days later without search clicks.