

The News Media Association’s Response to the CMA’s [Invitation to comment](#): ‘SMS investigation into Google’s general search and search advertising services’

Background

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 46.2 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.

Q1: Do you have views on the proposed scope of our investigation and candidate descriptions of Google’s general search services?

In general, we welcome the Competition and Markets Authority’s (“**CMA**”) scoping of general search and search advertising together, given the difficulty in implementing effective Conduct Requirements (“**CRs**”) were these two interdependent services to be designated separately. Google would not be able to provide a search engine at the scale it does if it did not carry out these activities in combination. Failing to do so - only designating either general search or search advertising - would allow Google to shift anti-competitive behaviour outside the scope of the designation, preventing the CMA from implementing effective CRs.

AI Interfaces

We also broadly agree with the CMA’s descriptions of both general search and search advertising as set out in the Investigation Notice. We welcome the acknowledgement of the integration of Foundation Model (“**FM**”) applications into Google’s search engine, in particular ‘AI Overviews’, and believe it is essential that all the AI interfaces that perform search related-activities, or are dependent on search in order to operate – including ‘AI Overviews’, ‘AI organised’ search results pages, and the Gemini AI Assistant – be included within the scope of the designation.

This is crucial, as the ability of Google to integrate its FM applications into its search engine is dependent on the underlying web crawling that it conducts to create its traditional search engine results page (“**SERP**”). Further, ‘AI Overviews’ are fully

integrated into the SERP, and placed above traditional search results. Therefore, Google's market power in search allows it to: (a) scrape news publisher data to fuel its FMs; (b) scrape news publisher data to ground its FM-powered AI interface through retrieval augmented generation (“**RAG**”); and (c) use ‘AI Overviews’ to compete with the very publishers whose data it has scraped.

To note, Google's market power in search allows it to use the same crawler for indexing as it does for RAG, meaning publishers are unable to opt-out of having their content used to ground AI Overviews; otherwise, they would be de-ranked or removed from the SERP altogether.

Gemini AI Assistant

The Invitation to Comment notes that the Gemini AI Assistant taps into Google search results to provide responses to queries.¹ Again, we consider that all web crawling or indexing conducted in connection with Google's search service should be in scope of the designation, as should all services that are dependent on them, including the Gemini AI Assistant. Google would be unable to create the Gemini AI Assistant, or tap into search results as it does, absent its search infrastructure which allows it to conduct mass web scraping and indexing. Moreover, it would be entirely inconsistent to have ‘AI Overviews’ encompassed by the designation, and not the Gemini AI Assistant. The means by which they access search results for contemporary, reliable information is the same, and the product they serve to users is distinguished only by Google's dictation of which queries in its SERP are answered with ‘AI Overviews’.

Google News Showcase and Google Discover

We would strongly advocate for Google News Showcase to be in scope of the designation, as it is inextricably linked to the provision of Google's general search service. Deals struck between news publishers and Google have up to now have primarily focused on the licencing of content for Google News Showcase (allowing Google to avoid any admission of the immense value that the presence of trusted news content brings to its SERP).

Low user engagement and low levels of traffic sent to publishers' websites should not cause the CMA to underestimate the importance of Google News Showcase to the relationship between platform and publisher. The purpose of Google News Showcase is to protect Google's core search service from payment for content claims. Therefore, Google News Showcase falling out of scope of the designation would give rise to concerns that Google could use its market power to set unfair terms for Google News Showcase which in turn could reduce the effectiveness of a potential CR for Google to

¹ Competition and Markets authority, ‘[Invitation to comment: SMS investigation into Google's general search and search advertising services](#)’, January 2025

‘trade on fair and reasonable terms’ with news publishers as regards payment for news content.

If the CMA is unable to bring Google News Showcase in scope of the designation, the CMA should consider a CR of the category allowed to prevent anti-competitive leveraging from a non-designated activity to a designated activity. Google must also not be able to diminish any value, including payment, it currently provides to smaller independent publishers as part of the new regime.

We anticipate that Google Discover will fall within the scope of the designation, and would strongly urge the CMA to account for the increasing importance of Google Discover as a source of publisher traffic. Google Discover is algorithmically served, and serves content based on a users’ anticipated interests – akin to a social media platform - making it a valuable source of users that are highly likely to be engaged with the content they are served. The CMA should take care to ensure that all publisher-relevant requirements placed on Google account for the increasing importance of Google Discover.

Q2: Do you have submissions or evidence relevant to the avenues of investigation set out in paragraphs 26-28? Are there other issues we should take into account, and if so why?

Since these avenues of investigation are primarily related to understanding Google’s impact on competitors in search, search advertising, and AI interfaces, the NMA is not best placed to provide evidence on most of the issues raised.

With regards to issue (f) – *“Whether Google can influence how other forms conduct themselves in relation to general search and search advertising”* – it is clear that Google has a position of strategic significance that provides it with substantial influence over how publishers interact with its general search service. Google’s huge market share means that publishers are compelled to offer their content in Google’s search engine on their terms, and do not receive a fair and reasonable share of the joint value created.

As set out in more detail in response to Q4, Google’s market power in search also means that publishers are compelled to allow the scraping of their content for Google’s AI services, as otherwise they would be de-ranked or removed from the SERP altogether.

The lack of warning and lack of information that Google provides over algorithm changes means that publishers must allocate significant resources to understanding changes to Google’s ranking algorithms, and optimising their content for Google Search.

Regarding issue (e), as set out above we do not view the application of Google’s FMs as distinct from search, and it is clear that Google has been able to leverage its market power in search to build up its FMs (and could not have done so absent this market power).

Q3: Do you have views on how Google’s general search services might be affected by the development of AI interfaces providing alternative means of returning information?

Over the next five years or more, the development of AI interfaces providing alternative means of returning information will not – or will not significantly – diminish Google’s market power and resulting strategic position in the provision of search and information services. Indeed, by leveraging its market power in search, Google has been able to build FMs that have AI applications which only entrench its market power, such as ‘AI Overviews’. As already set out, its market power in search, and the data it has been able to gather, is critical to the development of its Gemini AI Assistant. Therefore, it is foreseeable that Google’s development of AI interfaces will only entrench its power in search and that its market power in search will allow it to become dominant in alternative means of providing information.

Any shift towards AI search engines by consumers will not change the structural chokepoints in the search market. Google’s market power in search provides it with the search index data necessary to build its search product, and competitors are not able to access this data to build AI search engines.

Moreover, Google’s position is so deeply entrenched as the dominant search provider in the UK that it is simply not foreseeable that its dominance will be meaningfully challenged by an AI search engine or underlying FM, and Google has been able to leverage its dominance in adjacent markets and integrate the vertically integrate the AI stack. The search advertising market is forecast to grow well above the wider market in coming years according to the latest AA/Warc figures published last month.²

It may be that search and AI search will never become direct competitors, with market share shifting from traditional search engines to new AI search engines. Perhaps more likely will be a hybrid model, where users will use AI search to get a broad overview of a topic, and then use search results for a more detailed, coherent, verifiable answer to their query. The integration of ‘AI Overviews’ into Google’s SERP is an example of this hybrid model. Clearly, this scenario would by no means diminish Google’s market power in search.

Q4: Do you have views on whether the issues outlined in this section are the right ones for the CMA to focus on, or whether there are others we should consider?

Broadly speaking, the CMA has correctly identified the issues that should be focused on to spur competition and reduce barriers to entry, secure open markets, and protect users. However, there are some notable omissions in the remedies the CMA has set out, and we believe the CMA could go further with interventions related to AI.

² Advertising Association, “[UK ad spend rose 9.7% to £10.6bn in Q3 2024](#)”, January 2025

In particular, we welcome proposals to:

- Prevent Google from sharing user data across services (paragraph 42, intervention (b));
- restrict the tying of Google’s search web crawling with web crawling for the purposes of its AI services (paragraph 42, intervention (c));
- place requirements on Google to have an effective complaints process for businesses who are listed in search results (paragraph 43, intervention (b));
- place requirements on Google to ensure search rankings are non-discriminatory (paragraph 43, intervention (c)); and
- place requirements on Google to ensure fair terms (including payment terms) for use of publisher content (paragraph 43, intervention (d)).

The need for interventions in most of these areas has already been well established and evidenced, and we refer the CMA to the submissions made by the NMA and news media businesses to its Online Platforms and Digital Advertising Market Study (“**the Market Study**”), as well as the conclusions the CMA drew in Advice from the CMA and Ofcom to DCMS on how a code of conduct (as it was then) could apply to platforms and content providers (“**the Advice**”).

Algorithmic Transparency

Whilst we welcome the proposal for requirements for there to be effective complaints handling processes for businesses listed in search results, and for requirements for search rankings to be non-discriminatory, we are concerned by the notable absence of a proposal for an intervention to ensure that search algorithm changes are: (a) flagged before implementation, with sufficient time for publishers to understand the changes before they are introduced; and (b) that sufficient explanation of the anticipated impact of algorithm changes is provided.

Algorithmic transparency was one of the key potential interventions raised by the CMA in the Advice. The CMA observed that “*changes to these algorithms can dramatically affect the ordering of content, with knock-on impacts on traffic directed to publishers and therefore to revenue generation*”.³ The impact of sudden and opaque algorithm changes has not diminished since the Advice was published, and a failure to introduce transparency requirements would be a significant lacuna. We further agree with the CMA’s statement in the Advice that: “*We would expect SMS firms to provide clear, relevant, accurate and accessible information to users on how the algorithms and processes which determine how content is surfaced operate, the most material factors involved, and changes to these over time*”.⁴ We also agree that “*we would expect SMS*

³ Competition and Markets Authority and Ofcom, ‘[Platforms and content providers, including news publishers: Advice to DCMS on the application of a code of conduct](#)’, May 2022

⁴ *ibid*

*firms to be required to give ‘fair warning’ of changes that ‘are likely to have a material impact’”.*⁵

We would also expect a CR that requires algorithmic transparency to apply to the use of news content in ‘AI Overviews’ – or for a similar, specific CR to be applied to ‘AI Overviews’ given that the same web crawler is used to crawl content for both traditional search results and ‘AI Overviews’. Being attributed in AI search engine such as ‘AI Overviews’ may become a significant source of traffic for news publishers, so the ability to understand how and why their content is surfaced in the results will be crucial.

We note concerns that algorithmic transparency could, if imposed in the wrong way, allow disreputable content providers to ‘game’ Google’s algorithms. It may be that the CMA should impose a CR which requires that Google to provide transparency only to trusted news publishers. A definition of a ‘recognised news publisher’ already exists in the Online Safety Act 2023.⁶ A precedent exists not only in the existence of such a definition, but in its purpose in the Act: ensuring that trusted news content is not unduly impacted by Google, and other large platform’s algorithms (in this case algorithms designed to comply with the safety duties in the Act).

The definition includes safeguards to ensure that only trusted news providers are captured by it, including that a ‘recognised news publisher’ is ‘subject to a standards code’ and ‘has policies and procedures for handling and resolving complaints’, as well as containing a definition of the ‘news-related material’ that a ‘recognised news publisher must have the ‘principal purpose’ of publishing. A requirement for algorithmic transparency applied only to ‘recognised news publishers’ would not only prevent ‘gaming’ by disreputable content providers, but also help ensure the UK public are served a broad range of trusted news content, supporting media plurality.

Sharing of data to Google’s competitors

We are concerned about the potential requirement for Google to make its data available on fair and reasonable terms to competitors to enable them to improve the quality of their search results, or to enable the development of innovative new AI search services that could compete with Google. Google is already selling RAG data to other tech giants, scraped through the web crawler which it uses for search, which publishers assert constitutes copyright infringement and is allowed to persist only because publishers are compelled to allow their data to be scraped for ‘AI Overviews’ in order to appear in search. Any requirement of this kind must account for UK copyright law: in short, news publishers’ data should only be shared with their consent.

⁵ *ibid*

⁶ HM Government, [‘Section 56, Online Safety Act 2023’](#), 2023

Restrict the tying of Google web crawling

In relation to the proposal to restrict the tying of Google’s search web crawling with web crawling for the purposes of its AI services, we recognise that this issue has developed since the CMA produced the Market Study and the Advice. We are grateful that the CMA has accounted for these developments. As well as using news content in the training of its FMs, Google’s ‘AI Overviews’ uses publisher content even when it may not have been present in the data set used to train its FMs. This RAG connects the underlying LLM with external publisher content once a query has been made by a user. Clearly, up to date, quality news content is absolutely essential for a generative search engine such as AI Overviews.

In September 2023, Google announced the creation of Google-Extended, which allows publishers to opt-out of having their content scraped to fuel Google Gemini projects, which includes Google’s key FMs. Leaving aside the fact that publishers should not have to opt-out of scraping, as UK copyright law is clear that AI firms must seek a licence before conducting any text and data mining, Google-Extended only allows publishers to opt-out of scraping for Google’s underlying FMs. It does not allow publishers to opt out of scraping for the use of news content in ‘AI Overviews’. To block scraping for ‘AI Overviews’, publishers must block Googlebot – the web crawler that is also used by Google to collect documents from the web to build a searchable index for the Google Search engine.

This means that publishers are unable to prevent their content being used to ground ‘AI Overviews’ results without their ranking in traditional Google Search results being impacted. As set out by the CMA in the Market Study, far before the advent of generative AI: *“If a publisher opts out of allowing snippets, it may be ranked lower in organic search results, reducing the publisher’s visibility to consumers and, accordingly, click-through rates, website traffic and monetisation opportunities.”*⁷

A CR to give publishers control over when their content is scraped, by separating out web crawlers for specific purposes – for example, search indexing, FM training, and RAG - (and without any discrimination if a publisher chooses to block web crawling) would be a natural evolution of the CR suggested by the CMA in the Market Study, which would grant *“publishers more control over how their content is used by the platforms, particularly when they have little discretion over the use of content, such as when it is obtained by scraping their sites”*.⁸ This CR was suggested in the context of publishers being given control over the length of snippets; it is even more vital in the context of scraping for AI.

⁷ Competition and Markets Authority, [‘Online platforms and digital advertising market study’](#), July 2020

⁸ *ibid*

Transparency over the material scraped through web crawling

However, we believe the CMA could also go further, by requiring Google to be transparent about the inputs that are used to create, fine-tune, and ground – through RAG – Google’s AI interfaces such as ‘AI Overviews’, in all instances when its ability to provide AI interfaces is dependent on its ability to crawl the web due to its market power in search. It is not sufficient for publishers to be aware of the web crawlers operated by Google, and for these to be deployed for specific purposes. Absent transparency, it will be impossible for publishers to understand if Google is complying with the instructions given to its web crawlers. This transparency will also be necessary for publishers to negotiate ‘fair and reasonable terms’ for the value of content used in Google’s FMs and AI interfaces, for example when it is used in ‘AI Overviews’. As set out in the Advice, the CMA could “*require the SMS firm to provide content providers with the information they need to make an assessment of the value of their content to SMS firms*”.⁹

Developing pre-existing proposals to account for AI

There are several other areas where the CMA could develop interventions it has already considered in the Report and the Advice, to account for the impact of AI. For example, in the Advice, the CMA stated that they could support publishers to present content in their preferred web format and enable “*news publishers to understand their users and build a customer relationship*”, including “*ensuring the news publisher’s content is adequately attributed*”.¹⁰ This CR could be developed to ensure that Google provides prominent attribution when it used a publisher’s content to ground an ‘AI Overviews’ response, with a link to the source.

Similarly, the CMA suggested in the Advice that publishers should be provided with “*access to data which would help them to understand their audience*”.¹¹ Google should be required to provide data on users’ interactions with ‘AI Overviews’ and the Gemini AI Assistant, as well as traditional search results.

In relation to the proposal to place requirements on Google to ensure fair terms (including payment terms) for use of publisher content, we are clear that – under the proposed designation – publishers should be able to negotiate for the use of content in all AI products provided by Google, when the creation and/or deployment of that AI is dependent on the web scraping that is facilitated by its market power in search.

⁹ Competition and Markets Authority and Ofcom, ‘[Platforms and content providers, including news publishers: Advice to DCMS on the application of a code of conduct](#)’, May 2022

¹⁰ Competition and Markets Authority and Ofcom, ‘[Platforms and content providers, including news publishers: Advice to DCMS on the application of a code of conduct](#)’, May 2022

¹¹ *ibid*

Fair and reasonable terms – non-payment terms

We note that this proposed intervention could also include non-payment terms regarding content presentation and attribution, and publishers' access to data. We remain supportive of the non-payment interventions proposed in the Advice.

Specifically, publishers should be able to enjoy full access to the data associated with their content, including when it is hosted by Google in an AMP format. Indeed, publishers should be able to choose which format their content is presented in. It should also be ensured that unjustified data leakage does not occur, including a restriction on the collection of data through cookies placed on a publishers' website. Google should not place any unjustified barrier in between a publisher and their consumers, including ensuring that news content is properly attributed.

Fair and reasonable terms – payment terms

We welcome the CMA's statement that payment terms could encompass the use of publisher content in 'AI Overviews', as well as when it is part of or placed alongside search results. The use of news content in 'AI Overviews' threatens to destroy the concept of 'joint value' altogether, with content being scraped in order to create a product that directly competes with publishers' in search results (and Google can use its market power to position this competitor product above traditional search results). The CMA should also ensure that the use of publisher content in the Gemini AI Assistant is encompassed by a requirement for fair and reasonable payment terms.

We agree with the CMA's assessment in the Advice that the "*assessment of joint value should take a broad view of value created through use of content by SMS firms and not just include advertising revenue from paid advertising placed around the content*".¹² Ignoring the broader benefits brought to Google through the hosting of news publisher content would see the value of news content to the search service be significantly undervalued.

Prioritisation

We note that interventions proposed by the CMA, as well as the additional interventions that we have set out, will likely be of differing salience to different news publishers, given the broad range of business models deployed across the UK news media ecosystem.

However, we are clear that the CMA should prioritise all interventions which will support the sustainability and plurality of UK news media. This will create benefits not just for the direct consumers of news content, but every citizen in the UK given the centrality of trusted news content to a functioning democracy and society. We would urge the CMA to continue to engage with individual publishers of all sizes in order to understand which

¹² Competition and Markets Authority and Ofcom, '[Platforms and content providers, including news publishers: Advice to DCMS on the application of a code of conduct](#)', May 2022

interventions are likely to have a significant positive impact for the largest number of publishers.

Q5: Do you have views on whether the potential interventions are likely to be effective, proportionate and have benefits for users, including consumers and business search users? Are there other measures the CMA should consider that would be more effective or proportionate, or that would deliver greater benefits for users?

We believe that the potential intervention will be effective, proportionate, and have benefits for consumers and business users.

Q6: What are the key lessons the CMA should draw from measures imposed in relation to general search services in other jurisdictions? Are there specific areas where imposing a similar measure in the UK is more or less important for their overall effectiveness?

Whilst we appreciate that, in developing interventions the CMA is bound to initially concentrating on CRs as opposed to Pro-Competitive Interventions (“**PCIs**”). Whilst appreciating the requirement for the CMA to ensure that its interventions are proportionate, we would urge the CMA to monitor the need for structural interventions and monitor the structural remedies being considered on other jurisdictions (such as the US where the Department of Justice has proposed that Google be compelled to sell its Chrome browser). It may well be that structural remedies are the only route through which it is possible to effectively create the desired benefits for consumers and business search users. The CMA should not adopt a ‘wait and see’ approach following the setting of the initial round of CRs, but proactively explore the need for PCIs as soon as possible.

We would also note the significant enforcement challenges that have been faced by the European Commission in enforcing the requirements introduced for ‘gatekeepers’, including Google, under the EU Digital Markets Act. Google’s ability to avoid compliance with the spirit and letter of the law is, in part, due to the nature of the DMA, which sets high-level requirements in the law itself, giving Google and other gatekeepers the scope to skirt their obligations. Whilst appreciating that setting CRs with too much detail could cause the CRs to become out of date swiftly, or allow Google to comply in a very narrow manner which does not bring the desired benefits, we would urge the CMA to be as specific as possible and also make full use of Interperatoive Notes to set out further detail on how Google could effectively comply.

We are grateful for the CMA’s continued engagement with the NMA and news media publishers and look forward to further engagement as the investigation progresses.

**News Media Association
February 2025**