

From Search Traffic to Synthetic Circulation: AI Content Licensing and the Political Economy of Verification

The Economy Research Editorial^{1,2}

¹The Economy Research, 71 Lower Baggot Street, Dublin 2, Co. Dublin, D02 P593, Ireland

²Swiss Institute of Artificial Intelligence, Chaltenbodenstrasse 26, 8834 Schindellegi, Schwyz, Switzerland

Abstract

Generative search is weakening the economic bargain of the open web: publishers once provided indexable content and received referral traffic, advertising reach and direct audience relationships in return; the emerging system of AI-mediated synthesis absorbs and reproduces information without returning users to the original source. This article analyzes the structural shift from link-based discovery to synthetic circulation, through the Open Markets analyses to assess whether market mechanisms, including in the form of bilateral licensing arrangements, collective licensing mechanisms and interoperable provenance standards, can restore value to publishers. The analysis concludes that licensing is a necessary economic condition for sustainable synthetic circulation, but epistemically insufficient. Reliable reporting is not just a matter of providing credible texts but is dependent upon costly processes of sourcing, editing, correction and validation, legal review and editorial responsibility. Accordingly, the article distinguishes misinformation, disinformation and model hallucination, to illustrate why access to sources will not by itself be sufficient for reliable models and propose how a credible information ecosystem must compensate publishers while establishing enforceable access controls, transparent training practices, source attribution, interoperable provenance standards, editorial verification, competition oversight and investment in public-interest verification. The fundamental policy dilemma is not between protecting and innovating, but building AI systems that sustain digital information sources or systems that exploit their structural vulnerability.

1 Introduction — From Search Traffic to Synthetic Circulation

Claims that SEO is dying overstate the transformation. Search optimization remains relevant, but the economic exchange between visibility and referral traffic is weakening. Google's own guidance now explicitly considers visibility in the AI Overviews and AI Mode to be the same "search problem," while treating what the industry calls AEO or GEO as part of its existing search and quality systems.^[1] Rather, something more limited and more significant is being undermined: the previous traffic deal. In that deal, discoverability in search had conferred a reliable gateway to a consistent, monetizable source of demand, through clicks, subscriptions, advertising revenue and brand value. That deal sustained the economics of the open web for twenty years. It began to break between 2023 and 2026. Pew's 2025 behavioral survey recorded an 8% clickthrough rate among Google users who encountered AI summaries, versus 15% when no AI summaries appeared; just 1% clicked a source link within the summary itself.^[2] Meanwhile, the Reuters Institute survey of media leaders in early 2026 anticipated search-referral declines of 43% over the next three years, after Chartbeat already showed Google organic search traffic to over 2,500 sites in late 2024/early 2025 to have fallen a third worldwide.^[3]

That transition is probably best understood as a transition in circulation, rather than simply in interface design. Circulation in the print age depended on physical pathways, households subscribing to a paper and the ability to get the paper into a reader's hand. In the early age of the web, circulation was a matter of rank: a page needed neither a truck nor a newsstand, but simply had to appear in search and SEO was the design technology of choice for circulation. With the rise of the AI age, the same shift in circulation seems to be advancing another step: it is no longer a matter of whether a page appears in an index and ranks, but whether it appears at all in a trusted answer system that replies with an answer that may satisfy a user without a visit. According to the Reuters Institute, in 2026, only 10% of individuals in 45 markets used AI chatbots for news and only 1% used AI as their primary news source, but even at this early stage of adoption, the structural impact is already apparent, because answer engines increasingly absorb high-value informational queries and change users' expectations of what "search" should achieve.^[4]

It seems, then, that the familiar story of gradual migration from Google to questions to chatbots is less straightforward than it sounded. The piecemeal, overall movement is not yet evidence of the total displacement of traditional search by AI chatbots. The more pressing shift is more consequential: the search itself is becoming answer-shaped. The very same Google platform that mediated traffic onward to the web increasingly retains informational content within its own interface, while independent AI platforms extend that same paradigm into conversational interfaces. In the eyes of publishers and other knowledge producers, the end result is that traffic may diminish before the number of information seekers declines.

This is the environment that AEO and GEO are emerging in. These terms are themselves vague and Google even warns against treating them as a separate technical discipline.^[5] However, the general ambition behind them still reflects a real strategic shift. Rather than aiming simply to top a query, Google's own generative search guidance seeks to turn a website into a source that AI systems can identify and cite. Its top-level guidance for website content recommends "unique, valuable, non-commodity content," a clear structure, a first-hand experience and people-first writing. More detailed guidance for website content highlights original reporting,

real added value, clear authorship, clear expertise and the E-E-A-T framework, with trust emphasized as the key factor.^[6] The implications for authors and publishers are comparatively simple. Content designed primarily to secure ranking advantages is less defensible than material with evidence of experience, editorial oversight and reputational accountability.

That does not mean content producers no longer need traffic. For certain types of commercial players, brand presence in AI systems may, in some ways, replace readership; Similarweb has suggested that AI referrals have the potential to drive conversions that make sense for certain transactional categories.^[7] However, journalism, public service analysis, local reporting and other costly knowledge goods cannot be supported by simply existing within someone else's answer layer. Artificial circulation without an economic model is not distribution in any meaningful sense; it is value extraction. The key policy challenge, then, isn't whether the internet has to change to accommodate AI search; it is that it will. It is whether, in a political economy powered by a growing share of answer-based discovery, the market can sustain an ecosystem of trustworthy original reporting. This is why the licensing debate is critical and why licensing alone won't suffice.

2 The AI Content Licensing Market: The Open Markets Argument and Its Limits

The Open Markets analysis are based on a compelling intuition. If the web is shifting from link-based referral to answer-based extraction, then the prior equilibrium under which publishers gave content away in return for traffic is beginning to disappear. The Open Markets report outlines a nascent licensing market, which its authors argue seeks to secure compensation for content used for model training and retrieval. They contend that this market is forming in conditions parallel to the earlier platform-journalism settlement in which publishers' bargaining leverage waned. Radsch and Montoya distills the argument into a more pointed warning: the firms that reduce publisher traffic are now carving out the space to set the terms of the replacement revenue, consolidating their control over both sides of the value chain. That diagnosis carries weight because it approaches AI not as a unique copyright circumstance but as a market-structure problem.^[8]

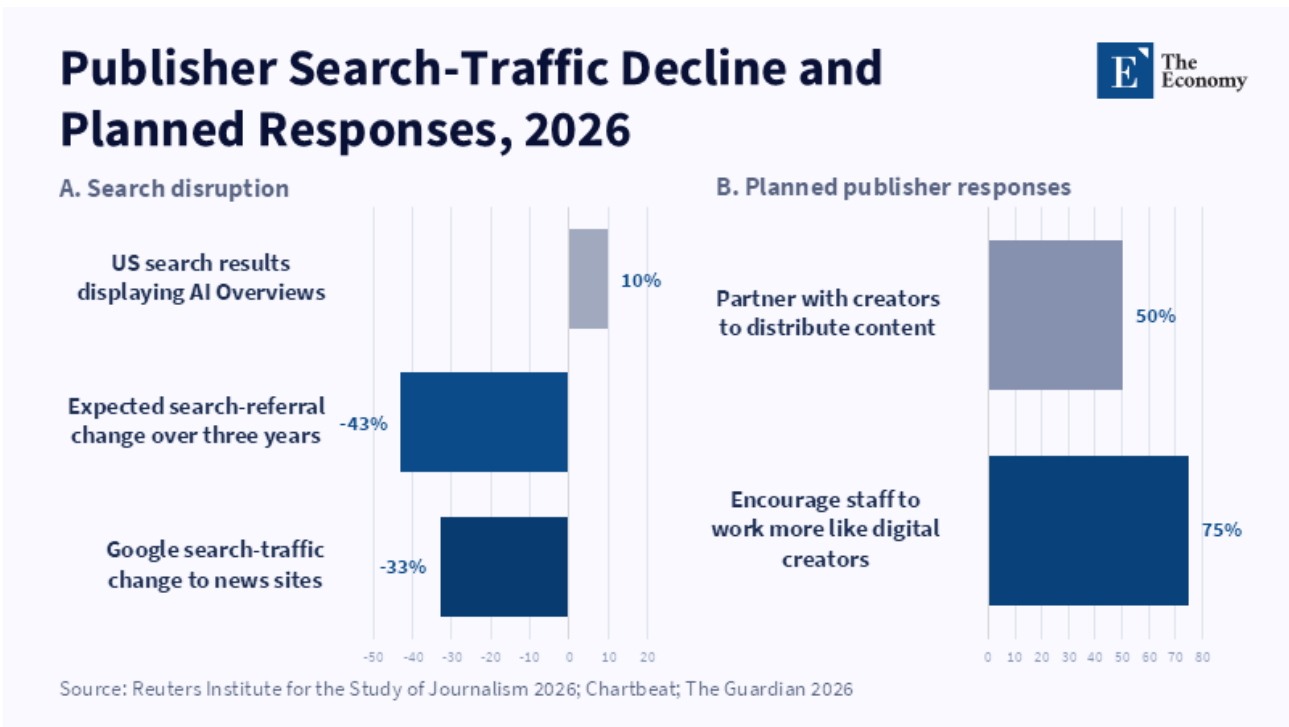


Figure 1: Search referrals are expected to contract further, accelerating publishers’ shift toward creator-led distribution and direct audience relationships.

The empirical imbalance is difficult to ignore. Open Markets reports that for each human referral OpenAI generated to publisher sites, its crawlers accessed approximately 1,700 pages; for Anthropic, that number was about 73,000:1, for Perplexity, it was about 369:1.^[9] Open Markets even suggests that AI zero-click activity represented just 0.04% of total referral traffic to publishers from other sites (compared to 85% from Google).^[10] Later, Cloudflare’s data pointed in the same direction. By mid-2025, training-related activity represented nearly 80% of AI bot crawling, while Google referrals to news sites in Cloudflare’s dataset had already fallen relative to January levels (March down by approximately 9%, April down by about 15%).^[11] That is, the extraction layer is growing faster than the traffic layer that is providing the compensation. That is the economic logic underlying the licensing claims.

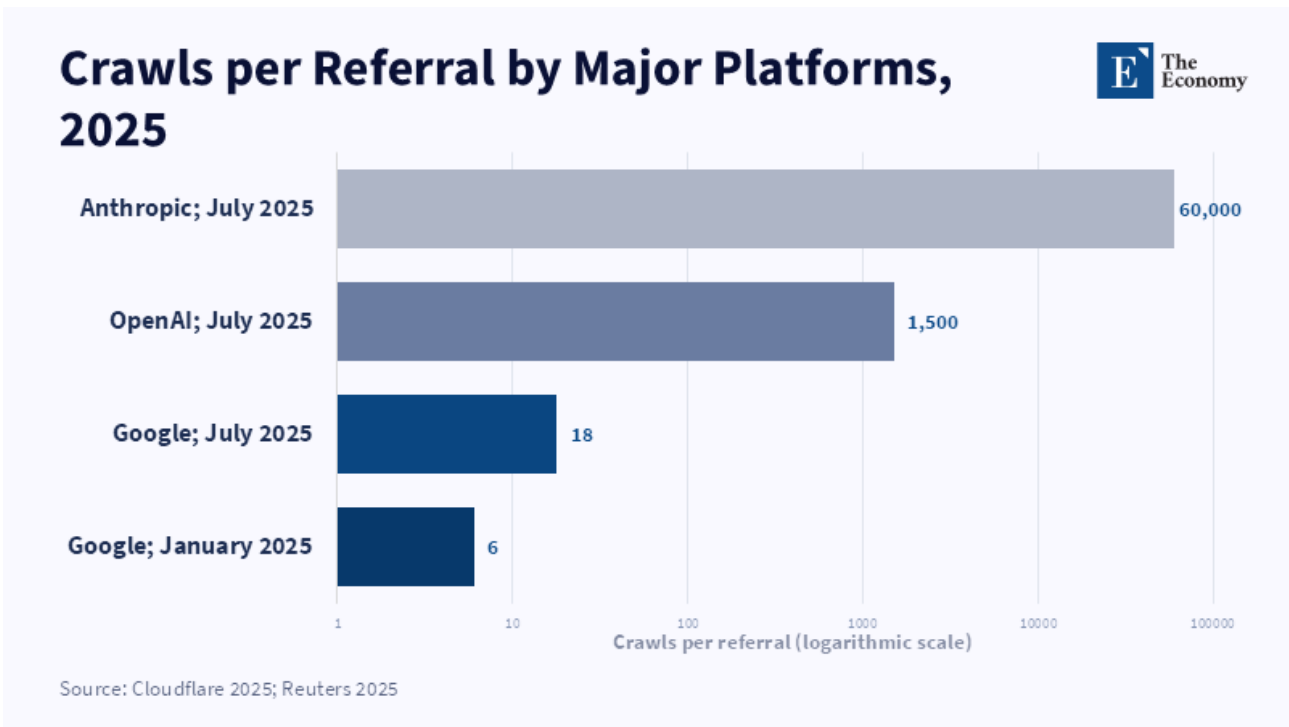


Figure 2: AI platforms extract vastly more pages than the referral traffic they return, weakening the exchange that previously supported publisher visibility.

Equally, a true licensing market is no longer utopian. According to research cited by the Reuters Institute, in October 2024, even the Institute noted that 26 or more international publishing groups had concluded licensing agreements with AI companies like OpenAI, Microsoft and Perplexity.^[12] By mid-2025, the market had expanded still further. The New York Times, while remaining involved in a separate set of lawsuits against OpenAI and Microsoft, had entered into a multi-year licensing agreement with Amazon to supply editorial content for use in AI products.^[13] These developments provide evidence that leading AI companies are willing to pay to access premium content on terms that make sense in certain circumstances. The real issue is not whether payment is feasible; it is whether the transactional arrangements, once established, remain resilient, transparent and scalable or whether they are unpredictable, opaque and driven by strategic self-interest.

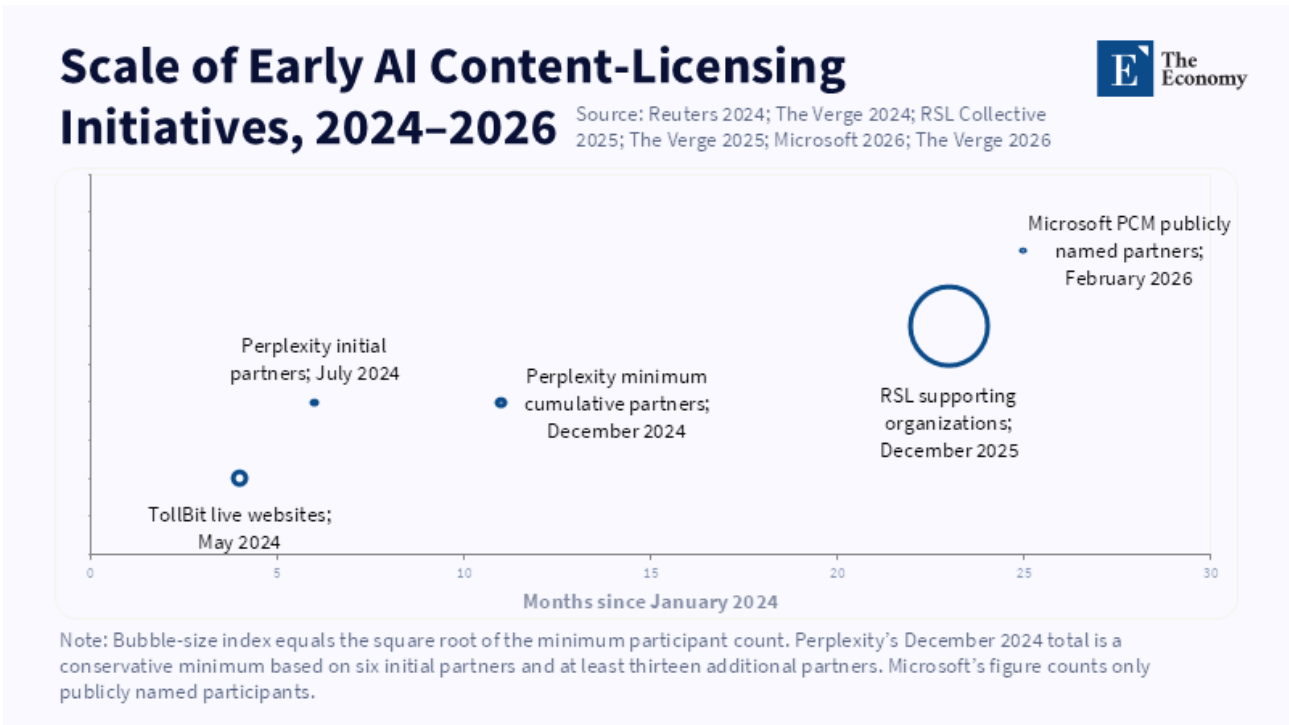


Figure 3: Content licensing is expanding from isolated publisher agreements toward shared standards and structured marketplaces.

The analysis is most persuasive at this point. It explains that licensing markets can be organized in different forms, from bilateral transactions to intermediary exchanges and protocol-level models such as pay-per-crawl or machine-readable rights reservation.^[14] It notes that bilateral negotiations tend to skew toward a handful of publishers with national reach, established brands, legal resources or strategic visibility and that auctions and other intermediaries with different take rates can create a wide range of results, with some reports showing that in comparable markets publishers can retain all or most of the proceeds and in other cases that intermediaries can keep a significant percentage, such as ProRata (50%) or Cloudflare (around 30%).^[15] The authors do not use these examples to establish a single preferred market design, but rather to warn that a two-sided market designed by platforms or adjacent companies can quickly lock into an alternative form of dependency, especially when the process is opaque to participants.

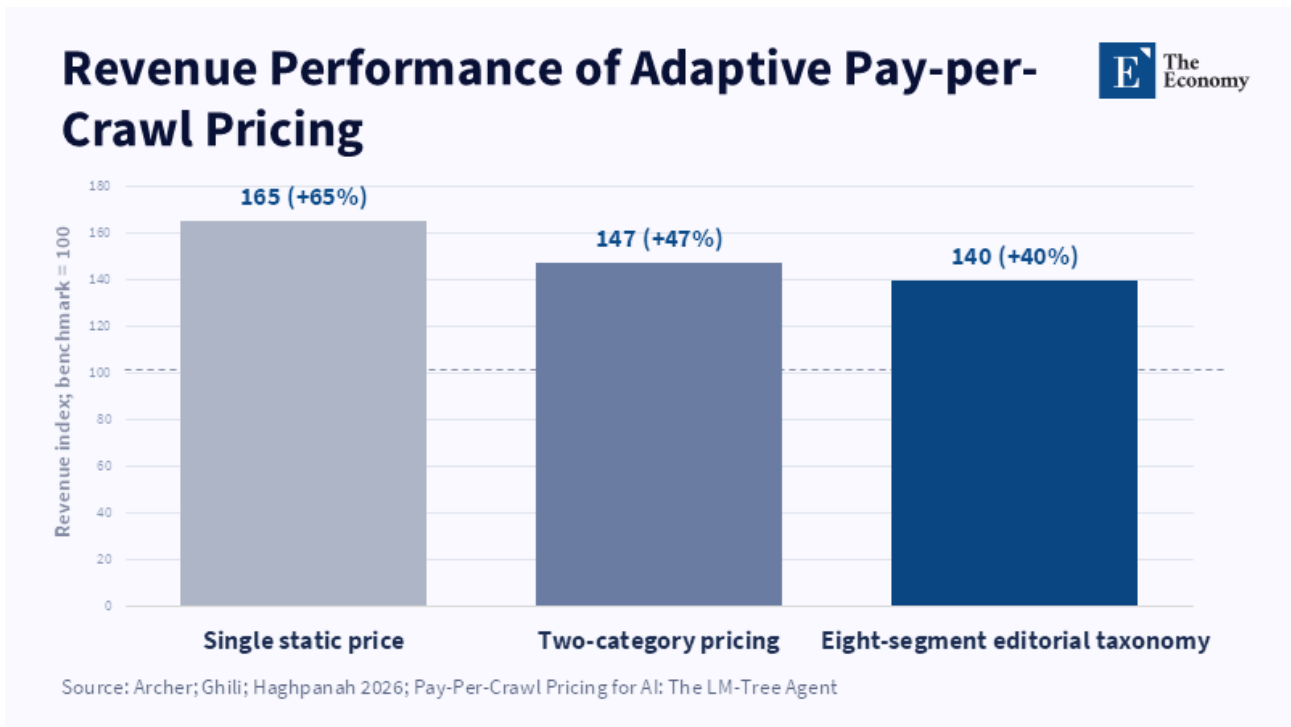


Figure 4: Adaptive pricing can outperform fixed licensing models, but the distribution of gains depends heavily on market design.

Nevertheless, the licensing thesis, compelling as it is, suffers from a necessary caveat. What AI companies are gravitating toward is not simply text itself, but something much more akin to well-vetted, up-to-date, accountable data. The Open Markets report is not just concerned with the proliferation of material; it also highlights temporal relevance and public trust, noting that models automatically acquire user confidence every time they anchor a response to a trusted news source.^[16] Conceptually, the market proposed is not just a market in licensed content; it is a market in provenance, timeliness and brand reliability. Consequently, describing it as akin to the print publisher who purchases hard-copy editions is only partially accurate: where a printed publication packaged reporting, editorial history, correction procedures, distribution and consumer confidence into a single commodity, AI systems are trying to separate those functions: to take in the sources, synthesize the information, inherit the associated credibility and externalize the costs of the institutions that make them reliable.

The comparative experience on this point is instructive even at this stage. Google’s own statements about the way that it ranks news are appropriately modest. Google states that Google News and its associated surfaces use technology to select items from reputable sources, expose users to multiple views and show the full evolution of a story. It explicitly states that the method it relies on the most is simply to reflect the ever-changing distribution of news with algorithms and to entrust the underlying editorial judgment to the original publishers.^[17] Full Coverage and Top Stories offer, therefore, a case of plausibility: if several authoritative publishers converge around a narrative, then the user is offered a limited indication that a statement is part of a broader visible consensus of reporting. That is not the same as the certification of truth. That is a substantial supplemental effort of judgment and institutional reputation that no ranking algorithm can deliver.

That distinction is important because the strongest form of the licensing argument can slide from compen-

sation to epistemology. Once the challenge is framed as AI simply needing reliable sources, a tempting but perilous subsequent step is to envisage a market in trusted arbiters, human or corporate agencies, who classify what sources can be used. It makes some sense in those circumstances, particularly in high-velocity fields with uneven sources. But it also risks excessive centralization. A business model that sells access to approved sources is likely to replicate the resource concentration that already exists on the Web: most revenue goes to the biggest corporations, smaller organizations are left out and the climb from reputation to financial benefit grows still steeper. Radsch and Montoya are correct to advise being wary of substitutes for the core mediating institutions. That risk is not just unfair distribution of fees, but that informational authority will be established by businesses providing the frameworks for access at scale.^[18]

The scale of payment from major LLM providers will also be influenced by law, competition and place. Payments that are voluntary are likely to be uneven because the benefits of high-quality content are often too diffuse and firms are generally able to free-ride except where forced to pay the true costs. This explains the importance of new regulations. The UK Competition and Markets Authority's (CMA) June 2026 conduct requirements afford publishers a new right within UK search services to prevent having their content used to power Google's AI search product, enforce proper attribution with clear links and bolster publishers' bargaining position vis-à-vis AI uses of their content.^[19] The EU's AI Act, on the other hand, made GPAI obligations applicable in August 2025 and linked transparency and copyright obligations to requirements for public summaries of training content.^[20] Likewise, the UK government's 2026 document on copyright and AI also anticipates that additional transparency is needed for rights holders to be able to assert their rights.^[21] Although none of these establishes a complete market, they indicate that the practice of uncompensated extraction is a politically contentious one.

Nevertheless, there is a cross-border limitation on enforcement. The UK House of Lords argued in 2026 that transparency should, in theory, be applied across all models entering the UK market, irrespective of where they were trained and acknowledged the restrictions of copyright law on overseeing training conducted entirely abroad.^[22] This conclusion gets to the core of the enforcement dilemma. A licensing-first regime binding only firms willing or required to participate can promote the best interests of the licensed or accessible markets, but cannot, on its own, stop foreign or non-transparent models from using content without compensation elsewhere. It is a universal problem of jurisdictional imbalance. If content regulation remains enforceable within sovereign boundaries while system deployment operates across borders, compliant firms will internalize the compliance costs that their non-compliant counterparts will not. In the end, any meaningful licensing regime requires not only contracts and platforms, but market-access rules, transparency mandates and compatible technical requirements deployed to all markets. Otherwise, licensing stays just a rational choice of select firms and select outputs.

So the right conclusion is not that the Brookings/Open Markets plan is exactly wrong, nor that it is enough. The right conclusion points to where the policy problem lies: the old referral deal is breaking; high-quality human reporting must not become a free upstream subsidy to synthetic answers. But the seriousness of the normative and economic burden now being placed on private licensing markets, absent strong competition policy, transparency obligations, opt-out options, attribution rules, rules for inclusion and exclusion and a collective publishing approach that includes smaller publishers, will turn this process into another form of platform

dependency. It will produce utility for some incumbents, but not a plural and trustworthy informational ecosystem.

3 Fake Content, Truth and the Economics of Verification

The web is not the origin of falsehoods and it will not be its terminus with AI. Rumor, propaganda, sensationalism and strategic deception all existed centuries before the digital age came along. What has changed from decade to decade is not so much the presence of lies but the institutional machinery through which claims to truth are organized and promoted. Print-era newspapers served as robust holders of reputational capital. The masthead concentrated reporting and editorial processes, editing and correction norms, legal liabilities and durable brand memories. The nascent internet unbundled those functions by collapsing logistics and deflecting discovery through search engines and social media. The outcome was not the removal of all institutional bases for truth but the attenuation of the financial underpinning underneath them. The Pew Research Center estimated that total US weekday newspaper circulation fell to approximately 20.9 million in 2022.^[23]

That shift in reach was important because truth remains costly: professional content production and circulation still require original reporting, editing, legal review and verification. Search and social media did not eliminate those costs; they merely altered how those costs were financed. During the search period, publishers could endure the mismatch a little longer because referral traffic was worth quite a lot. During the platform era, they persisted for a while longer until platform social referrals imploded. The Reuters Institute, drawing on Chartbeat data, concluded that aggregate Facebook traffic to news and media sites declined 67% in two years and traffic deriving from X plummeted 50%.^[24]

Even so, to speak as if you could buy this social good off the shelf, from a single source of known reliability, would be a grave mistake. Trustworthy media are not infallible and there is no institutional label that guarantees trustworthy reporting. According to the Reuters Institute's 2026 Digital News Report, just 37% said that they trusted most news most of the time, whereas a mere 20% trusted AI chatbots for news.^[25] The conclusion is stark: the public does not live in the world where one source can be designated within a framework of universal trust, but rather in a world where the authority of different institutions differs according to the values of different audiences, where even the most highly trusted can always be doubted and where confidence is a psychological effect garnered through ongoing processes of comparison, rather than a definitive judgment acquired once and for all. The social problem, therefore, remains, as it has always been, how to sustain procedures that generate credible reliance, not how to produce credibility where there is none.

That's why mechanisms like Google News do and should matter, but matter only so much. Google states that its news services aim to expose users to trusted sources, related coverage, additional context and multiple viewpoints.^[26] In addition, Top Stories, full coverage and fact-check links can help a user see whether something is being widely corroborated, debated, or treated as an outlier.^[27] That has some utility. It reduces search costs for plausibility checks and also provides a rough public indication of the position of an individual story in the reporting ecosystem. But it should not be mistaken for editorial verification. Google's own words remind us that algorithms are nothing more than a reflection of the news universe, not arbiters of its accuracy. A

mainstream or popular source could still be in error; a local or minority source might still be correct. Consensus visibility is a probabilistic indicator, not a judgment of truth.

This plural reality, in fact, finds expression in how the public behaves. The Reuters Institute 2025 research into how people attempted to verify items they suspected may be false revealed that trusted news sources, official sources such as government websites and search engines were the most common places respondents would turn, with most citing two or more potential sources. Fact-checking sites, albeit with significantly smaller audiences, still counted at moments of doubt. However, the research also found that 62% of respondents did not turn to publishers as the first port of call.^[28] This indicates both that journalism remains a central institution of verification and that its epistemic primacy can no longer be assumed. This illustrates how verification has become a distributed process involving other agencies, search engines and fact-checkers. Any policy framework that assumes a single marker of verification is therefore too narrow.

The publisher monetization dilemma needs to be understood in this context. A licensing market might help reward producers of trusted and time-sensitive information and in the best case, restore part of the value now captured by answer engines, but it cannot determine which sources are trustworthy. Functionally, the trustworthiness of sources is a cumulative institutional trait: transparency of authorship, correction standards, demonstrated expertise, sourcing discipline and institutional reputation. Google's own people-first content guideline echoes this reasoning. Original, valuable content along with transparency of authorship, expertise, source transparency, E-E-A-T and trust receive substantial emphasis.^[29] Accumulating the characteristics it highlights takes years of institutional cultivation. Trust is not a data field that can be added to search results. It is an organizational function.

That is why the proposition that "content creators will not need traffic anymore, as long as LLMs circulate the content for them" is only true conditionally.^[30] For some sorts of specialized marketing, software documentation, or discovery, inclusion in an answer engine might indeed deliver enough downstream value (by way of reputation or conversion) to compensate. Google has logged a fast rise in AI referrals, with some commercial building blocks trending upward.^[31] But journalism is different. A journalistic investigation, a town-hall story, a war dispatch, or a courtroom report does not build enough value simply by being captured in a chatbot output, when the publishing institution stands to shed the user relationship, the page view, the subscription path and the brand equity accessible through direct engagement. Artificial circulation that bypasses direct audience engagement disproportionately favors already prominent publishers while offering insufficient support to smaller, slower and more locally situated forms of reporting.

This then becomes an issue of inequity as well as revenue. As AI providers favor image-heavy, high-authority sources, given that they are safer to use, well-known and have benefits to reputation, the likely result is convergence. Open Markets warns that bilateral transactions tend to concentrate payment between publishers with established reputations and substantial legal resources, while leaving the rest of the information ecosystem out.^[32] The Lords likewise underlined any licensing-first scheme as something that had to be effective not just for the rights-holders but for the creators of different sizes in the different markets in which they operate, rather than relying on a single top-down marketplace scheme.^[33] Should this warning be disregarded, the answer-to-search transition will not just redistribute dollars, but also visibility, institutional weight and chances at survival

to the large incumbents, marginalizing local, niche, minority-linguistic and investigative operators who are often the sources of information that democracies depend on.

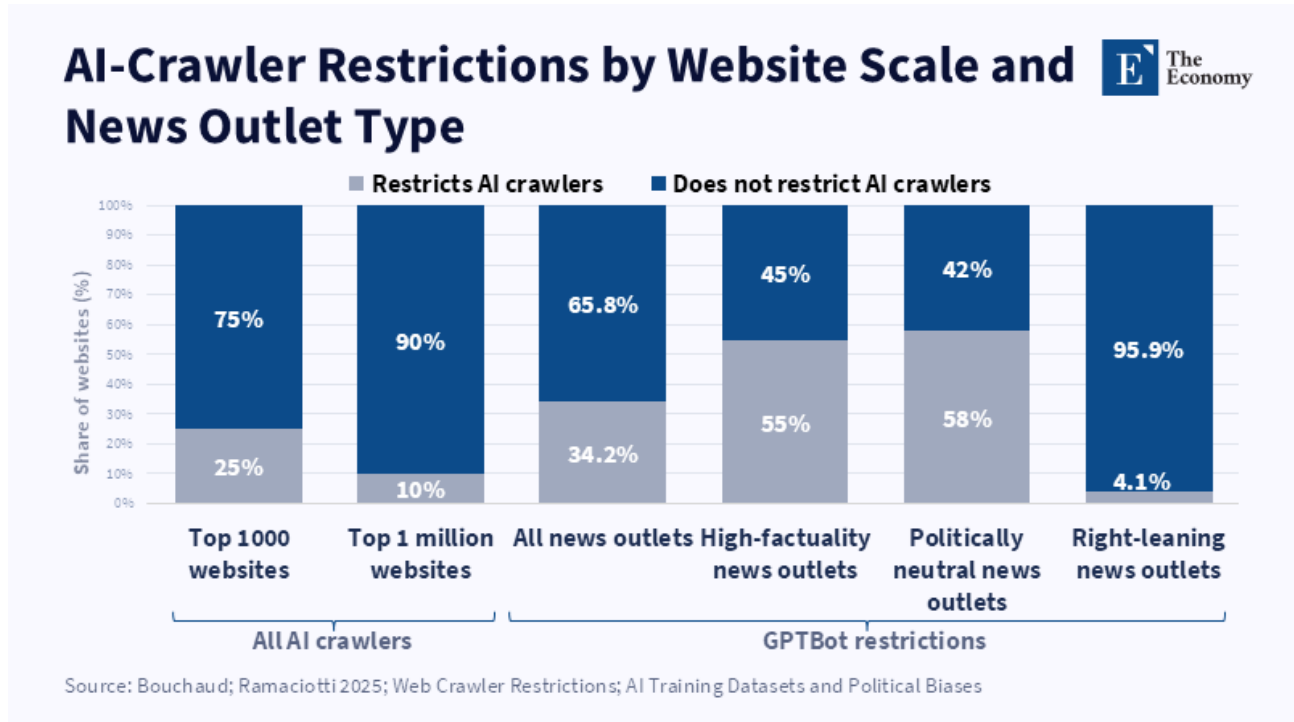


Figure 5: Higher-quality and politically neutral outlets restrict AI crawlers more frequently, potentially shifting model access toward less reliable sources.

The deeper insight is that truth is inseparable from the conditions of its construction. If any society wishes to have credible reporting in the age of AI, it will need to fund the institutions that can make warranted claims amid uncertainty. This will not be achieved by a moral endorsement of journalism, nor by a wistful longing for the hyperlink. It requires an economic acknowledgment that verification capability is a public good that bears private costs of provision. When one recognizes this, licensing is perhaps not the silver bullet, but merely one option in a broader settlement for safeguarding the material infrastructure of truth.

4 Discernment in the Age of AI: Provenance, Verification and Model Uncertainty

Any serious engagement with discernment during the AI age must clarify three problem domains that have been confused throughout the discussion. While not universally accepted terminology, misinformation has been used broadly to describe false or misleading content shared without malicious intent. Disinformation, in contrast, covers intentional falsehoods that distort beliefs or behaviors. Hallucination differs. As popularly used, hallucination covers plausible false outputs generated by models without deliberate human deception. UNESCO explicitly differentiates between disinformation and misinformation according to intent.^[34] Probabilistic generation and model-evaluation incentives are structural causes of hallucination, even when the underlying sources are reliable. Combining all three failures into a singular fake content category is ill-advised because the tools for mitigation are so different.

The scale of AI-enabled misinformation is already sufficiently large not to be complacent. NewsGuard's AI Tracking Center (updated in June 2026) identified thousands of AI content-farm news and information sites in 16 languages.^[35] It notes that these sites often operate with minimal or no human oversight, publish content at industrial speed and produce falsehoods about public health and safety, politics, brands and celebrities. The incentive structure is depressingly familiar: low-cost generative production plus programmatic advertising make cheap reach monetizable. UNESCO's survey of digital content creators in 2024 found that two-thirds did not fact-check before sharing, leading the organization to begin hosting training sessions on sourcing, verification and transparency.^[36] The unequivocal takeaway is simple. AI did not create the demand for cheap reach; it makes it extremely cheap to produce great quantities of convincing misinformation.

However, at the same time, reliable source licensing does not address hallucinations. A 2024 Nature paper found that existing methods for promoting truthfulness had only been partially successful, which the authors used as motivation for developing an uncertainty-based detection method, because 'confidently fabricated content persists'.^[37] An additional analysis published by OpenAI in 2025 incorporates the same line of reasoning to explain the incentives present in standard training and evaluation protocols, which 'reward more bold guesses over cautious guesses'.^[38] This becomes so important for governing the technology: even a model trained on the right sources can still invent connections, dates, names and attributions, in order to complete rather than abstain.

Evidence on news use reinforces these concerns. While the Reuters Institute 2026 evidence shows that AI chatbot news use remains a minority behavior, users who do use chatbots are doing more than seeking morning headlines.^[39] They are using them to ask experts follow-up questions, summarize complex topics, translate material, compare coverage and assess source credibility. AI systems aren't just an additional portal into existing journalism; they are an interpretive layer inserted between the user and the original source. Because that interpretation occurs inside the interface, the informational product users engage with is not the article but a synthetic integration of it. That further devalues checks for origin and attribution and further encourages users not to look underneath.

Click-through behavior reaffirms substitution pressure. In 27 markets, the Reuters Institute discovered that a mere 4% of respondents overall said they regularly or frequently would click through from an AI chatbot to the original source, compared with 19% from search and 17% from social media.^[40] Another investigation by Pew into US browsing habits, based on actual behavior rather than self-reporting, found that users who received an AI summary on Google clicked a news item link 8% of the time and clicked a link within the summary itself only 1% of the time.^[41] This data point shows why the economic and epistemic harm cannot be blunted by the presence of citations: attribution, which users rarely follow, still enables the answer layer to emerge as the primary site of consumption, while the authentic publisher becomes a back-end validator.

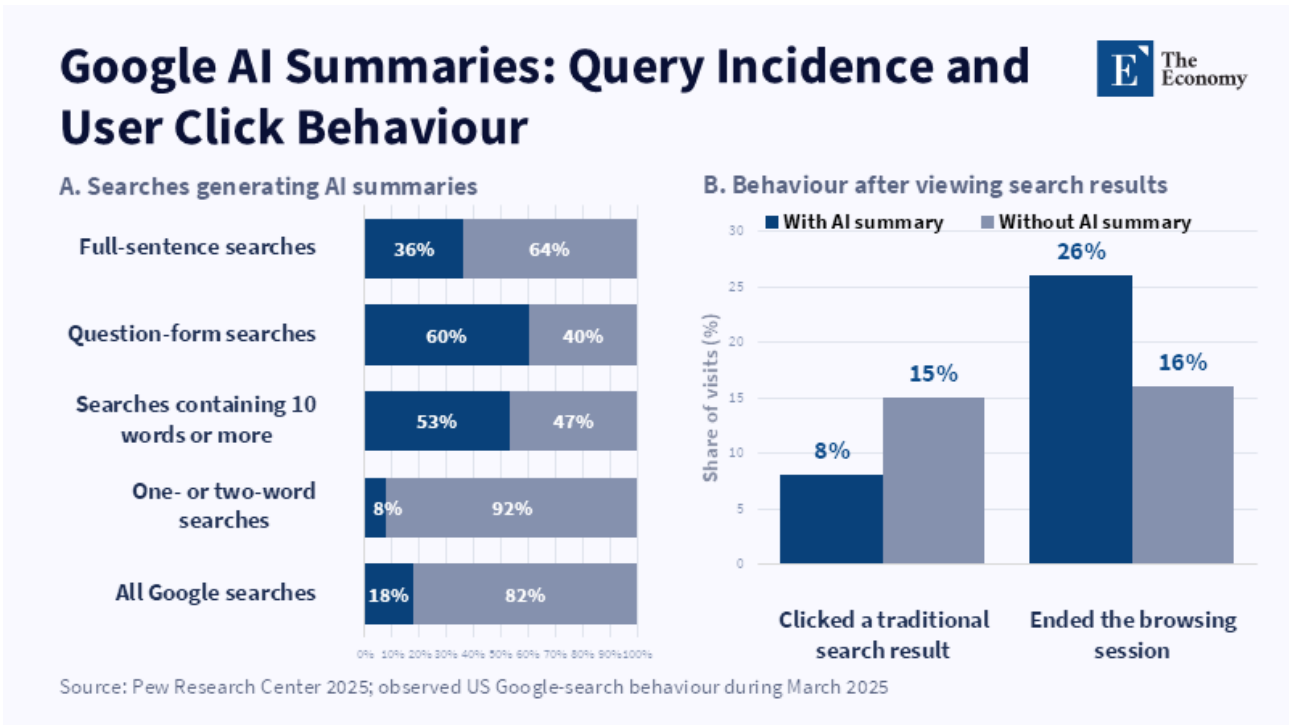


Figure 6: AI summaries appear most often for complex queries while reducing external clicks and increasing the likelihood that users end their search within Google.

There are also indicators that platform designers themselves are wary of the limits of an entirely synthetic account of current events. Reuters Institute observes that AI-generated summaries in search engines perhaps already influence clickthrough decisions on those queries and it also reports that Google’s AI Overviews are turned off on some breaking and developing new events, instead pulling up the first stories on the item in question.^[42] This is a crucial institutional indicator. None of this prevents the danger. All it does is reaffirm that on the very issues where public interest is concentrated, journalism is still irreplaceable.

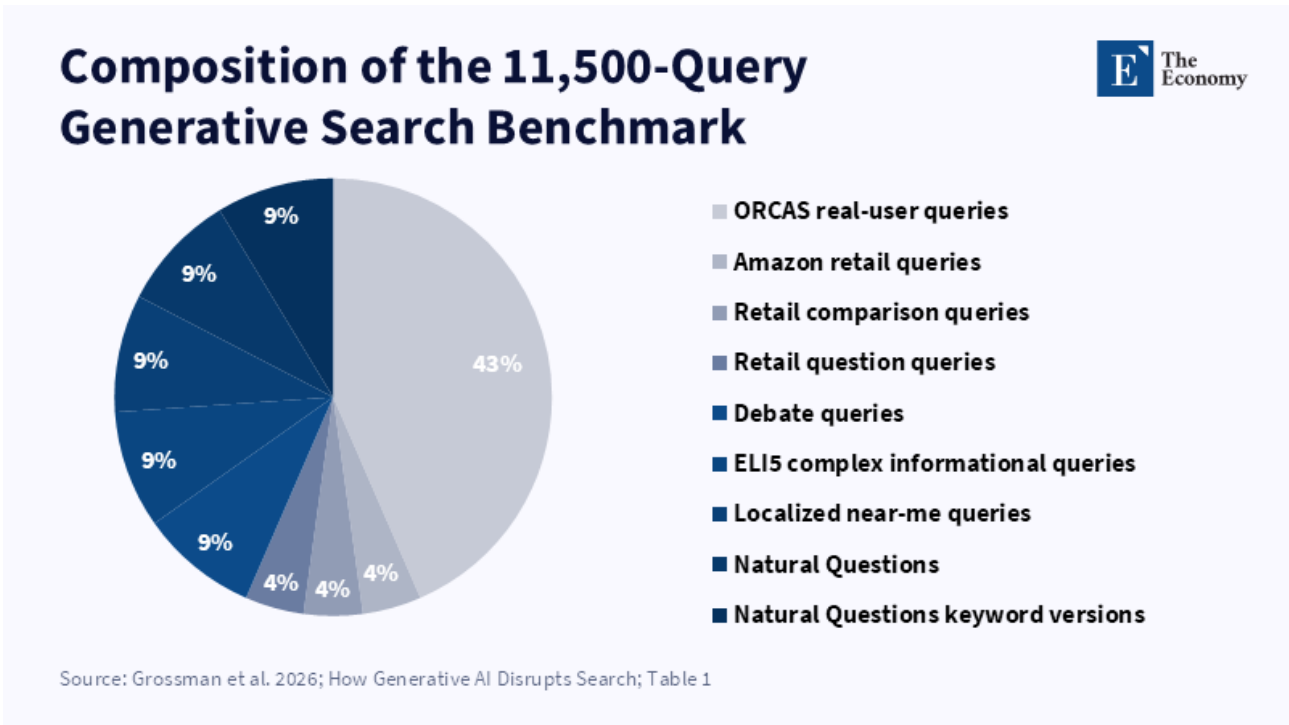


Figure 7: The benchmark spans real-user, commercial, political, local, and complex informational searches, providing a broad test of generative search behaviour.

Independent evaluations of the performance of AI news only intensify this warning. By October 2025, a European Broadcasting Union study involving 22 public-service media organizations across 18 countries and 14 languages analyzed more than 3,000 responses from leading AI assistants and found that 45% contained at least one significant issue, 33% had serious sourcing problems and 20% contained fundamental factual errors, including hallucinated data and obsolete information.^[43] Cross-lingual evidence of this type is important because it indicates this is not a problem that can be explained by incidental bugs in one model or the unique informational ecology of one nation. The pattern appears systemic. When users routinely use AI assistants as intermediaries for news, recurring sourcing weaknesses and factual errors are not minor flaws; they are structural risks within the informational environment.

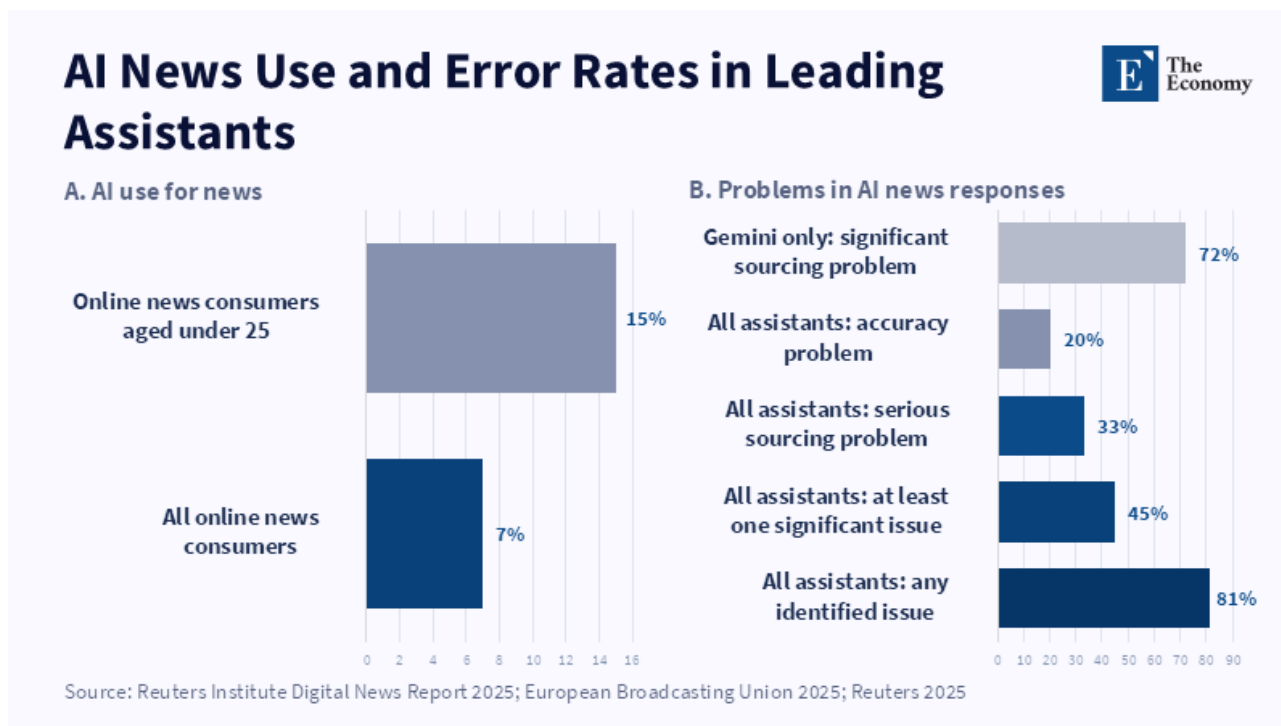


Figure 8: AI news use remains concentrated among younger audiences even as leading assistants continue to show substantial sourcing and accuracy failures.

The policy takeaway is that effective identification of false information cannot rely on retroactive fact-checking alone. Fact-checking is still critical, but it is inherently resource-intensive. An investigation published in 2025 of fact-checks used as data sources notes that the process of fact-checking is resource-intensive and a 2025 report from the Reuters Institute finds that fact-check brands have minimal general awareness, even when they matter in times of doubt.^[44] Human oversight must be targeted, adequately funded and institutionally reinforced. Not all queries, answers and misleading synthesized paraphrases can be effectively verified at scale, efficiently and inexpensively. That is precisely the reason why entirely reactive models of accountability are deficient. As both the magnitude of fabricated information and hallucinated paraphrases increase, so too will the downstream verification burden outpace the institutions that can support it.

What is required instead is a layered architecture of discernment. At the content layer, provenance and attribution need to be machine-readable and resistant to obfuscation. At the model layer, developers need transparency about training and grounding and a disincentive to guess where evidence is weak. At the interface layer, systems need to be clearer about when they are providing a direct quotation, when they are providing a summary, when they are making an inference and when they are synthesizing uncertain information. At the institutional layer, societies need better-funded public-interest verification institutions. The House of Lords in 2026 recommended exactly the sort of technical and legal scaffolding this would require: mandatory transparency about training data, open and interoperable standards for rights reservation, provenance and labeling and a licensing-first approach rather than a weak opt-out burden on creators.^[45] The EU AI Act builds on that rationale by mandating transparency and identifiability for AI-generated content, as well as attaching copyright obligations to GPAI providers.^[46] These do not solve discernment in their own right, but they begin to push policy toward an upstream model of accountability rather than a downstream scramble to clean up synthetic

misinformation after it has already occurred.

This upstream work is particularly critical for text. While both image and audio can deploy provenance, deepfake labeling and forensic tools effectively, even these are becoming subject to contestation. Text is more difficult. A good paragraph, even when properly rendered, can be copied, paraphrased, translated and re-contextualized with negligible signal degradation. That makes standardized fact-checking at scale particularly costly. Similarly, the EBU's Spotlight network, operating from 2025, which coordinates fact-checking and open-source intelligence capacity for public broadcasters, indicates the scale of institutional coordination required.^[47] Discernment in the age of AI depends less on any single automated classifier than on whether strong institutions can cooperate quickly enough to maintain verification as a social good.

On those grounds, the case for an AI content licensing market can be identified more narrowly. It would be justified if it subsidized the systems of production and verification required for credible synthetic retrieval; it would be unjustified if it simply commodified truth as a reputational premium sold off to a handful of dominant platforms, while magnifying the concentration and precarity of the underlying informational ecology. The best deal would therefore involve a combination of compensation and responsibility: enforceable control over access for AI retrieval and training, common standards for the reporting of crawl and referral data, visible attribution to original sources in synthetic answers, public summaries of training data, targeted regulatory disclosures where necessary and competition oversight to prevent platforms from acting as both market participant and rule-setter. These ought to be the same rules attached to every model in a market, rather than only those developed by domestic firms; otherwise, so-called compliant providers would be subsidizing non-compliant ones. That is not a panacea, but it is certainly the minimum needed to ensure that expertise and agency are not available only as a luxury for those with the time, staff and legal expertise to manage it.

5 Conclusion - Toward a Sustainable AI Content-Licensing Settlement

The current transition is often called the death of search engine optimization. But the fundamental shift is the rise of synthetic circulation: search is not disappearing. Instead, it is shifting to answers that encapsulate more and more information before the user even taps. This also recalls an enduring epistemic challenge: falsehood has always spread. What AI changes is the velocity, scale and presentation of mediation: fabricated content is cheaper to produce; even with reliable sources, hallucinatory outputs remain possible; reputable journalism is increasingly treated as feedstock.

Therefore, an AI-content licensing market is increasingly necessary. If answer engines depend on human-produced reporting for temporal accuracy, provenance and legitimacy, some mechanism must compensate the institutions that supply those conditions. But licensing will not work in a market dominated by private negotiations among major platforms and a small set of favored publishers. The evidence points toward a larger bargain: transparency on training data and system instructions, enforceable permissions for AI content use, accurate source attribution, interoperable provenance standards, protections to prevent monopolization and

public investments in verification capacity. The choice no longer lies between innovation and protection; it is between an AI system that supports the conditions of truth on which it depends, or one that depletes them until they weaken.

References

- [1, 5] Google Search Central (2025a) ‘AI features and your website’. Google.
- [2, 41] Pew Research Center (2025) ‘Google users are less likely to click on links when an AI summary appears in the results’, 22 July.
- [3] Newman, N. (2026) *Journalism, Media, and Technology Trends and Predictions 2026*. Oxford: Reuters Institute for the Study of Journalism.
- [4, 25, 39, 40, 42] Reuters Institute for the Study of Journalism (2026) *Digital News Report 2026*. Oxford: University of Oxford.
- [6, 29] Google Search Central (2025b) ‘Creating helpful, reliable, people-first content’. Google.
- [7, 31] Similarweb (2025) *2025 AI Traffic Report*. Similarweb.
- [8, 9, 10, 12, 14, 15, 16, 18, 30, 32] Radsch, C. and Montoya, K. (2026) *Same Gatekeepers, New Tollbooths: Mapping the AI Content Licensing Market*. Washington, DC: Center for Journalism and Liberty, Open Markets Institute.
- [11] Cloudflare (2025) ‘Content Independence Day: New tools for content creators to control their content’, 1 July.
- [13] Reuters (2025a) ‘New York Times partners with Amazon for first AI licensing deal’, 29 May.
- [17, 26, 27] Google (2025) ‘How Google News works’. Google.
- [19] Competition and Markets Authority (2026) *Conduct Requirements for Google Search Services*. London: Competition and Markets Authority.
- [20, 46] European Parliament and Council of the European Union (2024) ‘Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence’, *Official Journal of the European Union*, L 2024/1689.
- [21] Department for Science, Innovation and Technology and Intellectual Property Office (2026) *Copyright and Artificial Intelligence*. London: UK Government.
- [22, 33, 45] House of Lords Communications and Digital Committee (2026) *Copyright and Artificial Intelligence*. London: House of Lords.
- [23] Pew Research Center (2023) ‘Newspapers fact sheet’, 13 June.
- [24] Newman, N. (2025) *Journalism, Media, and Technology Trends and Predictions 2025*. Oxford: Reuters

Institute for the Study of Journalism.

- [28] Reuters Institute for the Study of Journalism (2025) *Digital News Report 2025*. Oxford: University of Oxford.
- [34] UNESCO (2018) *Journalism, 'Fake News' and Disinformation: A Handbook for Journalism Education and Training*. Paris: UNESCO.
- [35] NewsGuard (2026) 'AI Tracking Center'. New York: NewsGuard Technologies.
- [36] UNESCO (2024) *Behind the Screens: Insights from Digital Content Creators*. Paris: UNESCO.
- [37] Farquhar, S., Kossen, J., Kuhn, L. and Gal, Y. (2024) 'Detecting hallucinations in large language models using semantic entropy', *Nature*, 630, pp. 625–630.
- [38] Kalai, A.T., Nachum, O., Vempala, S.S. and Zhang, E. (2025) 'Why language models hallucinate', *arXiv preprint arXiv:2509.04664*.
- [43] European Broadcasting Union (2025a) *News Integrity in AI Assistants: An International Public-Service Media Study*. Geneva: European Broadcasting Union.
- [44] Altuncu, E., Başkent, C., Bhattacharjee, S., Li, S. and Roy, D. (2025) 'FACTors: A new dataset for studying the fact-checking ecosystem', *arXiv preprint arXiv:2505.09414*.
- [47] European Broadcasting Union (2025b) 'EBU launches Spotlight fact-checking network to combat misinformation and support trusted news', 11 April.