Artificial Poets



WHITE PAPER

You Can’t Get the Traffic Back. You Can Get the Session Back.

The economics of session depth, and the causal evidence behind it.

Executive summary

Media companies are losing audience to a change they did not cause and cannot reverse. AI-mediated discovery has pushed zero-click searches up roughly **60%**, cutting publisher traffic by about **26%**. [1] [2] Pages holding the top organic position can lose as much as **60%** of their traffic when an AI Overview appears above them. [3] CPMs fell **15–20%** over the same period, even as total digital ad investment grew. [4]

Most responses to this focus on replacing the lost traffic. This paper argues the opposite: the traffic is not coming back, and the remaining lever is what happens *after* a reader arrives.

The obstacle is structural. Most publishers run static content architectures organised by keyword and category, unable to read the intent behind an individual visit. What elite consumer platforms have done for a decade — predict what a person wants next and serve it immediately — has not been available to publishers at a workable cost.

The Artificial Poets Platform closes that gap, combining semantic understanding of content, anonymous interest profiles and real-time intent inference to decide what each reader sees next — without storing personally identifiable information.

Deployed on a live publisher property, the Platform produced a causally estimated **+53% increase in pageviews per user**, with a 95% credible interval of **[+48%, +58%]** and a posterior probability of a true effect above **99.9%**. Downstream engagement improved as a direct consequence. Deployment takes eight weeks.

+53%	>99.9%	8 weeks
Pageviews per user — 95% CI [+48%, +58%]	Posterior probability of a true causal effect	From engagement to measurable effect

What changed, and why the usual answer doesn't work

Publishers have spent two years watching search referrals fall. That conversation is exhausted, and most audience teams have concluded — correctly — that there is little they can do about the top of the funnel.

May 2024 marked the turn. Google AI Overviews began answering queries directly, interrupting the organic traffic flow that had sustained media businesses for two decades. [1] Zero-click searches rose sharply, [2] and a page holding the number-one position can now lose up to 60% of its traffic when an AI-generated answer appears above it. [3]

Referrals from AI assistants have grown — ChatGPT referrals to publishers rose **25%** year over year [1] — but they arrive at a fraction of the volume being lost.

Reported figures often understate the damage. Chartbeat has described traffic to large publishers as broadly stable, [5] but that figure depends on including Google Discover, which masks the decline in traditional search.

The commercial pressure compounds it. Publishers and ad-tech firms reported CPM declines of **15–20%**, concentrated in the fourth quarter, despite aggregate growth in 2024. [4] The

result is an environment where every remaining visitor is both scarcer and more valuable.

The second problem, which gets less attention

When a reader arrives from a search result that now competes with an AI-generated answer, they arrive with a narrower question and less appetite. The path that once produced three or four pageviews produces one or two.

Publishers are losing visitors *and* losing depth from the visitors who remain. The first loss is outside their control. The second is not.

Why personalization is the growth engine

The economic case is already demonstrated. Sophi, the system built by The Globe and Mail, produced a **222%** increase in user registrations and a **51%** increase in subscription conversions. [6]

Research on modern recommendation systems shows they materially increase engagement, [7] and that correctly applied personalization raises revenue per user by **5–15%**, reaching **25%** in the strongest cases. [8] Related work finds that when analytics genuinely inform editorial decisions, publishers see measurable gains in both engagement and monetization. [9]

Maximising revenue, however, requires more than technique. The experience has to feel relevant and uninvasive rather than mechanical. [10] A reader who notices they are being optimised is a reader you have lost.

How the Platform works

The Artificial Poets Platform does not bring more traffic. It makes each visit more valuable, which decouples revenue growth from traffic volume. It operates in four stages.

- **Content vectorization** — Language models and embeddings capture the intent, context and nuance of both content and reader queries, enabling semantic matching that goes well beyond keywords.
- **User fingerprinting** — A granular, privacy-preserving interest profile is built and continuously updated in real time, without storing personally identifiable information. Compatible with GDPR and CCPA.
- **Intent matching** — AI models infer the purpose behind a visit — informational, navigational, transactional or commercial — from behavioural signals, trend movement and micro-moments of intent.
- **Content orchestration** — The system matches and prioritises editorially relevant assets dynamically, optimising for engagement, reader satisfaction and monetization together rather than trading one against the others.

There is no redesign, no migration, and no change to how a site is authored. The reader does not click anything. They keep scrolling, and there is something else there.

The value cascade

The primary mechanism is pageview growth per user. Everything else follows from it. The table below is an **illustrative model**, not measured customer data. It shows how the mechanism translates at a hypothetical one-million-user scale, at an assumed $10 CPM.

Metric	Baseline	Post-Platform	Change
Monthly active users	1.00 M	1.00 M	0% — stable base
Pageviews generated	1.25 M	1.48 M	+18.4%
Estimated ad revenue	$10,000	$13,800	+38.0%

Because the model is multiplicative, a publisher with higher-value inventory sees proportionally more. The figure that matters is the pageview lift — the revenue column is arithmetic applied to it.

Beyond the revenue line

Improvements appear across the digital-experience metrics editorial and product teams already track: bounce rate falls as initial content relevance improves, pages per session rise as journey depth increases, and session duration extends as attention is sustained by dynamically prioritised content.

Two organisational benefits arrive alongside them. **Data-driven editorial intelligence** — insight replaces guesswork, and editorial teams act on real signals rather than intuition. And **operational efficiency** — automated analytics reduce manual effort in curation, tagging and prioritisation.

The evidence

Causal impact on pageviews per user

A **Bayesian causal impact model** estimated the isolated effect of deployment by contrasting observed outcomes against a counterfactual scenario in which the intervention did not occur. The analysis used **pageviews per user** rather than raw pageviews, which controls for traffic volume and isolates genuine engagement change from ordinary fluctuation in visits.

Pageviews per user per day	Value
Expected (modelled counterfactual)	1.38
Observed	2.11
Effect	+53%

How the effect is computed. The effect is the gap between what happened and the modelled counterfactual; the relative lift is that gap over the counterfactual.

$$\hat{\tau} \;=\; Y - \hat{Y} \;=\; 2.11 - 1.38 \;=\; \mathbf{+0.73}\ \text{pageviews per user per day}$$

$$\hat{\tau} \,/\, \hat{Y} \;=\; 0.73 \,/\, 1.38 \;=\; \mathbf{+53\%}$$

The credible interval on the relative figure is wider than the ratio of the absolute bounds, because the counterfactual carries uncertainty of its own.

The uplift is statistically robust: an average increase of **+53%** in pageviews per user, and a posterior probability of a true causal effect above **99.9%**. The increase is highly unlikely to be random variation.

Practical relevance

A statistically significant result is not automatically a commercially meaningful one, so a **practical relevance threshold** was defined in advance — the minimum uplift required to justify deployment economically. Combining the model's posterior effect trajectories with actual daily user counts yields the full posterior distribution of the causal effect, and with it the probability that the effect clears that economic bar.

It does, with 99.99% certainty, across every scenario evaluated. The measured uplift exceeds the profitability threshold by roughly seven times at the lower bound of the credible interval.

The cascade is causal, not correlational

A hierarchical Bayesian framework with full posterior uncertainty propagation was used to estimate the Platform's indirect effect on secondary engagement metrics — effects **entirely mediated** by its impact on pageviews per user. The **natural indirect effect** for each metric was estimated by Monte Carlo simulation, quantifying the causal lift attributable exclusively to that indirect route.

Metric	Mean indirect effect	95% credible interval
Bounce rate	−10.6 pp	[−13.9, −6.8] pp
Time on page	+38 s	[+33.8, +42.4] s
Session duration	+36.3 s	[+30.0, +43.7] s
Pageviews per session	+0.48	[+0.44, +0.52]

This is the part that separates the result from a correlation. The secondary metrics did not merely move alongside the deployment — they moved *because* pageviews per user moved, and the model quantifies how much of each is attributable to that path.

Implementation

Eight weeks from authorisation to measurable effect. The publisher's contribution is access and a decision.

Authorisation to proceed, CMS access, and infrastructure access. That is the list. The Platform is a JavaScript integration. The specifics vary by content management system, but nothing is rebuilt, migrated or re-authored. Integration is carried out by our team using documented standard implementations and is included in the pilot — the publisher's engineering team is not asked for sprint capacity. Where a deployment has genuinely non-standard requirements, forward-deployed engineering is scoped separately.

Ad serving is handled the same way. The Platform integrates with Google Ad Manager or an equivalent stack, and can render additional inventory and manage refresh behaviour within it. Legal documentation is prepared per customer, against that customer's own requirements and contract terms.

- **Weeks 1–4 · Instrument and index.** The Platform goes live in collection mode on day zero and does two things at once. It builds the behavioural baseline against which every later claim will be measured, and it indexes the site's full content library into embeddings. Indexing runs in the background and does not affect site performance. No AI feature is serving during this period. Readers see no change.
- **Week 4 · Enablement.** AI features are switched on, configured against the data collected rather than against defaults. This is the intervention date, and it is the reason the effect can be isolated afterwards.
- **Weeks 5–8 · Tune.** The models continue adjusting through reinforcement learning against live outcomes. Effect becomes measurable within this window.
- **Ongoing.** Results are available through a dashboard covering engagement, session depth, content performance and infrastructure health.

Why the baseline comes first

Most vendors switch on and start reporting. Four weeks of measurement *before* anything changes is what makes the result attributable when it arrives — it is the difference between a number that went up and a number that went up because of something. It also means the enablement date is unambiguous. Every result in this paper, and in the case studies that accompany it, rests on knowing exactly when the intervention happened.

Setting expectations

- **You can start small, and we would recommend it.** A publisher running several titles can enable on a subset and hold the rest as a comparison group. One portfolio in our data did exactly that — two titles enabled, the rest unchanged — which is why the effect on those titles could be isolated at all. It costs a few months of delay on the titles held back, and it buys a counterfactual.
- **The first quarter is not the run rate.** Launch quarters overstate steady state, in this product as in every other. On one measured property the first-quarter effect settled to roughly a fifth of its peak over the following six months, and it is the settled figure that should carry a business case.
- **And it is reversible without residue.** On one property the Platform stopped serving for four weeks. Engagement returned to its pre-enablement baseline and recovered within a fortnight of serving resuming — no degradation, no relearning period, no lasting effect on the weeks either side.

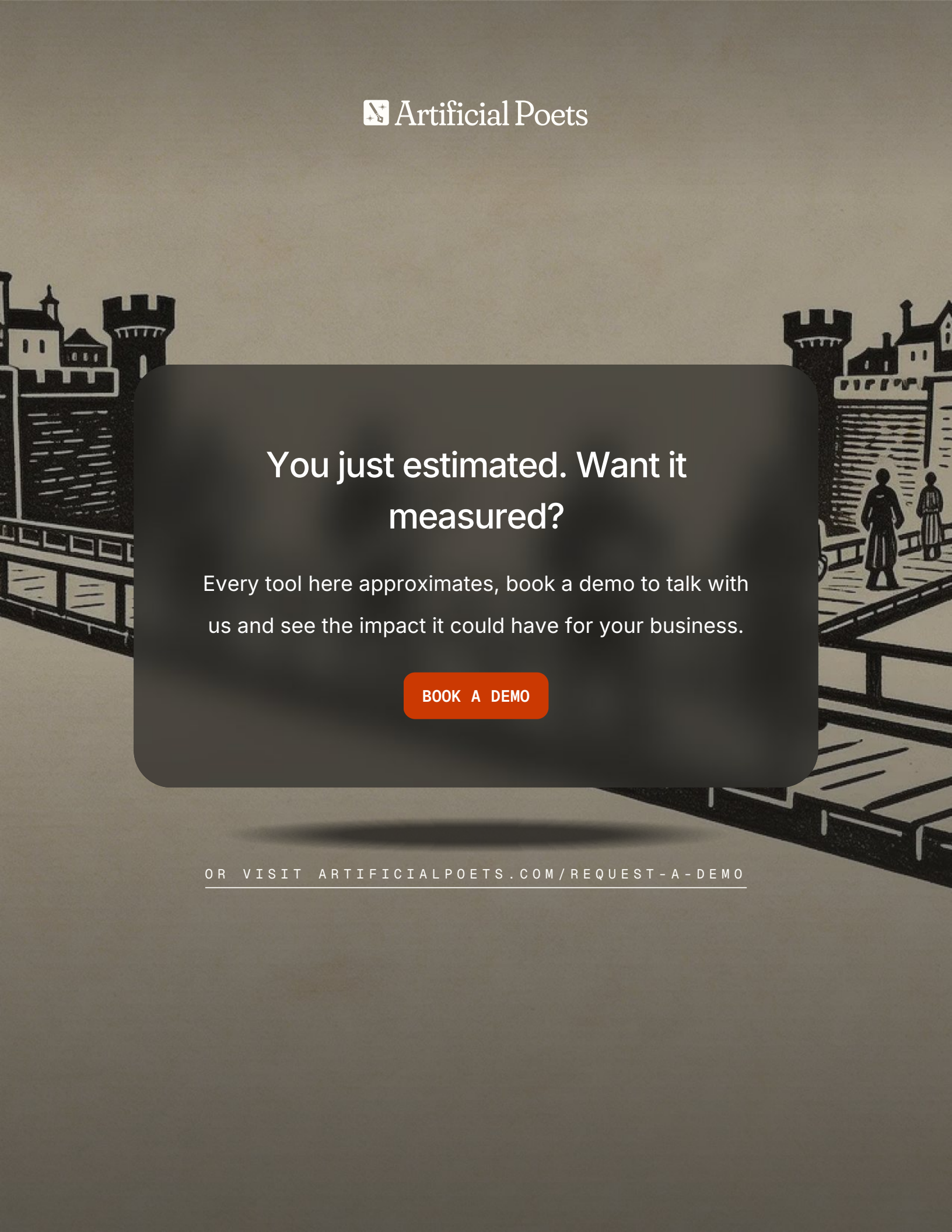
What this means

The new era of digital attention has redefined how audiences discover and consume content, and the consequences for media businesses are unambiguous: declining CPMs, sustained traffic loss, and a rising share of searches that end without a click.

Knowing precisely what each reader wants next is no longer a differentiator. It is a requirement. What was until recently exclusive to elite technology platforms is now available to publishers, supported by empirical and causal evidence rather than assertion.

You cannot make Google send you the visitors it used to. You can decide what happens to the ones who arrive.

Artificial Poets



You just estimated. Want it measured?

Every tool here approximates, book a demo to talk with us and see the impact it could have for your business.

BOOK A DEMO

OR VISIT ARTIFICIALPOETS.COM/REQUEST-A-DEMO

Appendix — modelling assumptions

The validity of the causal inference rests on four assumptions, stated plainly so they can be challenged.

- **Stable relationship.** The correlation between predictors and pageviews held consistently across the pre- and post-intervention periods.
- **No unobserved external shocks.** No major external event — traffic spike, platform change or seasonal anomaly — coincided with the intervention.
- **Accurate counterfactual.** The pre-intervention model captured baseline dynamics including seasonality and autoregressive structure (AR(1)).
- **Independent predictors.** Control variables were not themselves influenced by the intervention, which is what keeps the effect estimate unbiased.

References

1. Similarweb, “The Impact of Generative AI on Publishers” (2025).
2. D. Goodwin, “Nearly 60% of Google searches end without a click in 2024,” Search Engine Land (2024).
3. Pew Research Center, “Google users are less likely to click on links when an AI summary appears in the results” (July 2025) — the study behind The Guardian’s AI Overviews coverage.
4. IAB / PwC, “Internet Advertising Revenue Report: Full Year 2024” (April 2025).
5. J. Schwartz / Chartbeat data, “Publisher traffic sources 2019–2025 analysed,” Press Gazette (2025).
6. A. Hermida & F. M. Simon, “AI in the Newsroom: Lessons from the Adoption of The Globe and Mail’s Sophi,” Journalism Practice 19(10) (2025).
7. Liu et al. (ByteDance / TikTok), “Monolith: Real Time Recommendation System With Collisionless Embedding Table,” ORSUM @ ACM RecSys (2022).
8. McKinsey & Company, “The value of getting personalization right — or wrong — is multiplying” (November 2021).
9. M. K. Maurya & A. Kumar, “Audience Data and Editorial Decision-Making: Evolution and Applications of Web Analytics in Newsrooms,” Journal of Media Ethics 39(4) (2024).
10. J. Yin, X. Qiu & Y. Wang, “The Impact of AI-Personalized Recommendations on Clicking Intentions,” JTAER 20(1) (2025).

Results are from a live publisher deployment; the customer is not named. Property-level results are published separately as named case studies.