

News Media Association Response to the CMA Consultation - Google's general search services: proposed Conduct Requirements – Publisher Conduct Requirement

The News Media Association (“**NMA**”) is the voice of UK national, regional, and local news media in all their print and digital forms – a £4 billion sector read by more than 46.4 million adults every month. Our members publish around 900 news media titles, ranging from well-known national and international brands to independent local papers of record, including The Guardian, Financial Times, The Daily Telegraph and the Daily Mirror, to the Manchester Evening News, Kent Messenger, and the Monmouthshire Beacon.

1. Summary

The NMA welcomes the CMA's proposed Publisher Conduct Requirement (“**Publisher CR**”) as an crucial intervention to address clear market harms created by Google's Strategic Market Status (“**SMS**”) in general search. The CMA correctly identifies that publishers must allow Google to crawl their content for search visibility, and that this dependency is now being leveraged as that same content is used in search generative AI features (e.g. AI Overviews and AI Mode) and broader generative AI services (e.g. Gemini), with significant adverse effects on publisher monetisation and publishers' ability to invest in quality content.

The CMA's notes that clicks to publisher sites fall materially when an AI Overview is shown for relevant queries. The CMA also notes the resulting risks to publishers' brands, transparency for users, and longer-term investment in high-quality content - particularly news.

However, the Publisher CR will only achieve its aims if it is operationally enforceable, resistant to circumvention, and future-proof as generative AI systems evolve. The NMA therefore recommends strengthening the proposal in these core areas:

Effectiveness of controls and swift escalation in cases of non-compliance

The requirement for effective controls must be underpinned by verifiable implementation of compliance and robust enforcement. Where behavioural solutions do not deliver meaningful publisher choice in practice, the CMA should make clear that it will escalate swiftly to more robust remedies, including structural measures if necessary.

Granularity of publisher choice across AI use cases

Publishers must be able to exercise distinct, granular controls over economically different uses of their content, including grounding, training, and fine-tuning, and be able to do so at both directory and page level across all use cases. Bundled or imprecise controls would significantly weaken bargaining leverage and undermine the CR's objectives.

Fine-tuning as an economically meaningful use of content

Fine-tuning represents a distinct and increasingly important use of publisher content, with the potential to substitute for grounding over time. The absence of a dedicated control risks eroding the value of publisher controls and weakening the long-term effectiveness of the CR.

Preventing circumvention and data leakage

The CR must not be undermined by routes through which Google can reacquire opted-out publisher content, including via open source datasets. The Publisher CR must also prevent secondary source scraping, or onward sale of data to third parties by Google.

Transparency sufficient to support bargaining and verification

Transparency obligations must go beyond high-level metrics and provide publishers with the information necessary to assess commercial impact and negotiate effectively, and identify potential non-compliance. This includes per-feature reporting and metrics that provide meaningful contextual information about Google's use of – and user interactions with – publisher content.

Attribution that is prominent and encourages user engagement

Attribution is central to preserving publisher value and consumer trust. Requirements should ensure attribution is not only accurate, but prominent and designed to encourage user engagement with original sources.

Territorial effectiveness and international alignment

The CR should operate in a way that avoids territorial loopholes where UK publisher content is used in non-UK AI services. Greater clarity on territorial application, alongside international regulatory alignment, will support effectiveness and legal certainty.

Independent and ongoing compliance oversight

Given the technical complexity of AI systems and Google's information advantage, baseline auditing from an independent third-party alone is insufficient. Ongoing independent compliance audits are necessary to ensure that publisher choices are respected and that controls operate as intended.

Implementation timetable

A six-month implementation period risks unnecessarily delaying relief from well-evidenced harm. The core requirements of the Publisher CR build on existing systems – precisely because it will be easier for Google to implement - and could be implemented more rapidly.

Relationship with a fair and reasonable terms CR

Publisher control and remuneration are closely linked. While the Publisher CR may improve negotiating leverage, delaying consideration of a conduct requirement on fair and

reasonable terms risks leaving the underlying imbalance in bargaining power unresolved and weakening the overall policy framework envisaged by Parliament.

2. The NMA supports the CMA's theory of harm and analysis of consumer benefits

The CMA's diagnosis is well-founded: publishers' dependence on Google Search creates compelled consent to crawling. Where crawling is a prerequisite to reach audiences in Search, Google's decision to use crawled content to generate AI answers that reduce click-through directly impacts publishers' ability to monetise through advertising and subscriptions. The CMA's consultation expressly acknowledges these dynamics and the resulting risk that publishers cannot sustain investment in quality journalism.

The Publisher CR supports consumers by sustaining the production and availability of high-quality, trusted content. The CMA correctly identifies that reduced referral traffic from search generative AI features weakens publishers' ability to monetise content and therefore to invest in journalism and other forms of authoritative information. By restoring meaningful choice and bargaining power to publishers, the CR directly supports continued investment in original reporting, underpinned by robust editorial standards. Consumers will benefit from a richer, more reliable information environment and from the preservation of plural, independent sources of news rather than increased reliance on synthetic summaries alone.

The Publisher CR will also improve consumers' ability to verify and trust information encountered through Search. Stronger requirements on attribution - combined with publishers' ability to withdraw content if attribution is insufficient - create incentives for Google to design AI features that make sources intelligible to users. This enables consumers to distinguish authoritative reporting from lower-quality material. Over time, this supports digital literacy and trust, both of which are increasingly important as generative AI becomes a primary gateway to information.

In terms of the AI market, the Publisher CR will promote more effective competition and innovation in AI services, which benefits consumers through better product quality and choice. By reducing the extent to which Google can leverage its SMS in search to secure unique advantages in AI development, the CR contributes to a more level playing field for rival AI assistants. This dynamic competition encourages innovation in user experience, rather than entrenching a single dominant firm. Consumers will ultimately benefit from improved services and from competitive pressures.

These effects will reinforce one another. Sustained investment in quality content improves the underlying information ecosystem on which AI systems rely; better attribution and transparency improve user trust; and stronger competition ensures that these gains are not eroded over time by unilateral platform incentives. For these reasons, the NMA considers that the Publisher CR is not only proportionate but necessary to deliver significant, durable consumer benefits. This is entirely consistent with the Digital Markets, Competition and Consumer Act ("**DMCCA**") statutory objectives.

3. Control

3.1 Crawler separation: swift escalation as a credible backstop in cases on non-compliance

The CMA proposes that Google provide publishers with effective controls to withhold use of content for AI services without affecting search ranking or display, but leaves implementation detail deliberately open.

We recognise the CMA's proportionality constraints. However, effectiveness depends on enforceability: if Google retains discretion over how preferences are implemented and validated, there is a significant risk that controls will not be meaningful. The NMA therefore urges the CMA to make clear in its Conduct Requirement Notice that the CMA will move swiftly to more robust structural remedies in the event of non-compliance or ineffective implementation, including structural crawler separation.

3.2 Ensuring that publisher choice is not undermined by CR loopholes

3.2.1 Preventing circumvention: open-source dataset loophole

The Publisher CR states Google shall not attempt to circumvent a publisher's choice by acquiring opted-out content through other sources, but the Interpretive Note ("IN") indicates it may be reasonable for Google to acquire such content through "open-source datasets" where those datasets have obtained content "legally".

This a serious risk to the remedy's effectiveness. If Google can respect opt-out signals on publisher sites while ingesting the same content via open-source data sets - even where legality is contested or the scraping is accepted in weaker jurisdictions - then publisher choice becomes an illusion.

The CMA should amend the IN to state that when Google acquires open-source datasets, it must remove and delete opted-out publisher content prior to any use in training, fine-tuning, or grounding, including where the same content is present on pirate sites in the dataset. This solution means that Google will not be placed at a competitive disadvantage versus AI competitors that use open-source datasets, whilst ensuring that publisher choices are respected.

3.2.2 Secondary-source scraping via Google surfaces can undermine the CR

Secondary source crawling can target surfaces where publisher content is visible - including Google's own results interfaces - undermining publishers' control and reducing licensing value. We recommend that the Publisher CR include a requirement that Google take reasonable steps to prevent third parties from exploiting publisher content through Google-controlled surfaces where that exploitation would undermine publisher controls. It is no less damaging if publisher content surfaced on Google is used without consent by a third party, rather than Google itself.

3.2.3 Prohibit onward sale/sublicensing without informed publisher consent

The CMA must add an explicit requirement that onward sale, disclosure, or sublicensing of publisher data scraped by Google for any use must require proactive publisher permission. The value proposition and competitive consequences are fundamentally different from use within Google's own services alone, making a publisher's initial decision to opt-in to Google's use of content insufficient if that content is to be used by other businesses (and with Google profiting from the sale).

3.2.4 Anticipating a Search-specific underlying model

The Publisher CR currently reflects the reality that Gemini – using data scraped via Google-Extended – is the model that underpins AI Mode and AI Overviews. However, if Google were able to shorten the training-to-development cycle and create an underlying Large Language Model that only underpins AI Mode and AI Overviews, publishers would not have control over the training and fine-tuning of a Search-specific model (only the grounding). It may be that this eventually is covered by the anti-circumvention measures in the Publisher CR and INs, but it would be prudent for the CMA to ensure that its remedy is future-proofed in this regard.

3.2.5 Anti-circumvention must cover alternative ingestion pathways via other Google products and services

Anti-circumvention obligations should apply across all mechanisms by which Google may access or ingest publisher content, not only through web crawling. This includes content collected via pipelines associated with other Google products and services - such as Google Ads, Google Shopping, YouTube, or cloud-based tools - which operate through distinct technical channels but may nonetheless be used to support AI systems. The Publisher CR should explicitly prevent the use of alternative routes to content ingestion through Google's own products and services.

3.3 Granularity

3.3.1 Fine-tuning: a missing control that will become more important over time

The CMA indicates it is not minded to provide separate controls for fine-tuning, reflecting Google's claims that: (a) fine-tuning concerns internal processing and creates no realistic harm to publishers, and (b) fine-tuning is used across functions such as query understanding and ranking, so an opt-out would risk downranking of publisher content or would require duplicative models to be produced by Google. The CMA should not accept these assertions at face value, as Google's statements do not withstand scrutiny.

It is wrong to say that publisher value is only realised through queries about current news. In practice, news publishers are trusted, high-quality sources across a wide range of queries, and fine-tuning affects what information is prioritised and how content is ranked and presented, meaning publisher content can be highly valuable in contexts far beyond contemporary events.

We also doubt that Google would have to develop and maintain duplicative models to distinguish generative and ranking functions. Search ranking systems are optimised for retrieval and ordering (semantic relevance between queries and documents), whereas generative systems synthesise and produce novel responses and are optimised for language generation and reasoning. These functions are distinct and are treated differently within Google’s architectures; the claim of unavoidable duplication is therefore an attempt of shielding economically meaningful use from effective control.

Even if Google were able to substantiate claims of additional engineering cost to create duplicative models for search versus generative AI, the benefit to publishers would justify the cost. The central benefit the CMA intends the Publisher CR to deliver is improving the bargaining position of publishers; any incremental compliance burden should be weighed against that objective, particularly where the status quo places the costs and risks on publishers with limited visibility of use and limited ability to detect circumvention of controls.

Moreover, the CMA’s inclusion of a Google-Extended control confirms that publisher content has value in model development, whether via pre-training or other forms of updating. Fine-tuning is likely to be of much higher market value than pre-training for publisher content, because AI models disproportionately from high-quality, domain-specific, and up-to-date material. This makes fine-tuning closer in economic effect to grounding than to initial training.

Further, AI models are being fine-tuned more frequently on high-quality, domain-specific content, meaning fine-tuning may well progressively reduce reliance on grounding, diminishing the value of the grounding opt-out over time. It is also possible that requirements on Google to attribute content will incentivise them not to use grounding to surface specific content and instead rely on fine-tuning to avoid negotiations over payment and ensure that users stay within Google’s services.

The CMA should therefore introduce a separate fine-tuning control (distinct from training and grounding). The CMA should also test, rather than assume, Google’s “false dichotomy” claim about separating generative and ranking systems. Remedy design should not be anchored to unverified assertions, where publishers are unable to test the veracity of Google’s claims due to a lack of transparency.

3.3.2 Publishers need distinct controls by use case and value

The CMA anticipates Google maintaining Google-Extended as a control for training and grounding outside of general search, and implementing a new control for opting-out of grounding in search generative AI features. We support this, but stress that meaningful choice requires granular, separate controls for each economically distinct use of content scraped via Google-Extended. These uses have different substitution effects (and therefore commercial value), and different implications for publisher bargaining.

Use-specific controls is also essential to mitigate legal and commercial risk for publishers. Many publishers operate under licensing arrangements with freelancers and agencies that restrict reuse beyond standard online publication; allowing AI-driven reproduction or downstream use without differentiation could place publishers in breach of those agreements. In the absence of precise controls, publishers may be pushed toward blanket restrictions to avoid liability, limiting the supply of trusted journalism available to users and to Google itself. Given Google's established practice of offering page-level controls in other contexts, it is reasonable for the CMA to proceed on the basis that comparable technical solutions are achievable for AI uses. As AI products proliferate, meaningful publisher choice depends on clearly separated, purpose-specific controls for each distinct application, rather than aggregating legally and economically different uses under a single, general AI permission.

Publishers should be able to authorise or withhold each use independently - especially where content is used to ground Gemini outputs as distinct from being used to train the model (as well as fine-tuning). This can be done by making Google-Extended itself more granular – as the CMA proposes for the Googlebot. While publishers consider data captured by Google-Extended to fall within the scope of UK copyright protections, Google is not currently operating in a way that makes those protections effective. Absent remedies from the UK Government that make copyright enforceable and operational in this context, granular controls become a necessary backstop: they can create a workable pathway for publishers to negotiate over the value of content for grounding In Gemini, even if Google remains resistant to negotiating over training value due to the risk of setting a precedent by paying for training data. As adoption of Gemini grows, the commercial value of grounding in Gemini could become a significant revenue stream, but only if publishers have specific and enforceable controls that translate into bargaining leverage.

3.3.3 Page-level controls for Google-Extended, as well as search generative AI features are essential for premium content

The CMA proposes directory-level and page-level controls for opting out of grounding in search generative AI features, recognising that costs and benefits vary across content types. This is necessary, as in news publishing, value can be concentrated in particular content verticals. Page-level controls enable publishers to manage the trade-offs between discoverability and value extraction and are therefore central to the Publisher CR's bargaining objective.

Page-level controls are therefore also a necessary to complement the granular, use-specific controls set out above for Google-Extended. Bargaining leverage depends on the ability to make targeted decisions. If publishers are forced into generalised, site-wide choices, they cannot protect the most valuable pages while retaining the discovery benefits of other content, This will diminish the credibility of any threat to withhold, and weaken publishers' negotiation over terms for grounding.

4. Publisher Choice Tool

We are concerned that some smaller publishers (especially those with unique, e.g local, content and thus highly authoritative and valuable) may be unaware of controls or lack the technical capacity to manage them via robots.txt or metatag systems. The NMA therefore recommends that Google be required to provide an accessible ‘Publisher Choice Tool’ via Search Console and to provide prominent, repeated communication to ensure publishers are aware of and can exercise their choice. This is similar to the choice that publishers have to present to users under data protection law. This supports the CMA’s objective of meaningful choice without requiring more intrusive remedies at the outset.

Without such a Tool, many publishers may be unaware of the controls or unable to exercise them effectively. This would severely reduce the efficacy of the Publisher CR, even for those publishers that do utilise the controls – if only a few publishers are able to use the controls, then they will still have little leverage in negotiations as Google will be able to substitute in other content to ground AI Overviews and AI Mode responses.

5. Transparency metrics: ensuring publishers can measure AI impact and negotiate fairly

5.1 Transparency parity is essential for effective publisher choice

The transparency remedy should be grounded in parity between Search and AI features: Google should provide the same type and scope of information for generative AI features as it already does for Search via Search Console. This reporting should be product-agnostic and available in real-time, including for surfaces such as Google Discover. As AI interfaces increasingly become a primary point of user interaction, equivalent transparency is necessary for publishers to understand traffic shifts and negotiate on a more equal footing.

5.2 Require additional fields essential for commercial assessment

The CMA proposes that publishers receive metrics on impressions, engagement, click-throughs and click-through rates where content is used in search generative AI features. While welcome, these metrics must be expanded if the remedy is to be meaningful for commercial decision-making, and compliance monitoring.

In particular, reporting should include:

1. Aggregate and comparative performance metrics, including total impressions and referrals generated by generative AI features, and data enabling comparison between click-through rates from traditional organic search results and AI-mediated surfaces such as AI Overviews or AI Mode;
2. Advertising context, including whether advertising is shown alongside generative outputs that rely on publisher content, and how frequently this occurs, given its direct impact on value extraction and substitution effects;

3. Query and prompt characteristics, presented in an appropriately privacy-protected form, to allow publishers to understand the types of user intents that trigger the use of their content in AI outputs;
4. Audience type, indicating whether content was surfaced in response to a human user or via an AI agent interaction, as these pathways may have materially different effects on referral behaviour and monetisation;
5. Feature-specific referral sources, with separate reporting for distinct generative AI products and interfaces, enabling publishers to assess the impact of each feature independently;
6. Citation and visibility data, including how often publisher content is cited or referenced in AI outputs, the prominence of those citations, and their placement relative to other content, as a proxy for visibility and share of voice;
7. Contextual and behavioural indicators, such as visit duration, return-to-results behaviour, and downstream navigation patterns, to help publishers assess whether AI features are displacing engagement or altering user journeys; and
8. Feature prevalence over time, including whether and how often specific AI features appeared in response to relevant queries during the reporting period.

5.3 Per-feature transparency is necessary

The CMA indicates it does not currently consider it necessary to provide transparency on a per-feature basis (e.g. separate reporting for AI Overviews versus AI Mode), questioning what actionable benefit it provides.

Yet it is clear that AI Overviews and AI Mode differ in interface, user behaviour and experience, and attribution and referrals. Without per-feature transparency, publishers cannot identify which features cause harm to their business model and cannot optimise content and commercial strategy accordingly. They will also be put at even further disadvantage in negotiations with Google due to informational asymmetries. Per-feature reporting is therefore a necessity if the Publisher CR is to strengthen bargaining power.

6. Attribution: strengthen the standard and make it measurable

The CMA requires Google to take reasonable steps to ensure content is sufficiently attributed, with the INs suggesting attribution should be accurate and provide a clear means of accessing sources when content makes a significant contribution.

The obligation should be strengthened to require “prominent” attribution. “Sufficient” is not strong enough in a context where attribution is central to: (a) preserving referral value; (b) protecting publisher brands; and (c) ensuring users can trust and verify sources. The CMA must require that attribution be “prominent”, ensuring that users are encouraged to click through to the original source.

Prominent attribution must also be operationalised in the design of AI outputs. Where a sentence or substantive portion of an AI response draws on a publisher's content, users should be able to directly access that source via a clearly visible, in-line link. Attribution obligations should further be made measurable: publishers should receive data on citation and attribution rates, benchmarked against the extent to which their content is relied upon in grounding. This would allow publishers and the CMA to assess whether attribution practices are functioning as intended, and would support mechanisms that visibly reward original reporting, analogous to established citation-based signals in traditional search.

7. Jurisdiction

When assessing costs, the CMA notes that the Publisher CR design allows Google to comply on a global basis in order to capture all publishers that make content available to UK users, rather than limiting implementation strictly to UK-specific services or publishers.

This issue is significant because, if the Publisher CR were interpreted as applying only to the provision of Google Search in the UK, the controls would apply only where scraped content is used in UK Search. Content scraped from UK publishers but used to support Search or AI services in other jurisdictions would fall outside the scope of the remedy. Given the global reach and influence of UK content, particularly in English-language markets such as the USA or Australia, such an interpretation would risk creating a substantial loophole.

The current drafting of the CR and INs appears intended to avoid this, by tying the scope of protection to whether publishers make content available to UK users, rather than to the location of end-use. With this interpretation, the effect of the controls would not be limited to UK-only search services. At the same time, the absence of explicit clarification on territorial application may create uncertainty in interpretation (leading to a challenge of the Publisher CR by Google).

The CMA must engage with the European Commission and other jurisdictions to promote alignment. Search and generative AI services operate on a global basis, so greater regulatory coherence would reduce incentives for Google to evade compliance and increase legal certainty.

8. Compliance and enforcement: independent auditing must be ongoing

The CMA proposes a baseline independent compliance audit, alongside six-monthly compliance reporting by Google and stakeholder engagement.

The technical complexity of crawling, indexing, model training, fine-tuning, and grounding makes compliance difficult for third parties to observe from the outside. In these circumstances, self-reporting risks become the key limiting factor in enforceability (and publisher trust).

We propose a proportionate compromise: require an independent third-party compliance audit every six months. This would be materially less burdensome than structural crawler

separation while delivering credible oversight. Over time, if it is clear that Google is complying with the Publisher CR, the compliance audit could become more and more narrowly focused on the highest risk areas in terms of potential breaches, reducing costs for Google.

Under this approach, Google should also be required to publish a six-monthly compliance notice alongside the independent compliance audit. The timing of these publications could be staggered so that the independent third party is able to assess a recent compliance notice when commencing each audit cycle, strengthening scrutiny.

The CMA should appoint the relevant experts rather than leaving their selection to Google, with Google nonetheless required to fund all compliance reporting. Reporting should cover the incidence of zero-click searches and include a detailed assessment of any internal data combination or merging undertaken by Google after collection.

9. Implementation period

The CMA proposes an implementation period of six months for the Publisher CR. The NMA considers that this timeframe is longer than necessary for the core obligations set out in the CR, and risks delaying the delivery of benefits to publishers and consumers in circumstances where harm is already well-evidenced.

The proposed measures do not require the development of entirely new systems. Google already operates a sophisticated crawling and indexing infrastructure. They support publisher-level controls through as robots.txt, and already give reporting and transparency tools via Search Console. The Publisher CR primarily requires adaptations to existing systems - most notably to ensure that publishers can express more granular preferences and that those preferences are respected across specific AI use cases without adverse ranking effects. Therefore, a six-month implementation period appears excessive; we propose 3 months.

10. The delay of a CR to ‘trade on fair and reasonable terms’

The delay of a CR addressing fair and reasonable terms for the use of publisher content represents a significant gap. This CR was originally considered a ‘Category 2’ intervention, but has since been deferred even further, with the CMA indicating in January 2026 that it will not be considered for at least 12 months on the basis that other proposed CRs would deliver quicker benefits for UK businesses and consumers.

We understand the CMA’s aim to prioritise measures capable of rapid implementation but it is disappointing that the CMA has not recognised that a fair and reasonable terms CR is essential to directly address a long-established imbalance in bargaining power between Google and news publishers. That imbalance has constrained publishers’ ability to negotiate fair terms for the value their content brings to Search services, with direct implications for the financial sustainability of UK news publishing. Addressing this imbalance would support investment in journalism, with clear consumer benefits of increased quality and plurality of news.

The CMA must recognise that control and fair-trading terms (including potential remuneration) are closely linked. While the proposed Publisher CR may improve publishers' negotiating position to some extent, this alone does not resolve the underlying imbalance in traditional search that Parliament sought to address through the inclusion of a fair and reasonable terms power in the DMCCA (the proposition that controls for AI Overviews and AI Mode facilitate negotiations is a laudable one, but does not address the imbalance of value exchange in core, traditional search).

Publishers have always been able, in theory, to opt-out of appearing in Search altogether, but this has never translated into an ability to negotiate fair terms for the value of their content, precisely because of Google's market power. Therefore, delaying consideration of a fair and reasonable terms CR risks postponing negotiations over the value of search content that the DMCCA was designed to enable.

The CMA should proceed with work and consultation on the detail of this proposed CR, even if a later view is taken to delay enforcement, in order to support publisher negotiations and avoid delay at any later stage.

News Media Association
February 2026

