September 2026

Is This the Death of Open Web Programmatic?

The auction ends before the truth begins: agentic traffic, the auction's blind window, and why the open auction is unlikely to fix itself in time

STFP
network

The real-time auction completes in 200-1400 ms - before the visitor has touched the page. Every signal that can expose a sophisticated AI agent comes into existence after that window has closed. The problem is not detection technology. It is that the commercial decision is locked in before the evidence exists - and the invoice chain cannot carry a correction backwards.

In one Danish campaign, the layers below the ad server counted up to 2.3 times what the ad server verified - with no one in the chain obliged to resolve the gap.

Contents

STFP

network

STEP

network

Executive Summary

For more than 15 years, advertising on the open web has been bought and sold on a single assumption: that the visitors to websites are people. That assumption is breaking - and the part of the market least able to absorb the damage is the independent open auction that funds most of the open web.

Three facts frame the problem. More than half of all traffic on the internet is now automated rather than human (Cloudflare Radar, 2026). Traffic from AI agents - software that browses, compares and buys on a person's behalf - grew 7,851% in a year (HUMAN Security, 2026). And when researchers tested seven widely used AI browsing agents, a leading commercial bot defence recognised only one of them; analysing their behaviour afterwards caught all seven (Wang et al., 2026).

One Danish campaign shows what the money side of this looks like in practice:

> ***Same campaign. Same inventory. The creative layer counted 1.8-2.3 times the impressions the publisher's ad server had verified.***
>
> *Nobody in the chain is contractually obliged to resolve a gap like that. Today it is a measurement nuisance. As agent traffic grows, it becomes the hole synthetic traffic is billed through. (DAMA Working Group, 2026; dama.media - anonymised data in Appendix A.)*

The claim of this paper, in plain words: the open auction decides and bills faster than the truth can arrive and the billing chain cannot carry a correction backwards.

The four steps:

1. The auction is over before the visitor has touched the page. Buyers must price and respond within 200-1400 ms. The signals that reveal an AI agent - how it scrolls, types and moves the mouse - only come into existence seconds later. The decision is locked in before the evidence exists.

2. The auction sees almost nothing, and checking afterwards uses the same thin data. The only information buyers receive before deciding is a short technical message, the "bid request", and an agent running in a person's own browser inherits everything in it: the cookies, the logins, the network address. Every field reads "human". The behavioural evidence that could reveal the agent stays on the publisher's page and never reaches the buyers, not in real time and not afterwards.

3. The money flows one way; the truth sits at the other end. Money moves down a chain of middlemen (advertiser → buying platform → marketplace → publisher), and each link settles on counts made inside the blind window. The publisher, who knows most about the traffic, has no say over the invoice. When we mapped the five setups used in practice, only one could correct mistakes automatically after the fact: Google's own end-to-end path. The open path most publishers depend on has no correction channel at all, and gaps of 5-15% between counts are normal (STEP Network, 2026; see the source note in 6.3).

4. If nothing changes, the winner is Google and the other walled gardens. Inside one company's stack, counting, validation and billing reconcile automatically. When buyers stop trusting the open auction, they move there in self-defence - and the open web becomes a tenant in someone else's house.

A necessary caveat on scale – and on urgency.

Undeclared agentic traffic is still a minority of open-web impressions, and for many publishers the immediate revenue impact is limited. This paper is an argument about the trajectory: agentic browsing is on course to become the normal way people use the web. When that happens, a flaw that is merely expensive today becomes existential – and the time to redesign settlement is while the problem is still small enough to get ahead of.

To move from argument to evidence, STEP Network ran a controlled experiment in June 2026: scripted, human-like browser traffic - labelled invisibly before the auction so every layer's reports could be segmented on it. It was sent through live publisher pages carrying a full programmatic stack, and every link's reports were pulled twice: the day after the traffic, and again after the 30-day correction window. Every synthetic impression that cleared the auction was bought and billed as human; a month later the ad server had quietly restated 1.95% of the traffic - and that correction reached no other ledger in the chain (methodology and results in Chapter 12).

This is deliberately a problem paper. Chapter 13 sets out the principles any viable successor must satisfy, and notes that STEP Network is actively developing such a model with Danish publishers and with the DAMA group (Danish Advertiser and Media Alliance).

1. The internet has changed character

Bot traffic is not new. Crawlers, scrapers and simple scripts have existed as long as the web itself, and the industry has built mature tooling against them: user-agent lists, IP reputation, headless-browser detection, technical fingerprints. What is new is the speed and the nature of the change. This is generally also what we would describe as Invalid Traffic (IVT) (HUMAN Security, n.d.).

1.1 The crossover has happened

On 3 June 2026, Cloudflare's co-founder and CEO stated that automated traffic now accounts for **57.5% of all HTTP requests** to web content, against 42.5% from humans - the first such crossover in the history of the internet, arriving roughly 18 months earlier than his own prediction, with agentic AI as the primary driver (Cloudflare Radar, 2026; Prince, 2026). One methodological caveat matters: this measures HTTP requests to HTML content, not time spent - humans still dominate by engagement. But for the advertising ecosystem, requests are precisely the relevant unit: page loads trigger auctions, and impressions get billed based on this functionality, not human engagement.

1.2 Agentic traffic is growing explosively

HUMAN Security's 2026 State of AI Traffic & Cyberthreat Benchmark Report, based on more than one quadrillion digital interactions analysed in 2025, documents the shift:

- Automated traffic grew 23.5% year over year - eight times faster than human traffic (3.1%).
- AI-driven traffic grew 187% in monthly volume through 2025, peaking at 3.61 times the January level in October.
- Traffic from AI agents and agentic browsers grew 7,851% year over year.
- 2.3% of all agentic activity now occurs on checkout pages - agents no longer merely read the web, they transact on it.
- More than 95% of AI-driven traffic in 2025 was concentrated in retail/e-commerce, streaming/media, and travel. Media sits at the epicentre.

Crucially, the growth is dominated by browser-based agents - agents operating inside real browsers, in many cases the user's own - rather than datacenter bots. In HUMAN Security's (2026) reporting, two agentic browsers alone accounted for two-thirds of observed agentic traffic, and browser agents made up roughly 71% of all agentic activity in April. This detail decides everything that follows: the fastest-growing traffic class inherits a real human's technical identity.

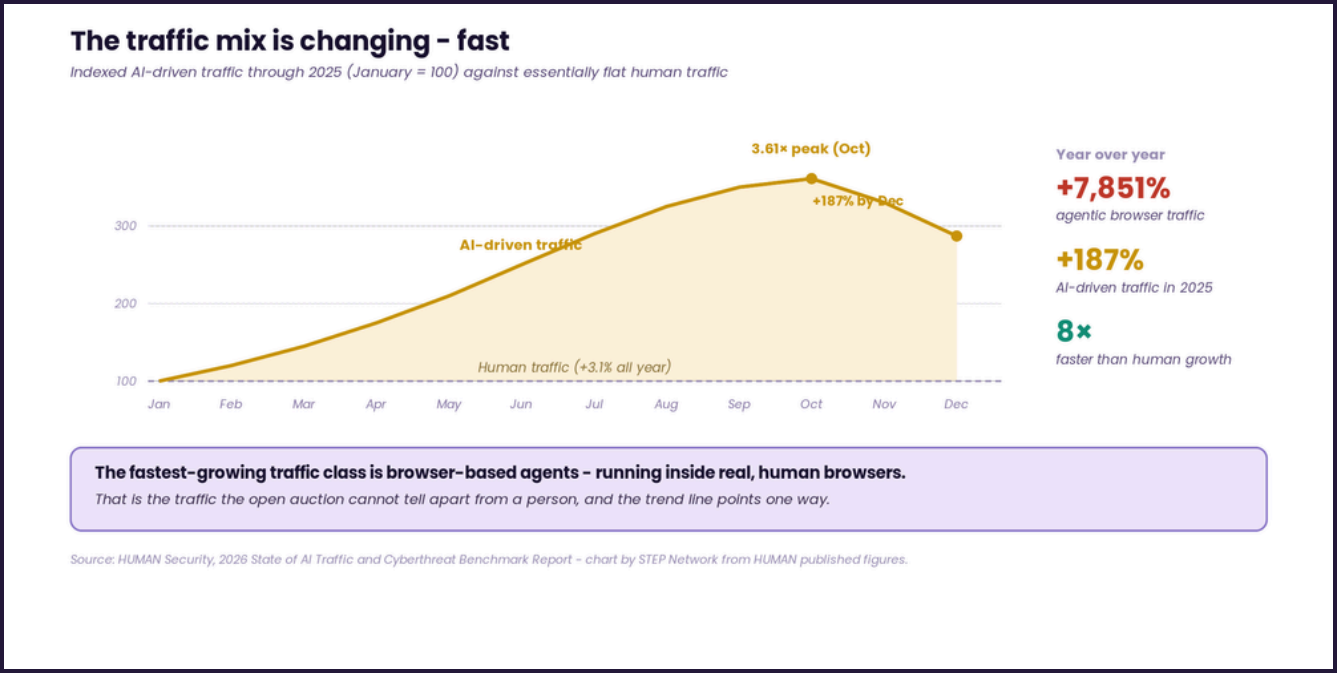

Figure 1. *Indexed AI-driven traffic through 2025 (January = 100), peaking at 3.61× in October and ending the year up 187%, against essentially flat human traffic (+3.1%). Agentic browser traffic grew 7,851% year over year. Source: HUMAN Security, 2026 State of AI Traffic & Cyberthreat Benchmark Report; chart by STEP Network from HUMAN's published figures.*

1.3 Classic bots vs. browsing agents

Traditional bots	Modern browsing agents
Crawl pages systematically	Perform concrete tasks on behalf of a user
Often self-identify via user agent	Present as an ordinary browser
Typically headless, datacenter IPs	Real browsers - often the user's own, on the user's IP
No cookies, no history	Inherit the user's cookies, logins and sessions
Leave technical fingerprints	Technical fingerprint is the user's fingerprint

The entire existing Invalid Traffic (IVT) defence, including the industry's shared bot lists and crawler registries, was built for the left-hand column in the table above. The right-hand column is the one growing at four-digit rates. Under the framework of MRC, the Media Rating Council, whose Invalid Traffic Detection and Filtration Guidelines set the industry standard for identifying and filtering invalid traffic (Media Rating Council, n.d.-a, n.d.-b), it is by definition SIVT: Sophisticated Invalid Traffic, requiring multi-point behavioural analysis rather than list lookups. The rest of this paper examines whether the open auction is capable of that analysis. The answer, we will argue, is no - for reasons of timing and information architecture, not technology.

2. The hidden bill: what one Danish campaign revealed

Before the theory, the evidence - with the numbers stated precisely, because they are striking enough without exaggeration.

In a recent Danish engagement (hereafter the DAMA case), the DAMA Working Group - a Danish advertiser-and-media collaboration (dama.media) - compared, line by line, what the publisher's ad server verified against what the creative-hosting layer counted, for the same campaigns, on the same inventory (DAMA Working Group, 2026; anonymised line-level data in Appendix A):

- On the affected setups, the creative layer counted 1.8-2.3 times the impressions the ad server had verified - that is, 80-130% more.
- And the control that proves the point: one identically formatted setup showed no gap at all (1.00×). The discrepancy is architectural - it depends on how the chain is wired, not on chance.

Same campaign. Same inventory. Two counts - and the invoice follows the higher one.

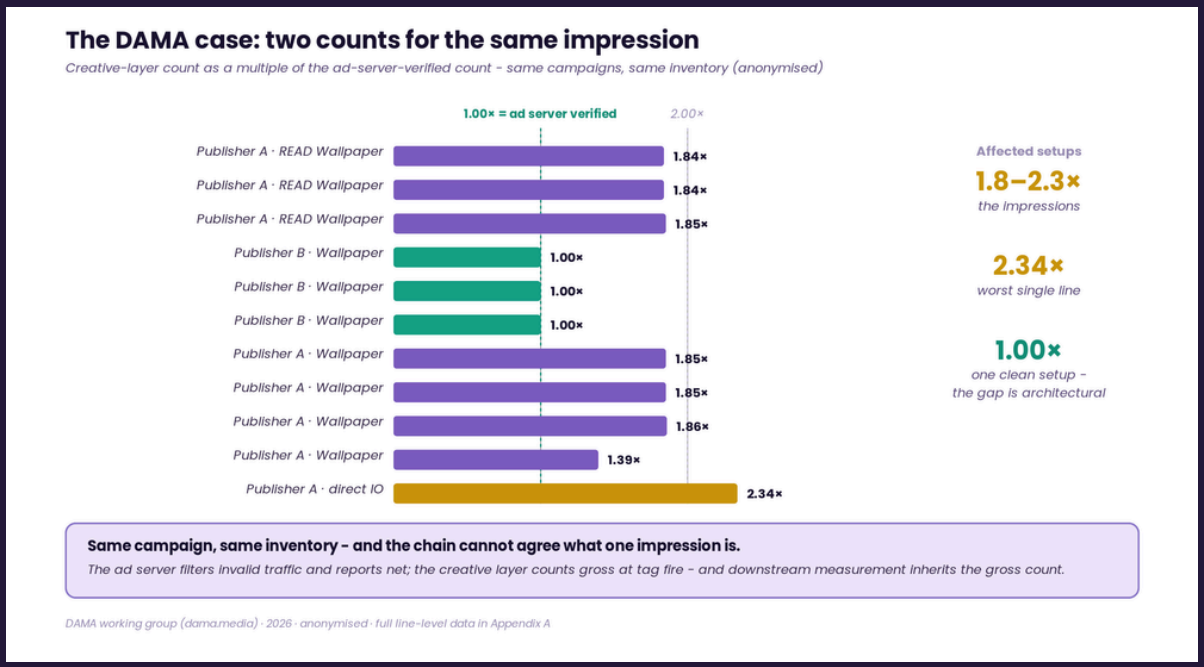

Figure 2. *The DAMA case, line by line (anonymised). Creative-layer count as a multiple of the ad-server-verified count. Three lines at 1.00× prove the gap is architectural; affected lines run 1.4-2.3× on impressions. Source: DAMA Working Group (dama.media), 2026.*

2.1 What actually causes the gap

It would be convenient to present these numbers as direct proof of bot or agent traffic. They are not, and this paper will not pretend otherwise. The mechanism is more mundane and for the argument that follows, more damning.

The publisher's ad server filters invalid traffic live and post-serve and reports net (Google, n.d.-b). The creative-hosting layer counts the moment its creative files fires - before the ad has even rendered successfully - applies little to no filtration, and reports gross (DAMA Working Group, 2026). And because verification and measurement tags physically execute inside the rendered creative, every downstream party inherits the gross count automatically. No channel reports the ad server's removals to any of them.

2.2 Why this is the agent problem in miniature

Read correctly, the DAMA case proves something more fundamental than any single bot rate: the chain cannot agree on what one impression is - even for ordinary traffic, even with nobody trying to deceive anyone. Different links count at different moments, filter at different points, and bill on their own numbers, with no mechanism to reconcile.

Now add agentic traffic. An undeclared agent's impressions land precisely in this gap: counted gross by the layers that measure and bill, filtered - if at all - only by the layer nobody settles on. Today the defect is a measurement nuisance. At agentic scale, it becomes the hole through which synthetic traffic is billed: invisibly, and with no contractual owner. The chapters that follow explain why the gap cannot be closed from the demand side, and why its economics fail as agent traffic grows.

3. The 1400-millisecond problem

This chapter contains the central argument of the whitepaper. Everything else follows from it.

3.1 Two timelines that never meet

Consider what happens, in sequence, when a visitor - human or agent - lands on a page monetised through header bidding. The auction infrastructure (Prebid and equivalents) fires bid requests to demand partners. Each demand-side platform must price and respond within a window typically of 200-1400 ms. The auction resolves, the winning creative is handed to the ad server, and the impression is served and counted - all within roughly the first second or two of the page's life. Now consider when the evidence arrives. The signals that, per current research, actually distinguish a sophisticated agent from a human are behavioural: scroll velocity and rhythm, typing cadence, mouse micro-movements, interaction depth, dwell time, render completion. None of these exist at auction time. They begin to accumulate seconds after the impression has been served, and they only become statistically reliable across the session - or across many sessions.

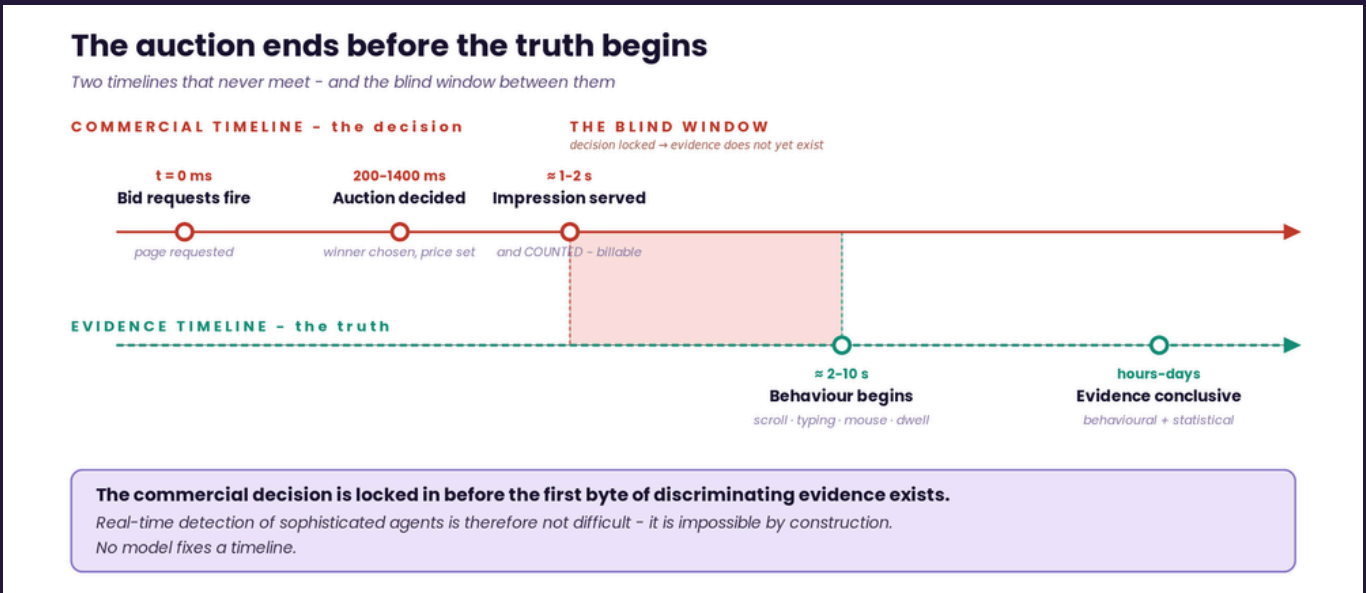

Figure 3. *The two timelines. The commercial decision (bid → auction → serve → count) completes within roughly one second; the behavioural evidence that distinguishes an agent only begins seconds later and becomes conclusive over hours to days. The gap is the blind window. Illustrative; positions not to linear scale.*

One further fact stretches the evidence timeline even further: the ad server's own numbers are provisional. Google documents that Ad Manager reporting data can continue to adjust for up to roughly 30 days, with statistics typically locking month by month (Google, n.d.-g); in revenue terms, Google states that revenue is only finalised once records are "processed and verified by the 2nd of the month" (Google, n.d.-c). Even the system's own ground truth is finalised weeks after the millisecond in which the money was committed.

3.2 Why this is a constraint, not a capability gap

It is tempting to read this as a technology problem: detection models will improve, signals will get richer, the window will be used better. That misses the point. The auction window cannot meaningfully expand - latency is user experience, and user experience is revenue. And the evidence cannot arrive earlier - behaviour that has not happened yet cannot be measured. The two timelines are fixed by physics and economics respectively, and they do not overlap.

The commercial decision - price, win, serve, count - is locked in before the first byte of discriminating evidence exists.

Real-time SIVT detection in the open auction is therefore not merely difficult - it runs against the grain of the protocol's own timeline. Settlement models that assume the auction can see the truth are betting against physics. Whether the gap can be engineered or legislated shut, and fast enough to matter, is the question Chapter 10 takes seriously - and answers with measured pessimism.

3.3 The narrowing exception and an important caveat

Two caveats keep the argument precise. First, unsophisticated automation - declared crawlers, headless browsers, datacenter IPs - is caught today by list-based and reputation-based filtering inside the window, and will continue to be. This class is known as GIVT, General Invalid Traffic (Media Rating Council, 2020, Section 1.1.2). The argument concerns the growing class that is not: agents in real browsers with inherited human identity. Second, a sufficiently sloppy agent (uniform timing, no jitter) can leak hints early. But sloppiness is a bug agents fix, and the trajectory of agent development points one way: towards behavioural indistinguishability inside the window. The exception narrows every quarter; the constraint is permanent.

3.4 Why the bid request cannot see a bot

There is only one reliable way to separate a sophisticated agent from a human: fingerprinting, specifically the behavioural fingerprints (typing, scroll, mouse) that the next chapter examines. The decisive fact for the auction is that none of those fingerprints exist in the bid request. OpenRTB, the message format every auction using Prebid travels in, defines a fixed set of fields: device, geo, a user object, IP, user agent, placement and a handful of identifiers (IAB Tech Lab, 2016). There is no field that carries a live behavioural signature, because at bid time the behaviour has not happened yet (Chapter 3). A buyer evaluating a bid request therefore has nothing fingerprint-grade to assess: the request is starved by design, not by oversight.

4. The auction is nearly blind and the agent looks exactly like you

4.1 A deliberately narrow channel

The timing problem is compounded by an information problem. The bid request - the only evidence the demand side ever receives before deciding - is a deliberately constrained channel, designed for speed: IP address, user agent, device signals, cookie/ID information, placement and page metadata. A few dozen fields. It was never designed to carry proof of humanity, because when the protocol was designed, humanity was the default.
Set that against what the publisher's side observes across a session: full page and render context, scroll and pointer behaviour, interaction depth, session history, login state, first-party behavioural baselines per user. The publisher and its ad server operate on orders of magnitude more evidence - all of it arriving after the auction.

In the bid request (pre-auction)	Only observable post-render (publisher side)
IP address	Scroll velocity, rhythm and depth
User agent	Mouse micro-movement and typing cadence
Device signals	Render completion and interaction timing
Cookie / ID information	Dwell time and engagement
Placement, format, page metadata	Session-over-session behavioural baselines

Table 1. *What the bid request can carry versus what only the rendered page reveals - the information asymmetry of the open auction.*

4.2 The agent inherits the user

Here is what makes the channel poverty fatal rather than merely limiting. The fastest-growing agents operate in the user's own browser, or in environments the user has authorised. Such an agent inherits the user's cookies and IDs, logins and sessions, IP address and device - increasingly even payment credentials, a prerequisite for agents that transact on the user's behalf. Every field the bid request can carry therefore reads "known human". There is nothing for a pre-bid filter to filter on: the agent's technical identity is the user's technical identity.
This also disposes of the intuitive remedy. Login separates humans from anonymous bots - but not the user from the user's own agent. An agent reading an article on behalf of a logged-in subscriber presents as an authenticated, known user with full history. First-party cookie absence can flag fresh headless automation, but that signal erodes precisely as agent browsing moves into the user's own browser.

A near-term accelerant deserves naming. The first large spike in agentic browsing is likely to arrive when Google rolls out its agentic-first search experience at scale - an agent that visits, reads and summarises pages on the user's behalf. Inside Google's own stack that traffic can be recognised and reconciled (Chapter 8). The blunt question for everyone else: Google can be assumed to account for its own agents inside its own ecosystem - but what about the rest of the internet, where those same agents arrive as ordinary, logged-in browser sessions with nothing to tell them apart from the human whose browser they run in? Nor is undeclared crawling hypothetical. Cloudflare has publicly accused Perplexity of using stealth, undeclared crawlers to evade websites' explicit no-crawl directives across tens of thousands of domains (Cloudflare, 2025), and Reddit is suing both Anthropic and Perplexity, together with three data-scraping firms, over large-scale collection of its content without permission (CNBC, 2025). Content is already being fetched at scale against publishers' declared preferences; undeclared agentic visits will not announce themselves either.

4.3 The empirical anchor: FP-Agent

In May 2026, researchers at UC Davis published the first controlled measurement study of AI browsing agents: FP-Agent: Fingerprinting AI Browsing Agents (Wang et al., 2026). Seven of the market's most widely used agents, plus human users, performed everyday tasks - flight booking, shopping, forum interaction - on an instrumented website collecting both technical and behavioural fingerprints. Two findings matter here:

1. Browser fingerprints are weak. Technical identity signals - the foundation of the classic defence, and the only class of signal a bid request can carry - provided limited discriminative power.
2. Behavioural fingerprints are distinctive. Typing rhythm, scroll patterns and mouse movement separated every agent from humans and from each other - signals that exist only post-render, only on the page.

The study's case study makes the point operational: a leading commercial bot defence, sitting in front of roughly a fifth of the web, identified 1 of the 7 agents. The behavioural classifier identified all 7. That is not a vendor failing. It is the timing constraint of Chapter 3 expressed as a detection rate: identity-based, request-time detection is structurally behind, and the signals that work live where the auction never looks.

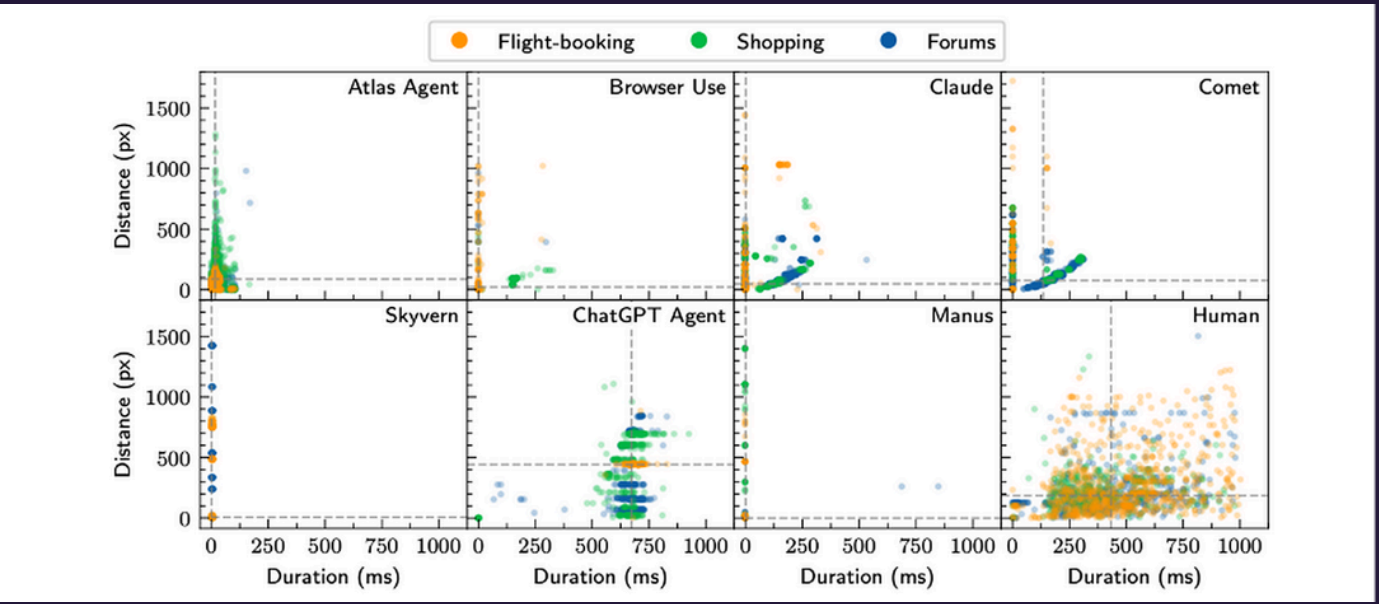

Figure 4. *Verdict on AI-agent detection (request-time detection caught 1/7; behavioural fingerprinting caught 7/7) with stylised keystroke-timing patterns for seven agents and a human baseline. Adapted from Wang et al. (2026).*

5. Why checking afterwards does not fix it

The standard industry response to Chapters 3 and 4 is: “we know real-time is hard - that is why we filter post-bid.” Exchanges and demand platforms do run post-auction invalid-traffic analysis, and supply platforms do deduct identified invalid traffic (IVT) from publisher payouts after the fact. This chapter explains why that practice, as architected today, cannot close the gap - and why claiming it does concede the argument.

5.1 Afterwards, the buyers re-analyse the same thin data

The decisive question is not when the analysis runs, but what data it runs on. Downstream post-bid filtering analyses the data the downstream party possesses: bid request fields, win/loss logs, aggregate request patterns, list lookups. It never possesses the post-render behavioural evidence - scroll, typing, mouse, dwell, render - because that evidence is generated on the publisher’s page and is never transmitted into the bidstream. If no discriminating signal can be sent in real time, there is no discriminating signal for the demand side to post-analyse. Running yesterday’s starved signals through a better model tonight does not create evidence that was never collected.
This is the precise meaning of the DAMA discrepancy: downstream layers counted up to 2.3 times what the ad server’s post-bid classification - informed by the publisher-side reality - had verified. Only one side had the evidence.

5.2 The learning-loop defence - and its missing closure

The sophisticated version of the industry response is: “post-bid analysis feeds a learning loop; we get better at request-time decisions over time.” Three things are wrong with this as a rescue of the open auction. First, the loop trains on the wrong data (see above). Second, the adversary is not stationary: agents are trained on human behaviour and improve at imitating it, so request-time signatures decay as fast as models learn them. Third, and decisively, a learning loop only closes if its corrections are binding: if every impression later identified as invalid produced an automatic, full correction of money and reporting, the loop would discipline itself. In a chain of independent companies, it does not, for the economic reasons set out in Chapter 6. A learning loop without binding settlement is an apology with a roadmap.

5.3 A web of competing thruths

Layer on the rest of the chain and the picture darkens further. Between advertiser and publisher sit not just a demand platform and a supply platform but verification vendors, creative-serving vendors, measurement and attribution vendors, each sampling different points of the transaction, each applying different methodologies and thresholds, each producing a different invalid-traffic number for the same campaign. One vendor flags 20%, another 12%; the demand side prefers one number, the supply side the other; the publisher's ad server saw a third reality. None of these truths is binding on any other. There is no arbiter, no escrow, and no contractual mechanism by which the most-informed measurement - the publisher-side one - prevails.

> ***Post-bid filtering, as practised downstream, is not a solution to the timing problem. It is a confession that the bid should never have been priced as human - dressed as due diligence.***

6. The money flows one way - the truth sits at the other end

Chapters 3-5 established an information asymmetry. This chapter establishes why the asymmetry is economically fatal: the settlement architecture points the wrong way.

6.1 Detection capability and invoice authority live in different places

Rank the chain by evidence and the order is: publisher and ad server first (full post-render reality), exchange and supply platform second (request patterns at scale), demand platform last (a starved bid request). Rank the chain by settlement authority and the order inverts: the advertiser pays the demand platform, the demand platform pays the supply platform, the supply platform pays the publisher - each link settling on counts made inside the blind window.

The party with the least information makes the binding financial decision. The party with the most information has no authority over the invoice. That single inversion is the structural rot of open web programmatic. Everything else - fraud rates, discrepancy disputes, verification disagreements - is downstream of it.

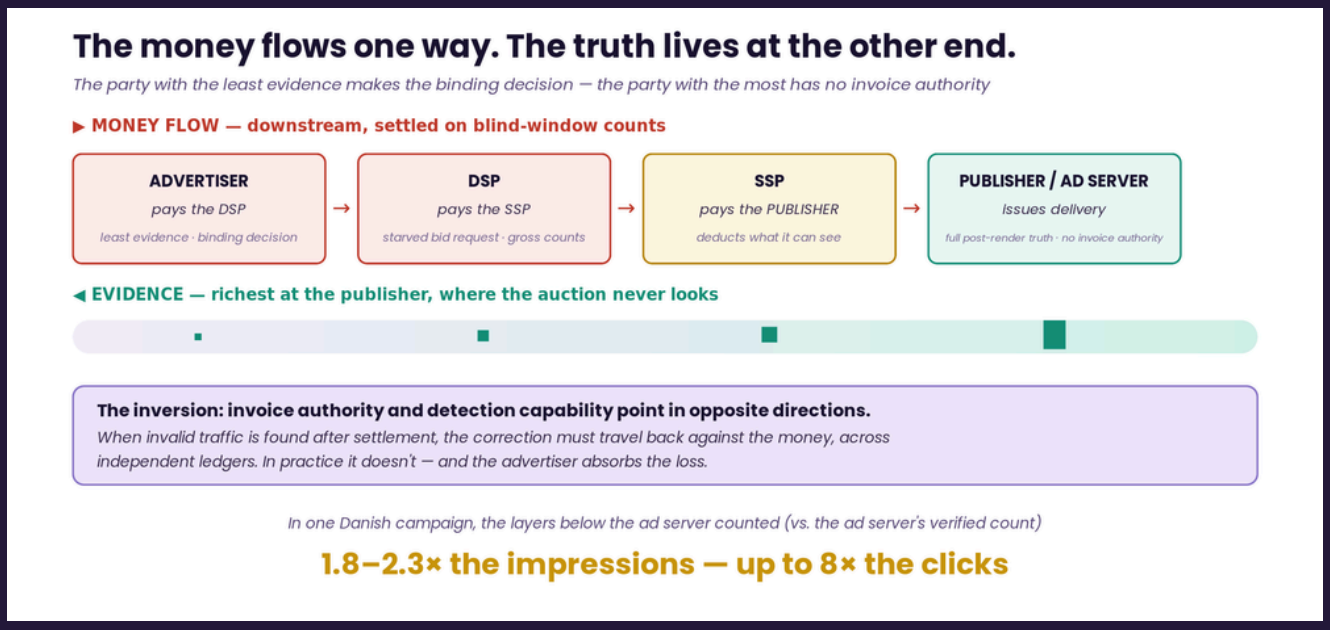

Figure 5. *The inversion. Money settles downstream on blind-window counts (red), while evidence is richest at the publisher, where the auction never looks (green). A correction found after settlement must travel back against the money across independent ledgers - and in practice does not.*

6.2 The deficit that lodges in the middle

Trace what happens when invalid traffic is identified after settlement as, per Chapter 3, it only can be. The advertiser's platform reports a flagged share and the advertiser requests a correction. The demand platform turns to the supply platform; the supply platform has already paid the publisher in near-real time and holds no escrow; the publisher's own ad server may classify the traffic differently in the first place. Each party faces the same menu: absorb the loss (margin destruction), claw it back from the next link (relationship destruction), or dispute the measurement (delay until the question dies). In practice, supply platforms deduct the invalid traffic they can detect from publisher payouts - based, inevitably, on the starved signal set - and the deeper corrections simply do not propagate. The deficit lodges in the middle of the chain, and the residual loss lands on the only party with no one left to charge:

> ***The structural loser of the open auction is the advertiser - paying real money, at every link's margin, for traffic that no link could verify and no contract can retroactively correct.***

6.3 Five real-world setups, one pattern

Assertions about “the chain” risk staying abstract, so we mapped the five concrete setups STEP Network actually operates - who counts, who bills on whose count, and whether a correction can ever travel back against the money.

Setup	Who is billed on whose count	Channel for corrections	Typical gap
All-Google: DV360 → AdX → Google Ad Manager	Google's count end to end	Automatic and internal - same logs; invalid-traffic removals and credits propagate by themselves	~1-2% (Google, n.d.-a, n.d.-i)
DV360 → third-party exchange → Open Bidding	Ad server's post-filtering “billed impressions”, settled three days later	Automatic via Google's Demand Discrepancy Resolution - beta, Google demand only	Small residual (Google, n.d.-e, n.d.-f)
Any DSP → SSP → header bidding (Prebid) → ad server	The SSP's own count at both hops; the ad server's number consulted by nobody	None	5-15% (Databeat, n.d.; STEP Network, 2026)
Non-Google DSP → AdX → ad server	Google's billed count; the advertiser sees the DSP's own count - two sets of books	None automated; manual invoice-level disputes	Varies (STEP Network, 2026)
Direct insertion order with advertiser-side tag	Ad server count (publisher invoices on it); advertiser verifies on its own tag	Manual monthly reconciliation	1-10%, direction flips (Google, n.d.-a; STEP Network, 2026)

Source basis: settlement mechanics are verified against vendor and standards documentation (Google, n.d.-a, n.d.-e, n.d.-f, n.d.-i; IAB Tech Lab, 2016, 2022); typical gaps reflect industry experience and STEP Network’s own measurement, including the DAMA case and the creative-host over-count described below (DAMA Working Group, 2026; STEP Network, 2026).

Two patterns leap out. First: the closer a path runs to one company's stack, the better corrections propagate: fully automatic inside Google's own path, nonexistent in the open header-bidding path this paper is about. Second: in every setup, the creative-hosting layer splices in one more count (gross, essentially unfiltered, taken the moment its template fires), and because verification tags execute inside its render, every downstream party inherits that gross count. That is precisely the DAMA mechanism of Chapter 2.

The protocol itself concedes that no shared definition exists. In OpenRTB, the win notice (nurl) is explicitly not billable, and the billing notice (burl) fires according to each exchange's own billable-event policy: "billable" is defined by the party sending the bill - not by any shared standard (IAB Tech Lab, n.d., 2016, 2022).

A further distortion sits inside the auction before any impression is even counted: bidders self-report whether their bids are gross or net of fees, and unless the publisher maintains per-bidder adjustment factors, gross bidders beat net bidders at equal real value - so the booked price and the paid-out price diverge by construction (Prebid.org, n.d.; Sovrn, n.d.).

6.4 Why this compounds at agentic scale

At low single-digit SIVT rates, the chain absorbs the friction: discrepancies are written off, disputes are rare, the leakage is priced in. The growth rates in Chapter 1 remove that comfort. As the agentic share of open-web traffic climbs, the unverifiable share of every gross count climbs with it - and a pricing model denominated in impressions, settled inside the blind window, has no mechanism to adjust. CPMs cannot simply fall to compensate: publisher costs do not scale down with traffic quality, buyers will not knowingly bid on unverifiable inventory at any price, and contracts promise human attention that the chain cannot document. A market that cannot reconcile its own counts gets repriced by its buyers - or abandoned by them.

7. A long tail without defences

One objection deserves its own short chapter: “publishers can defend themselves at the gate.” It is true that a publisher can deploy pre-auction defences - challenge flows, sophisticated gating, custom behavioural scoring wired into its ad serving. The most resourceful publishers in the world do some of this today.

But the open web is not five mega-publishers. It is a long tail running from large national houses to local news sites, vertical media and niche communities - and that tail is precisely what the open auction was supposed to monetise. For the long tail, gate-level defence is out of reach three times over: the engineering and data-science investment is unaffordable; the behavioural baselines require traffic scale a single small publisher does not have; and any defence aggressive enough to stop sophisticated agents will, run naively, also block real readers - revenue a small publisher cannot spare. Challenge walls also tax every legitimate visitor to catch the few.

The result is a two-tier open web: a handful of publishers who can partially defend themselves, and a long tail that structurally cannot - inheriting the full force of the problem with none of the tools. Any credible answer to the diagnosis in this paper must work for the tail, not only the head. A defence that requires scale must therefore be built collectively - a point we return to in Chapter 13.

8. Who wins if nothing changes: Google and the walled gardens

This chapter is scenario analysis - STEP Network's assessment of market drift if the structural fault is left unaddressed. It is an observation about architecture and incentives, not about any company's intent.

8.1 Precision about what is dying

It is important to be precise, because the protocol is not the casualty. OpenRTB, the message format, will keep clearing impressions, including inside vertically integrated stacks. What the preceding chapters condemn is a specific architecture built on top of it: the independent open auction: header bidding, the Prebid-era infrastructure, and the open web monetisation layer in which independent demand platforms, exchanges and supply platforms transact across company boundaries on blind-window counts.

A vertically integrated operator does not share that architecture's fatal flaw. In practice that means Google (DV360 buying through AdX into Google Ad Manager) and, in their own closed environments, the other walled gardens: Meta, Amazon and TikTok. An operator that owns the publisher-side ad server, the exchange and the buy side observes the post-render truth (page context, rendering, interaction) on the publisher side of its own stack, can feed it (lossily but meaningfully) into its own buying decisions, and can reconcile invalid traffic against billing inside one ledger, as automatic credits rather than inter-company disputes. Everything the open chain structurally cannot do, an integrated stack does internally as a matter of course. The timing constraint still binds at the moment of auction, but the settlement correction that the open chain cannot execute, one company's ledger executes trivially. Our scenario mapping confirms it from Google's own documentation: in the all-Google path, invalid-traffic removals and credits propagate to both sides automatically, with a residual gap of roughly 1-2%, while the open header-bidding path has no correction channel at all (see 6.3).

8.2 The drift, step by step

1. **T**rust in the open auction erodes as agentic share grows and DAMA-style discrepancies multiply: buyers see open-auction numbers they cannot reconcile and SIVT exposure they cannot quantify.

2. Demand migrates to where the ledger and the signals already live together. Buying inside an integrated stack offers one system of record, internal IVT crediting, and publisher-side signal access. Advertisers move not out of preference but out of self-protection.

3. Independent intermediaries become tenants. As independent paths thin out, exchanges and supply platforms face a rational temptation: plug into the dominant stack's server-side marketplace, where the integrated operator's publisher-side signals - the post-render evidence they could never obtain in the open auction - might partially solve their validation problem for them. They survive - inside someone else's auction, on someone else's terms, with someone else's visibility into their demand. Header bidding, the open web's great rebalancing achievement of the 2010s, loses its economic foundation.

4. Control exits the open internet. Pricing power, validation standards and the definition of a billable impression are then set inside one stack. Publishers negotiate from inside someone else's house. The open web does not go dark - it becomes a managed environment.

The irony is sharp: the independent open auction was built precisely to give publishers and buyers an alternative to platform dependence. An unresolved trust problem in that auction now risks delivering the market back to the platform - more completely than before, because this time the migration is justified as fraud-risk management. Architecture, not virtue, decides the winner.

> ***If the open chain cannot agree on what a valid impression is, someone else's chain will decide it for them.***

9. Standards will not save the open auction

2026 is the year the industry's standards bodies took the agentic agenda seriously. The work is necessary and welcome - and it is essential to be precise about what it does and does not address.

9.1 The current landscape

- Content monetisation standards (IAB Tech Lab's CoMP, v1.0 March 2026): a framework requiring AI systems to hold commercial agreements with publishers before crawling or content use - access control, licensing, settlement (IAB Tech Lab, 2026d). It governs how declared AI systems pay for content. It does not perform real-time agent identification, auction verification or IVT classification, and, as far as we are aware, none of the major AI players has publicly committed to it at the time of writing.

- Agentic transaction standards (IAB Tech Lab's Agentic Roadmap and RTB framework, January 2026; related agentic protocols demonstrated live in May 2026): reference implementations and frameworks letting buying and selling agents transact within and alongside existing infrastructure (IAB Tech Lab, 2026a, 2026b, 2026c).

- Open agent-to-agent protocols (an industry coalition's Ad Context Protocol, launched October 2025, building on MCP and A2A): standardising how buying and selling agents discover inventory, negotiate terms and execute campaigns - often outside the classic bidstream, asynchronously, with human-in-the-loop approval (AgenticAdvertising.org, n.d.).

Notably, the publisher side's posture is already shifting from blocklists to allowlists - from "block the known bad" to "admit only the known good" - a remarkable philosophical U-turn for the open web, and an implicit admission that detection has lost.

9.2 The shared limitation

Every initiative on the table works for agents that choose to cooperate. None of them identifies agents that do not declare themselves - and none of them repairs the open auction's settlement architecture.

Standards can govern the cooperative future. They cannot retrofit truth into a blind window, and they cannot make corrections binding across a chain of independent ledgers. The significant detail is rather that the new agentic protocols move transactions closer to the publisher - tacitly conceding where the truth lives.

10. Objections - the strongest case against this paper

A diagnosis this stark invites pushback, and the pushback deserves to be answered at full strength rather than in caricature. Below, the four strongest objections we expect from the intermediary side of the market - stated as their proponents would state them - and our responses.

10.1 “We own both sides of our pipe - our DSP and SSP reconcile internally”

The objection: an intermediary operating both a demand and a supply platform has a tight internal feedback loop: it sees both ends of its transactions, feeds findings from one side into the other, and adjusts bidding behaviour daily.

The response: owning two links of the chain narrows the inter-company settlement problem but does not touch the evidence problem. Both links still operate on bid-request-era signals; neither observes the post-render behavioural reality, which is generated on the publisher’s page and never enters the pipe. Internal reconciliation of two starved datasets produces a consistent number - not a true one. The DAMA case is the counterexample in production: internally consistent downstream counts running up to 2.3 times above the ad server’s post-render classification. Our own controlled experiment reproduces the failure end to end: wholly synthetic impressions were bought and billed as human, and the one correction the ad server later made reached no other ledger in the chain (Chapter 12).

10.2 “Scale beats you - we see patterns across billions of impressions”

The objection: large intermediaries process billions of requests daily across thousands of publishers; cross-network pattern detection catches what any single publisher - or any national coalition - could never see in isolation.

The response: scale is decisive against unsophisticated automation, and we expect intermediaries to keep winning that fight. But cross-network scale aggregates the signals the network carries - and the network carries bid requests. Against an agent whose every request-time field reads as a known human on a residential connection, a billion starved observations are a billion copies of the same blindness. Scale multiplies evidence; it does not create evidence that was never collected. The empirical anchor stands: the best-resourced request-time defence measured to date caught 1 of 7 agents.

10.3 “The learning loop closes the gap over time”

The objection: post-bid analysis continuously retrains request-time models; detection rates improve quarter over quarter; the gap is temporary.

The response: three failures, in ascending order. The loop trains on the wrong data (Chapter 5). The adversary is not stationary - agents improve at imitation as fast as models improve at discrimination, and the agent’s asymptote is perfect indistinguishability inside the window. And a learning loop only works if its corrections are binding: as long as flagged traffic produces partial, negotiated, lagging adjustments rather than automatic full corrections, the loop’s own training labels are corrupted by the commercial incentive to under-flag. An unfalsifiable promise of future improvement is not an architecture.

10.4 “Advertisers will not pay for your friction”

The objection: post-hoc validated settlement means delayed reconciliation, make-goods and extended delivery. Performance buyers want this week’s inventory this week; they will choose cheap, immediate, probabilistic inventory over verified, slower inventory.

The response: this is the strongest objection, and it is a question about price discovery, not architecture. It holds only while buyers cannot see the difference between the two products. The moment the unverifiable share of open-auction inventory is quantified - which is precisely what our experiment in Chapter 12 has begun to measure - “cheap and immediate” becomes “discounted for a measured defect rate”, and effective cost per verified human contact becomes the comparable metric. Some buyers will still choose volume. The buyers funding quality journalism will not knowingly fund synthetic eyeballs at any CPM - once they can tell.

11. Could the ecosystem fix itself?

This paper deliberately stops short of declaring the problem unsolvable. Several fixes are genuinely conceivable, and a careful assessment requires taking each seriously before reaching a conclusion. The argument is not that no fix exists - it is that none of the available fixes is likely to arrive, and to scale, fast enough to keep pace with the evolution of agentic behaviour. We examine four.

11.1 A two-stage bid: signal now, confirm later

The most promising technical idea keeps the real-time auction but splits the decision in two. A first bid clears inside the 200-1400 ms window on the signals available; a second, asynchronous exchange - seconds or minutes later, once post-render behavioural evidence exists - confirms, adjusts or invalidates the impression and reconciles the price. In effect, the protocol would grow a deferred-settlement layer that the timing constraint of Chapter 3 cannot otherwise accommodate.

This is viable, and worth pursuing - indeed it points in the same direction as the principles in Chapter 13. But two obstacles are serious. First, it requires the entire chain to adopt a new settlement semantics in concert: every demand platform, supply platform and verification vendor would need to support provisional pricing and binding post-hoc revision, across company boundaries and contracts. That is a multi-year standards-and-integration effort, not a feature. Second, even a perfect two-stage protocol still depends on the second-stage evidence being available to the party that prices - which, for the open auction, means transmitting publisher-side behavioural signals downstream, the very thing the bid request was never built to carry. The idea relocates the problem closer to a solution; it does not dissolve it, and it does not arrive quickly.

11.2 Richer real-time signals

A narrower version: enrich the bid request itself - publisher-side risk scores, first-party behavioural baselines, session context - so the demand side has more to decide on inside the window. This helps at the margin and STEP Network recommends it (Chapter 13). But it runs into the asymptote of Chapter 4: an agent operating in the user's own browser produces a session that, to the publisher, increasingly resembles the user's own. A richer signal raises the bar; it does not close a gap that narrows as agents improve. It buys time, not resolution.

11.3 Regulation forcing agent self-identification

The most consequential possible fix is not technical but political. If legislators - reacting to the broader risks of an agentic internet, of which ad fraud is only one - required AI agents to declare themselves, the detection problem would partly dissolve: a declared agent is trivially filtered. This is a plausible medium-term outcome, and the threats posed by the new agentic reality make some form of regulatory response likely.

But three gaps remain even in the optimistic case. First, timing: legislation moves in years, agentic capability in quarters; the rule arrives after the damage curve has bent. Second, enforcement against the cooperative only: even if the largest Western model providers (the likes of Google, OpenAI and Anthropic) join an identification regime, that does not bind providers outside the regulating jurisdictions. Agent companies based in China, India or anywhere beyond reach have no obligation to comply, and a residential browser agent does not wear a flag. Third, the adversarial floor: the moment identification is mandatory, non-compliance becomes the definition of a bad actor, and bad actors are precisely the ones who will not self-identify. Regulation can clean up the compliant majority. It cannot identify the agent that has every incentive to stay hidden.

11.4 Wait and see

The final option is to do nothing structural - to assume the ecosystem, under commercial pressure, evolves its own remedy in time, as it has adapted to past challenges. We take this possibility seriously rather than dismissing it. Markets do route around friction, and the incentives to solve verifiable delivery are real on all sides.

But the bet is asymmetric. If the ecosystem self-corrects and a publisher has meanwhile built verifiable settlement, the publisher has lost little - they hold a better product. If the ecosystem does not self-correct in time and a publisher has waited, the demand has already migrated to the walled auction (Chapter 8) and the independence is gone. Given the pace of agentic evolution against the pace of cross-industry coordination, our assessment is that self-correction at sufficient speed is unlikely - and that the cost of being wrong about waiting is far higher than the cost of being wrong about acting.

None of these fixes is impossible. The claim is narrower and more defensible: each is either slow, partial, or dependent on the cooperation of actors who have every reason not to cooperate - and the agentic curve does not wait.

The rational posture for a publisher is therefore not to bet the business on the ecosystem rescuing itself, but to build the verifiable alternative in parallel - cheap insurance against a likely failure, and a better product if the rescue somehow arrives.

12. STEP Network’s experiment: measuring the blind window

The FP-Agent study measured agents against a general-purpose defence on an instrumented site. The open question for the advertising economy is sharper: when synthetic traffic hits real pages carrying a real programmatic stack, how much is caught by the links that settle the money - inside the auction window, or at any point before the invoices are final? No public, empirical answer existed. In June 2026, STEP Network ran the test.

12.1 Hypotheses

- H1: Synthetic browsing traffic generated with a real, headful browser on an ordinary end-user connection is not classified as invalid by the layers that settle the money, and is therefore bought, counted and billed as human.

- H2: Whatever invalid-traffic classification does occur happens at the publisher’s ad server - and arrives only in the deferred correction window, after the money has already moved.

- H3: Corrections made by the ad server after the fact do not propagate to downstream ledgers: the downstream counts - and the invoices resting on them - stand unchanged. This is the settlement inversion of Chapter 6, reproduced under controlled conditions.

12.2 Setup

Real browser, human-like behaviour, ordinary connection. Traffic was generated by browser automation (Playwright) driving a full, headful build of Google Chrome - not a headless or stripped-down browser - from a single physical machine on an ordinary Danish connection. One persistent browser profile per site accepted the consent wall once and accumulated cookies across sessions, so the traffic presented as a returning, consented visitor. Ten scripted flows imitated human reading: front-page entries, section dives, article click-through chains, on-site searches, long reads with irregular scroll rhythms and randomised dwell times, multi-tab browsing and a mobile-emulation variant. Flow order was reshuffled between sessions and the browser fully restarted every four to five flows, so the traffic read as many short, distinct visits. No two sessions were identical, and no ad was ever clicked.

Live pages, full production stack, closed demand loop. Rather than a synthetic test site, the traffic ran against three live Danish publisher sites in STEP Network's portfolio, with the owners' authorisation: Google Ad Manager as ad server and the sites' production header-bidding setup, with demand supplied through a fixed-price (DKK 50 CPM) private deal at a major European supply-side platform, bought by a dedicated campaign on the same vendor's demand-side platform and funded by our own budget. Every link - ad server, SSP, DSP - counted the same inventory, and no external advertiser could be billed.

Controlled ground truth, labelled before the auction. Every session injected a key-value label at document-start - before the page's own ad stack initialized - into both the ad server's page-level targeting and the header-bidding wrapper's first-party data, so the label landed on the first auction and every subsequent ad request. The label makes the synthetic traffic precisely segmentable in every layer's reporting and guarantees it is excluded from client KPIs and billing. Whatever the chain does to this segment, it does to inventory that is 100% synthetic by construction.

The two-pull design. Because the ad server's own numbers remain provisional for up to roughly 30 days (Google, n.d.-c, n.d.-g; Chapter 3), every report was pulled twice: once the day after the traffic ran (18-19 June 2026), and again four weeks later (15 July 2026), after deferred invalid-traffic analysis had had the full correction window. The difference between the two pulls measures deferred detection at each layer; the difference between layers measures whether corrections propagate.

12.3 Traffic and session design

Component	Configuration
Traffic engine	Browser automation (Playwright) driving real, headful Google Chrome on one physical machine
Behaviour	10 human-like flows: front page, section dive, breaking/live, click-through chain, search, returning visitor, long read, deep read, multi-tab, mobile emulation
Identity	Persistent profile per site; consent accepted once; cookies accumulate across sessions; fresh browser instance every 4-5 flows
Inventory	Three live Danish publisher sites (authorised), desktop and mobile ad units
Ground-truth label	Key-value injected at document-start into ad-server targeting and header-bidding first-party data, on every ad request
Demand path	Ad server → header bidding → SSP (fixed-CPM private deal) → DSP (dedicated test campaign, own budget)
Run window	17 June 2026; modest, paced volume with randomised timing

12.4 Measurement points

- Per link: ad requests counted by the ad server under the label; bid requests, bid responses, impressions and revenue at the SSP (deal level); impressions bought and spend at the DSP (campaign level) - for exactly the same traffic.
- Real-time detection: whether any link blocked, discounted or flagged the labelled traffic before or at settlement.
- Deferred detection: the change in each layer's numbers between the two pulls - how much of the correction window each layer actually used.
- Propagation: whether a correction made at one layer appeared in any other layer's figures - the controlled counterpart of the DAMA reconciliation gap.

12.5 Ethics and scope

The test was designed to impose no costs on third parties: authorised inventory only, a closed deal loop with STEP's own budget on both the buy and sell side, traffic labelled precisely so it can be excluded from client KPIs and never billed to any external advertiser, no interaction with any creative, and volumes paced far below anything that could disturb production systems. Partners whose demand could be touched were informed in advance. Results are anonymised at vendor level. Scope: display on the open web, on one SSP-DSP path; apps and CTV are out of scope, and the agent-matrix extension (hosted agents and consumer agentic browsers) is future work.

12.6 Results

The traffic ran on 17 June 2026. The table shows what each layer counted for the labelled segment - the day after, and after the 30-day correction window; the full pull-level exports are reproduced in Appendix B.

Everything that cleared the auction was billed. 3,114 wholly synthetic impressions were bought through the deal, counted as valid and settled: the advertiser side paid DKK 179 for them and the publisher side booked DKK 145: two prices for the same impressions, the spread being the chain's fees, with no invalid-traffic deduction anywhere between. Neither the SSP nor the DSP blocked, discounted or flagged any of the traffic in real time, despite a signature (scripted automation, one machine, one connection, no ad interaction) far cruder than the browsing agents of Chapter 4.

Layer	Metric	Pull 1 (18–19 Jun)	Pull 2 (15 Jul)	Change
Ad server	Labelled impressions	19,962	19,573	-1.95%
SSP (deal)	Bid requests	37,047	37,047	0.00%
SSP (deal)	Bid responses	7,242	7,242	0.00%
SSP (deal)	Impressions counted and billed	3,114	3,114	0.00%
SSP (deal)	Publisher-side revenue	DKK 144.80	DKK 144.80	0.00%
DSP	Impressions bought	3,114	3,114	0.00%
DSP	Advertiser-side spend	DKK 179.06	DKK 179.06	0.00%

Everything that cleared the auction was billed. 3,114 wholly synthetic impressions were bought through the deal, counted as valid and settled: the advertiser side paid DKK 179 for them and the publisher side booked DKK 145: two prices for the same impressions, the spread being the chain's fees, with no invalid-traffic deduction anywhere between. Neither the SSP nor the DSP blocked, discounted or flagged any of the traffic in real time, despite a signature (scripted automation, one machine, one connection, no ad interaction) far cruder than the browsing agents of Chapter 4.

Deferred detection: the ad server restated 1.95%; the settling layers restated nothing. A month later, the ad server's deferred invalid-traffic analysis had removed 389 of the 19,962 labelled impressions - 1.95% - confirming on our own inventory that the ad server's numbers are provisional (Chapter 3), and that its deferred analysis does catch something the auction window missed. Every figure at the SSP and the DSP was identical to the digit: zero deferred correction, zero clawback, four weeks after traffic that was synthetic by construction.

The correction did not propagate. The one adjustment that did occur - at the ad server - reached no other ledger in the chain. The SSP's billed impressions, the publisher-side revenue and the advertiser-side spend all stood unchanged. The layer that detected corrected only its own reports; the layers that invoice never saw the correction. This is Chapter 6's inversion, reproduced end to end under controlled conditions.

Verdicts. H1 confirmed: *100% of the synthetic impressions that cleared the auction window were bought, counted and billed as human.* ***H2 confirmed:*** *the only invalid-traffic classification observed anywhere in the chain occurred at the ad server, inside the deferred correction window - after settlement.* ***H3 confirmed:*** *the ad server's correction propagated to no other ledger; every invoice-bearing count stood.*

Four caveats keep the result precise. First, scripted automation is the least sophisticated class in the FP-Agent taxonomy - which makes the result a conservative floor: traffic easier to catch than a modern browsing agent was billed end to end. Second, the layers count different units (the ad server metric here is ad impressions; the SSP and DSP settle on impressions), so the 1.95% is not directly a share of billed impressions. Third, any real-time filtering by the ad server is invisible in a two-pull design - the first pull was already net of it; the finding concerns what the settling layers did, which was nothing. Fourth, this is one SSP-DSP path at modest volume. None of these caveats weakens the propagation finding: whatever any layer removed, no other layer ever saw.

The blind window, measured: all surviving synthetic impressions billed; 1.95% restated a month later, at the one layer that does not send invoices; none of that correction reaching any ledger that does. Chapter 13 sets out what settlement must look like for detection to matter at all.

13. What a viable architecture must look like

This paper is deliberately a diagnosis, not a product pitch. But a diagnosis this structural implies design requirements for whatever comes next, and they can be stated plainly. Any settlement architecture for an agentic open web must satisfy five principles:

1. **Validation and invoice authority in the same boundary.** The party that observes the post-render truth must be the party - or stand inside the boundary - that issues and can correct the invoice. The inversion described in Chapter 6 is the disease; this is its negation.

2. **Settlement on verified human reach, not blind-window counts.** Real-time execution can and should remain - speed is not the enemy. But the billable unit must be what post-render evidence verifies, with delivery extended or made good until the contracted human reach is met. Pay for what was proven, not for what was counted in the dark.

3. **Corrections that are binding and automatic.** Invalid traffic identified post-hoc must adjust money and reporting by contract and by default - not through negotiation between independent ledgers. A correction that requires a dispute is not a correction; it is a write-off in progress.

4. **Behavioural evidence captured where it exists.** Publisher-side behavioural instrumentation - the only signal class with demonstrated discriminative power against agents - must become a first-class, auditable input to settlement, not an orphaned analytics layer.

5. **Collective scale for the long tail.** Behavioural baselines, detection infrastructure and negotiating leverage all require scale that individual publishers - especially the long tail of Chapter 7 - do not have alone. The defence must be built and operated collectively, or it will exist only for the head of the market.

Readers will notice what these principles quietly exclude: any architecture in which independent intermediaries settle on bid-request-era counts across company boundaries. They will also notice what the principles point towards - transaction models in which buying agents deal directly with the publisher side, execution remains automated, and settlement happens against the publisher's verified delivery. The emerging agentic protocols (Chapter 9) make such models technically practical for the first time.

STEP Network is actively developing an architecture along these lines together with Danish publishers & DAMA.

14. Conclusion: so - is this the death of open web programmatic?

Not of the protocol, and not of automation. OpenRTB will keep clearing impressions tomorrow morning, and real-time execution is worth keeping. What is dying is a specific architecture: the independent open auction - header bidding and the open web monetisation layer built on it - in which companies that never see the post-render truth settle binding invoices on counts made inside a 200-1400 millisecond blind window.

The diagnosis is structural, and it is three-fold. The auction closes before the evidence exists - a timing constraint that conceivable fixes (a two-stage bid, richer signals, regulation) can chip at but, as Chapter 11 argues, are unlikely to close at the pace agents evolve. The bid request cannot carry proof of humanity, and an agent in the user's own browser inherits everything the request can carry - so downstream filtering, real-time or "post-bid", analyses blindness retroactively. And the invoice chain runs against the truth: the most-informed party has no settlement authority, corrections cannot propagate backwards across independent ledgers, and the residual loss lands on the advertiser. The DAMA case shows the result in production; the growth rates of agentic traffic guarantee it compounds.

Standards will govern the cooperative agents and are welcome. They will not retrofit truth into the blind window. Vertically integrated stacks will reconcile internally and absorb the migrating demand - which is precisely the warning: if the open web cannot agree on what a valid impression is, a walled auction will decide it instead, and the independence that header bidding won in the 2010s will be surrendered in the name of fraud-risk management.

The way out is not to wait for better detection inside a constrained timeline, nor to bet the business on the ecosystem rescuing itself in time. It is to move the settlement boundary to where the evidence lives: execution in real time, validation post-render, invoicing on verified human reach, corrections binding by construction, operated collectively so the long tail is defended and not just the head. If the ecosystem does fix itself, a publisher who built this has lost nothing and holds a better product. If it does not (the likelier case), the publisher who waited has already lost the demand to the walled auction. The open web's answer to the agentic internet should be built by publishers, deliberately, now.

Open web programmatic was designed to trade human attention in real time. Attention is becoming agentic, and the truth about it arrives late.

Either the open web redesigns its settlement around that fact - or the only auctions left standing will be the ones somebody else owns.

The authors

Ulrik Kristensen
VP Product & Publishers

Jakob Peters
Senior AdTech Specialist

This paper was developed with the assistance of Claude, an AI system created by Anthropic. Claude was used to support research, drafting, and editing. All analysis, conclusions, and claims have been reviewed and are the responsibility of the authors.

STFP
network

Sources

Primary research and measurement data

Cloudflare. (2025, August 4). Perplexity is using stealth, undeclared crawlers to evade website no-crawl directives. Cloudflare Blog. https://blog.cloudflare.com/perplexity-is-using-stealth-undeclared-crawlers-to-evade-website-no-crawl-directives/

Cloudflare Radar. (2026, June 3). Automated vs. human traffic share [Data dashboard]. https://radar.cloudflare.com/

CNBC. (2025, October 23). Reddit accuses Perplexity of stealing user posts, expanding data rights battle with AI industry. https://www.cnbc.com/2025/10/23/reddit-user-data-battle-ai-industry-sues-perplexity-scraping-posts-openai-chatgpt-google-gemini-lawsuit.html

HUMAN Security. (n.d.). What is invalid traffic (IVT)? https://www.humansecurity.com/learn/topics/what-is-invalid-traffic/
HUMAN Security. (2026). 2026 state of AI traffic and cyberthreat benchmark report. https://www.humansecurity.com/learn/resources/2026-state-of-ai-traffic-cyberthreat-benchmarks/

Media Rating Council. (n.d.-a). Invalid traffic (IVT) detection and filtration guidelines: GIVT/SIVT definitions. Interactive Advertising Bureau. https://www.iab.com/guidelines/mrc-invalid-traffic-ivt-detection-and-filtration-guidelines-addendum/

Media Rating Council. (n.d.-b). MRC statement on pre-bid IVT requirements and processes. https://mediaratingcouncil.org/sites/default/files/News/MRC%20Statement%20on%20pre%20bid%20IVT%20requirements%20and%20processes.pdf

Media Rating Council. (2020, June 25). Invalid traffic detection and filtration guidelines addendum [Update]. https://mediaratingcouncil.org/sites/default/files/Standards/IVT%20Addendum%20Update%20062520.pdf

Prince, M. [@eastdakota]. (2026, June 3). Welp, that happened faster than I predicted. Thought it would be end of 2027, then early 2027, but agentic traffic [Post]. X. https://x.com/eastdakota/status/2062212701414187452

Wang, E., Shafiq, Z., & Vekaria, Y. (2026). FP-Agent: Fingerprinting AI browsing agents (arXiv:2605.01247). arXiv. https://arxiv.org/abs/2605.01247

Standards and agentic protocols

AgenticAdvertising.org. (n.d.). Ad Context Protocol (AdCP): Open standard for agentic ad trading. https://adcontextprotocol.org/

IAB Tech Lab. (2026a). AAMP 2.0 release brings transaction-ready buyer and seller agent SDKs. https://iabtechlab.com/aamp-2-0-release-brings-transaction-ready-buyer-and-seller-agent-sdks/

IAB Tech Lab. (2026b). Agentic advertising and AI [Standards portal]. https://iabtechlab.com/standards/agentic-advertising-and-ai/

IAB Tech Lab. (2026c). Agentic Real Time Framework (ARTF). https://iabtechlab.com/standards/artf/

IAB Tech Lab. (2026d). Content Monetization Protocol (CoMP) initiative, version 1.0. https://iabtechlab.com/standards/comp-content-monetization-protocols-initiative/

Google/IAB documentation

Google. (n.d.-a). Click and impression counting discrepancies. Campaign Manager 360 Help. https://support.google.com/campaignmanager/answer/2835377

Google. (n.d.-b). Counting impressions and clicks. Google Ad Manager Help. https://support.google.com/admanager/answer/2521337

Google. (n.d.-c). Difference between estimated and finalized revenue. Google Ad Manager Help. https://support.google.com/admanager/answer/2731724

Google. (n.d.-d). Investigate report discrepancies. Google Ad Manager Help. https://support.google.com/admanager/answer/6160380

Google. (n.d.-e). Open Bidding discrepancy resolution and impression reconciliation. Authorized Buyers Help. https://support.google.com/authorizedbuyers/answer/13531755

Google. (n.d.-f). Open Bidding payments. Google Ad Manager Help. https://support.google.com/admanager/answer/9752368

Google. (n.d.-g). Overview of Ad Manager reporting. Google Ad Manager Help. https://support.google.com/admanager/answer/2671992

Google. (n.d.-h). PG and Preferred Deals net revenue reporting. Google Ad Manager Help. https://support.google.com/admanager/answer/9771273

Google. (n.d.-i). Reporting discrepancies. Display & Video 360 Help. https://support.google.com/displayvideo/answer/2721749

IAB Tech Lab. (2016). OpenRTB API specification version 2.5. Interactive Advertising Bureau. https://www.iab.com/wp-content/uploads/2016/03/OpenRTB-API-Specification-Version-2-5-FINAL.pdf

IAB Tech Lab. (2022). OpenRTB 2.6 specification [nurl/burl, billable events]. https://iabtechlab.com/wp-content/uploads/2022/04/OpenRTB-2-6_FINAL.pdf

IAB Tech Lab. (n.d.). OpenRTB 2.x implementation guide. GitHub. https://github.com/InteractiveAdvertisingBureau/openrtb2.x/blob/main/implementation.md

Industry

Adform. (n.d.). Adform Help Center: Stats discrepancies and ad verification. https://www.adformhelp.com/hc/en-us

Databeat. (n.d.). GAM vs SSP discrepancies. https://databeat.io/knowledge-base/discrepancy-between-gam-vs-ssp-partners/

Index Exchange. (n.d.). Using burl to request billing notifications. Index Exchange Knowledge Base. https://kb.indexexchange.com/dsps/open-rtb/using_burl_to_request_billing_notifications.htm

Prebid.org. (n.d.). pbjs.bidderSettings: bidCpmAdjustment, netRevenue. https://docs.prebid.org/dev-docs/publisher-api-reference/bidderSettings.html

Sovrn. (n.d.). Gross vs net bids in Prebid. https://www.sovrn.com/blog/gross-net-publishers-losing-revenue-prebid/

STEP Network / DAMA (internal, primary)

DAMA Working Group. (2026). The DAMA case [Unpublished internal document]. Danish advertiser-and-media collaboration. https://dama.media/

STEP Network. (2026). Programmatic settlement scenario mapping [Unpublished internal analysis]. https://stepnetwork.dk/

STEP

network

About STEP Network

STEP Network is part of Jfm A/S and represents a broad network of Danish publishers. We work for an open, verifiable and commercially sustainable internet - from programmatic infrastructure to new, outcome-based trading models. Contact: stepnetwork.dk

Appendix A - the DAMA data (anonymised)

The table below is the anonymised line-level data behind the DAMA case. Publisher and advertiser identities are removed; counts and ratios are as measured. “Ad server” is the publisher’s ad server count (verified, net of invalid-traffic filtering); “creative layer” is the creative-hosting template’s own count (gross, taken at tag fire). Note the three clean lines at 1.00×: same format, different wiring, no gap - the discrepancy is architectural.

Setup	Format	Ad server impressions	Creative layer impressions	Impressions ratio
Publisher A	READ Wallpaper	35,041	64,474	1.84×
Publisher A	READ Wallpaper	34,768	64,079	1.84×
Publisher A	READ Wallpaper	34,802	64,232	1.85×
Publisher B	Wallpaper	51,159	51,217	1.00×
Publisher B	Wallpaper	51,508	51,572	1.00×
Publisher B	Wallpaper	51,394	51,433	1.00×
Publisher A	Wallpaper	51,259	94,875	1.85×
Publisher A	Wallpaper	51,498	95,124	1.85×
Publisher A	Wallpaper	51,208	95,089	1.86×
Publisher A	Wallpaper	32,257	44,684	1.39×
Publisher A (direct IO)	Wallpaper	3,263	7,629	2.34×
Total	-	**448,157**	**684,408**	**1.53×**

Ratios computed from the underlying campaign export; the full export is available for verification under NDA.

Appendix B - the experiment data (anonymised)

The tables below reproduce the pull-level data behind the controlled experiment in Chapter 12. Each report was exported twice from the respective platform's own reporting: Pull 1 on 18-19 June 2026, the day after the traffic ran, and Pull 2 on 15 July 2026, after the 30-day correction window. The traffic ran on 17 June 2026. The ad server report covers 17-18 June (Europe/Berlin); the SSP and DSP reports cover 17-19 June. Vendor identities are anonymised in line with 12.5; currency is DKK throughout.

B.1 Ad server (Google Ad Manager) - total ad requests by ground-truth label

Segment	Pull 1 (18 Jun)	Pull 2 (15 Jul)	Change
Labelled test traffic	19,962	19,573	-389 (-1.95%)

B.2 SSP - deal-level export (17-19 June 2026)

Metric	Pull 1 (19 Jun)	Pull 2 (15 Jul)	Change
Bid requests	37,047	37,047	0.00%
Bid responses	7,242	7,242	0.00%
Bid rate	19.55%	19.55%	0.00%
Win rate	43.00%	43.00%	0.00%
Impressions	3,114	3,114	0.00%
Buy rate	8.41%	8.41%	0.00%
Clicks / CTR	0 / 0.00%	0 / 0.00%	0.00%
eCPM	DKK 46.50	DKK 46.50	0.00%
Revenue	DKK 144.80	DKK 144.80	0.00%

B.3 DSP - campaign-level export (17-19 June 2026)

Metric	Pull 1 (19 Jun)	Pull 2 (15 Jul)	Change
Impressions bought (tracked ads)	3,114	3,114	0.00%
Win rate	42.97%	42.97%	0.00%
RTB cost (spend)	DKK 179.06	DKK 179.06	0.00%
Clicks / CTR	0 / 0.00%	0 / 0.00%	0.00%
eCPM	DKK 57.50	DKK 57.50	0.00%
eCPMV	DKK 81.84	DKK 81.84	0.00%

All 3,114 impressions were recorded on 17 June 2026 in both the SSP's and the DSP's daily breakdowns. The deal transacted at a fixed price of DKK 50 CPM; the DSP-side effective CPM (57.50) and the SSP-side effective CPM (46.50) bracket that price - the spread is the chain's fees, applied to counts that no layer ever adjusted for invalid traffic. The underlying exports are available for verification under NDA.