

## **The News Media Association's Response to the CMA**

### **Consultation: Proposed decision - Strategic market status investigation into Google's general search services**

The News Media Association (the “**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 46.2 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.

#### **Introduction**

The NMA welcomes the Competition and Markets Authority's (the “**CMA**”) proposed decision to designate Google Search as a crucial step forward in ensuring news media businesses can flourish and grow in the age of AI, ensuring Google's general search services do not inhibit the ability of publishers to monetise their highly valuable content. We agree with the CMA's findings that Google Search has substantial and entrenched market power and a position of strategic significance.

The inclusion of AI Overviews and AI Mode (if deployed in the UK) in the scope of the designation is particularly welcome. This recognises how these features have been incorporated into general search and now contribute to Google's market power. This is also a key step towards ensuring Google agrees fair and reasonable payment terms with news publishers, recognising the immense value that trusted news content brings to both traditional search and generative AI services.

We also welcome the CMA's Roadmap (which we will be responding to separately), in particular the prioritisation of ‘transparency, attribution and choice for publishers in how their content, collected for search, is used in Google's AI services’ as a Category 1 Conduct Requirement.

However, we are concerned that a lengthening of timelines for consultation on, and implementation of Conduct Requirements will leave publishers without the remedies they need for some time. It was originally envisaged that the CMA would be imposing the first wave of Conduct Requirements alongside making its decision on designation. Now, consultation on the first, limited set of Conduct Requirements will only begin once a decision on designation is confirmed. There is no schedule whatsoever for their finalisation or implementation. Crucial Conduct Requirements – in particular ensuring ‘fair and reasonable terms in relation to the use of publisher content’ – will not be consulted on until the first half of 2026 at the earliest. In addition, the initiation of a third Strategic Market Status investigation is being delayed by six months or more. We will set out these concerns in more detail in our response to the Roadmap.

In terms of the proposed decision, we are concerned that the exclusion of:

(a) the Gemini AI Assistant; and

(b) Google News as a standalone service,

will severely inhibit the ability of the CMA to ensure that Google's general search services trade on fair and reasonable terms in relation to the use of publisher content.

We appreciate that the CMA's proposed decision must be based on an assessment of whether these services form part of general search: the evidence we set out below clearly demonstrates that these services do fall within the scope of the relevant digital activity.

### **Gemini AI Assistant**

The CMA's provisional conclusion that the Gemini AI Assistant is not part of Google's general search services is, in our view, illogical, inconsistent with its own framework, and unsupported by the evidence.

At the heart of the CMA's reasoning is the assertion that Gemini is merely a "user" of Google's general search services, not part of it. This distinction is artificial. When Gemini generates a "grounded" response, it calls on Google's Search API and incorporates the results directly into the output. This mirrors the operation of AI Overviews and AI Mode, both of which are included in the proposed designation. Gemini is not a third-party passively using Search infrastructure; it is functionally executing a general search, albeit through a new Google interface. To put it another way, if Gemini did not have access to data scraped for the original purpose of a general search index, it could not function.

Under the CMA's own definition, general search is described as "a service that searches the world wide web, and can draw on other sources, to return information on any subject". The Gemini AI Assistant clearly falls within this definition. It submits user prompts, queries Google Search infrastructure, and returns results on any topic. Outputs by Gemini generally include hyperlinks and structured responses akin to SERPs, AI Overviews, or featured snippets.

Google's own explanation confirms that Gemini performs this function by calling Google Search APIs and incorporating results into the "context" used to generate responses. The CMA simply restates this process and then concludes, without clear reasoning, that Gemini's "blending" of this input with other sources is insufficient for inclusion. This conclusion lacks a transparent rationale.

The CMA appears to rely heavily on Google's claim that Gemini is "predominantly focused on content generation" across a broad range of use cases. But this claim is misleading.

First, the multi-functional nature of Gemini is not a valid reason for its exclusion. The relevant legal test is whether a product performs the digital activity, not whether it performs other activities as well.

Second, even what Google calls "content generation" often involves users making search-like queries. Moreover, Gemini is fundamentally built upon information retrieval, albeit across different types of sources which go beyond the general search architecture.

The CMA's own consumer research confirms that users do not clearly distinguish between responses from the SERP and those from Gemini. In fact, users are already forming search habits around Gemini, often replacing multi-step search processes with a single prompt.

The CMA also gives weight to Google’s evidence that grounding occurs in only a “minority” of Gemini use cases. We note that we cannot verify the accuracy of this claim. Regardless, this does not justify the exclusion of Gemini. If Gemini performs the general search activity in a material subset of its use cases, that is sufficient for inclusion under the CMA’s approach to other multi-purpose services. Indeed, there would be nothing to prevent Google from significantly increasing the level of responses that are grounded as soon as designation is made (and there would be a strong incentive to do so, given that grounding ensures answers are up to date, and increases accuracy). It is clear that the proportion of grounded responses produced by Gemini is already non-trivial; this use of the Google Search API should already be sufficient to bring Gemini in scope.

Further, the fact that Gemini is branded, accessed, and monetised separately from Google Search is not relevant to whether it falls within the digital activity. AI Overviews and AI Mode are also branded and monetised differently from the traditional SERP; Google Discover is accessed through a different interface and, unlike the SERP, is monetised directly through display or advertising within the Discover interface, yet all are rightly in scope. Gemini allows users to make queries, based on the same search infrastructure, returning information based on the blending of multiple pieces of information into an output. Gemini therefore performs the same core function and general search. The CMA’s highlighting of branding as a factor in its decision is concerning, as Google regularly rebrands products without any, or without substantial, functional changes – reliance on branding as a factor would make it simple for Google to bring products outside the scope of designation.

The CMA’s treatment of Gemini as merely an “access point” fails to reflect its true operation. Gemini not only initiates queries, it also consumes and delivers grounded search results. It is already the default assistant on Android, with clickthroughs to SERPs and tight integration into Google’s mobile ecosystem. Gemini is not just a gateway to Search, it is an evolution of the same product. It is an extended use of publisher content scraped for general search that creates new outputs which deliver little to no value back to source publishers.

It is therefore illogical for the CMA to accept Google’s claim that Gemini generates “original” responses, while simultaneously including AI Overviews, which are functionally identical. Both draw on the Google index and generate structured informational outputs. Whether the result is a link list or a generative narrative summary is irrelevant.

It is for this reason that the CMA’s comparison between ChatGPT and Gemini is incorrect. The fact that ChatGPT is more widely used in the UK than Gemini is irrelevant because it is the linkage between Gemini and the rest of Google’s search activities that is important in this context.

We are also concerned that the CMA’s analysis focuses exclusively on Gemini’s grounding at the point of a user query. This overlooks the fact that Google’s general search infrastructure has likely played a foundational role in training the underlying language model. The large volume of web data, including news publisher content, scraped via Google’s index or access by other means contributes to the model’s pre-training. Just because this data has been tokenised rather than drawn on directly does not mean the link to general search infrastructure, or its function as a general search service, is severed.

This point – that news and myriad other content accessed through Google’s general search architecture was used to train the large language model that underpins the Gemini AI Assistant – means it is irrelevant if only a “minority” of Gemini AI assistant responses are grounded. It is

unclear why the CMA has focused on grounding, rather than training of the underlying model, which could not have been created – or would be substantially different – if Google were not able to access substantial amounts of training data through its general search architecture. There should not be a distinction between grounded and non-grounded responses for the purposes of determining whether the Gemini AI Assistant is in scope.

This point is particularly pertinent given that Google has given particular weight to news content when training its underlying models. News and media content is the third most-used category of content in Google’s C4 dataset (which is used to train its underlying models) with half of the top ten websites represented in the dataset being news websites.<sup>1</sup> For example, [www.theguardian.com](http://www.theguardian.com) is the fifth most used website in the dataset (measured by the number of tokens representing content from that website present in the dataset). News content is overrepresented by a factor of five in C4, compared to the open-source dataset Common Crawl. Clearly, news content is not an incidental component of the millions of pages scraped by Google to form its underlying model that underpins Gemini: it is particularly valuable due to its accuracy, structure, and subject matter, and is accordingly given significant weight within the model.

Taken together, these arguments demonstrate that Gemini is functionally and technically part of Google’s general search services. And once that is accepted, demand-side arguments posited by both Google and the CMA lose relevance: Gemini and the SERP are interchangeable. Indeed, Google itself is showcasing Gemini as a SERP alternative. In any case, Gemini’s usage is growing and aligns with general search behaviour. As it becomes the default user interface for queries, the effectiveness of Conduct Requirements around transparency, control, and fair terms will erode rapidly if Gemini is not in scope.

If the CMA nonetheless persists in its position, it is critical that, as acknowledged, the input provided by Google Search to Gemini is included in scope. This must trigger the ‘fair and reasonable terms’ Conduct Requirement under the DMCCA. Even if Gemini itself is excluded, Google should not be able to extract value from publisher content and pass it to another Google product, even one it claims is distinct, without compensating publishers fairly. If Google gave that same data to a third party, publishers would expect to negotiate its use. That expectation is not erased because Google keeps the transaction in-house.

Given this, it would be useful if the CMA made clear that there is a distinction between the scope of the relevant digital activity, and the much broader scope of services that can be impacted by Conduct Requirements. It would be legitimate for the CMA, for example, to create Conduct Requirements that would impact Search access points such as Gemini. The CMA should explicitly recognise that the scoping of the relevant digital activity does not place a firm limit on the scope of remedies.

## **Google News**

The CMA should include Google News, not only when it appears in the SERP but as a standalone product, within the scope of the relevant digital activity.

Google News is not a narrow vertical search service like Google Flights or Google Shopping. Those services respond to very specific queries and provide links only to flights or products. They are presented to the user in unique ways that distinguish them from the general search results.

---

<sup>1</sup> News Media Alliance, “[How the Pervasive Copying of Expressive Works to Train And Fuel Generative Artificial Intelligence Systems Is Copyright Infringement and Not a Fair Use](#)”, October 2023

In contrast, Google News provides information in response to open-ended, natural language queries on almost any topic. It draws from a wide range of sources – not just news publishers, and functions far more like general search than a vertical aggregator. News results are presented as an integral part of the general search results.

Google News also uses the same core search infrastructure: web crawling, indexing, and ranking. The CMA describes standalone verticals as using proprietary data feeds from providers and not crawling the open web. This is demonstrably not the case for Google News.

Even if Google News has a more limited pool of sources, that does not fall outside the CMA's definition. The activity is defined by its breadth of topics, not the number of sources. Google News routinely returns relevant content in response to almost any query. The NMA has not been able to identify any queries that Google News fails to address meaningfully.

Moreover, Google itself has made clear that crawling is integral to how Google News works. In its Publisher Centre documentation, Google states: "Using Publisher Centre is entirely optional and does not affect your site's eligibility for Google News. Google automatically crawls the web regularly to include websites in Google News".

The presence of feature-specific policies (on ads, misleading content, or transparency) does not affect the underlying functionality: Google News searches the open web and returns information on any subject.

It would be inconsistent for the CMA to include Google Discover, a curated, passive news feed, while excluding Google News, which allows users to actively search and retrieve results. In fact, Google News even more clearly satisfies the definition of general search.

If the CMA chooses not to include Google News within scope, it must at least recognise the risk that unfair conduct within News could affect the general search service. Anti-leveraging Conduct Requirements might help, but they will be narrowly defined. The legal test in section 20(3)(c) is more burdensome for the CMA than the legal test for Conduct Requirements that exist fully within the scope of the designation.

For example, if Google News is not in scope, a Conduct Requirement providing 'transparency, attribution and choice for publishers in how their content, collected for search, is used in Google's AI services' would not cover news content in Google News. This could allow Google to continue to use the content of publishers who wish to withhold their content from AI services, or negotiate fair and reasonable terms for its use in AI services. It is far from clear that an anti-leveraging Conduct Requirement of the type allowed under 20(3)(c) could cover this, particularly given that the CMA does not currently intend for the Gemini AI Assistant to be in the scope of the designation. This would provide Google with a significant loophole, seriously undermining the efficacy of a Conduct Requirement focused on transparency, attribution, and choice.

We therefore urge the CMA to distinguish between genuine, narrowly scoped verticals like Google Flights and Google Shopping versus broad, topic-agnostic services like Google News, which plainly satisfy the definition of general search. This will give the CMA more flexibility to address harms throughout the designation period.

## **Google Discover**

We support the inclusion of Google Discover in the scope of the relevant digital activity.

The CMA defines general search as a service that “searches the world wide web, and can draw on other sources, to return information on any subject.” The CMA also proposes to include “all information [Google Search] returns through its underlying infrastructure.” Discover clearly satisfies both elements of this definition.

Google Search Central, with features Google’s advice to publishers, states: “Discover is a part of Google Search that shows people content related to their interests, based on their Web and App Activity.” It further explains that content is automatically eligible to appear in Discover if it is indexed by Google, that Discover includes a wide range of topics, tailored to individual interests.

A user’s search term limits the scope of traditional SERP results. In a similar way, Discover is constrained only by prior user behaviour. The underlying infrastructure is the same. Moreover, the Search bar is prominently displayed at the top of the Discover page, reinforcing that Discover is part of the general search experience. For many users, Discover is effectively the home screen of Google Search. Given the volume of traffic it directs to publishers, its functional integration with Search, and its infrastructure, Discover clearly falls within scope.

### **AI Overviews and AI Mode**

We agree with the CMA’s conclusion that AI Overviews and AI Mode are part of general search, for the reasons set out in our previous response to the Invitation to Comment.

**News Media Association**  
**July 2025**

[REDACTED]