

Submission to the European Commission on the Proposed Measures for Google Search Data Sharing (Article 6(11) of the DMA)

Dr Christophe Carugati

The Commission's proposed measures for Google's search data sharing with third-party online search engines require further calibration to preserve competition in data collection, protect user privacy, and ensure fairer pricing.

Introduction

The European Commission is seeking stakeholder feedback on its proposed measures for data sharing by Alphabet-owned Google Search under Article 6(11) of the Digital Markets Act (Regulation (EU) 2022/1925, DMA) until 1 May 2026 ("the consultation")¹. Under this provision, Google Search must share ranking, query, click, and view data generated by its end users in relation to both paid and free search ("Search Dataset") with third-party providers of online search engines ("OSEs"), upon request. This obligation applies under fair, reasonable, and non-discriminatory (FRAND) terms and requires the anonymisation of personal data.

The consultation forms part of the ongoing specification proceeding concerning Alphabet's data-sharing obligations under Article 8(2) of the DMA, which the Commission opened on 27 January 2026². This proceeding aims to clarify how the data-sharing requirement can effectively promote contestability and fairness in online search engine services while safeguarding user privacy.

¹ DMA.100209 – Consultation on the Proposed Measures for Google Search Data Sharing (Article 6(11) of the DMA), *European Commission* (accessed 24 April 2026). Available at: https://digital-markets-act.ec.europa.eu/dma100209-consultation-proposed-measures-google-search-data-sharing_en

² Commission Opens Proceedings to Assist Google in Complying with Interoperability and Online Search Data Sharing Obligations Under the Digital Markets Act, *European Commission*, 27 January 2026 (accessed 24 April 2026). Available at: https://ec.europa.eu/commission/presscorner/detail/en/ip_26_202

This marks the first specification proceeding at the intersection of the DMA and the General Data Protection Regulation (Regulation (EU) 2016/679, GDPR), following the joint draft guidelines issued by the Commission and the European Data Protection Board (EDPB) in October 2025³. The guidelines seek to provide greater legal certainty for market participants on the interplay between the DMA and the GDPR, including anonymisation. The proceeding also reflects recent developments in the AI sector and follows the Commission's consultation on the DMA's applicability to AI, aimed at ensuring the regulation remains fit for purpose⁴.

Against this backdrop, the Commission proposes measures addressing eligibility, the scope and conditions of data sharing, anonymisation requirements, FRAND pricing, and the processes for data acquisition and pre-acquisition testing.

These measures attempt to balance the DMA's objectives with the need to protect user privacy. However, further recalibration is necessary along three key dimensions. First, the balance between contestability and competition requires safeguards to preserve incentives for both Alphabet and third-party OSEs to invest in data collection, including protections against free-riding. Second, the balance between contestability and privacy calls for stronger protections against re-identification risks, including a more limited data scope and more robust anonymisation techniques. Third, the balance between contestability and fairness should better reflect Alphabet's investment in data collection, thereby calling for the incorporation of a return on forward-looking investments into the FRAND pricing methodology.

This submission assesses the main issues raised in the consultation and proposes targeted improvements. Each section begins with a critical analysis of the Commission's proposed measures, followed by specific policy recommendations.

Eligibility

The Commission proposes that Alphabet should not exclude third-party providers of AI chatbots with online search engine (OSE) functionalities, provided they meet the definition of an OSE under Article 2(5). This provision defines an OSE as a digital service

³ For an analysis of the interplay, see Christophe Carugati, Reconciling Competition and Privacy in Accessing Search Data in Europe, *Digital Competition*, 19 March 2026 (accessed 28 April 2026). Available at: <https://www.digital-competition.com/articles/reconciling-competition-and-privacy-in-accessing-search-data-in-europe>

⁴ Consultation on the first review of the Digital Markets Act, *European Commission* (accessed 28 April 2026). Available at: https://digital-markets-act.ec.europa.eu/consultation-first-review-digital-markets-act_en

that allows users to input queries to search, in principle, across all websites, or all websites in a particular language, on any subject, using keywords, voice commands, phrases, or other inputs, and to receive results in any format that provides access to the requested information.

Critical Analysis

AI chatbots, such as OpenAI's ChatGPT, increasingly incorporate OSE functionalities, including web search. These systems perform search-related tasks on behalf of users by grounding their outputs in underlying search indexes. Given the technologically neutral wording of the OSE definition, such services are likely to fall within its scope. However, AI chatbots also offer a wide range of non-OSÉ functionalities, such as image generation. As a result, granting them access to the Search Dataset creates a risk that the data may be used beyond OSÉ-related purposes.

The Commission's proposed measures do not define AI chatbots or establish clear eligibility criteria, creating significant uncertainty about the scope of potential eligible beneficiaries. Without a definition, the framework could capture a broad range of AI systems, including both general-purpose and highly specialised AI chatbots (e.g. legal or medical chatbots), regardless of their relevance to general search services. Similarly, the absence of eligibility criteria could allow any undertaking that integrates a customised AI chatbot with an OSÉ functionality into its services to request access to the Search Dataset. This would considerably broaden the category of eligible OSÉs, despite Article 6(11) aiming to promote contestability in general search services, rather than in specialised search services or those where a search-related AI chatbot is ancillary to the main service.

The Commission's proposed contractual safeguards for anonymisation partially address concerns about misuse of the Search Dataset for non-OSÉ purposes. These include a purpose limitation requiring third-party OSÉs to use the data solely to improve or optimise their search services, as well as governance and accountability obligations to document and ensure compliance with this purpose. The requirement to provide a reasonable assurance report further strengthens these safeguards by obliging third-party OSÉs to demonstrate credible plans for compliant data use (second assurance objective) and to implement effective technical, organisational, and contractual controls (third assurance objective).

However, these contractual safeguards primarily operate *ex post*. They do not adequately address risks at the time of access request and therefore fall short as *ex ante* protections against potential misuse.

Policy Recommendations

The Commission should define AI chatbot providers and establish clear eligibility criteria to ensure that only qualifying services can access the Search Dataset. In particular, access should be explicitly limited to OSE functionalities. To align with the aim of Article 6(11), eligibility should be restricted to undertakings that provide general-purpose AI chatbots with OSE functionalities.

Data Scope and Conditions of Sharing

The Commission proposes that Alphabet grant third-party OSE providers access to data in accordance with a parity principle, enabling them to optimise their services. This principle requires Alphabet to share all Search Data it collects, including data on queries, views, clicks, and rankings.

Regarding scope, Alphabet must share query data covering all access points, any modifications made by end users or by Alphabet, and any associated metadata relating to the user or the query. View data includes Uniform Resource Locators (“URLs”) and visual content displayed on Google Search Engine Results Pages (“SERPs”). Click data encompasses all user interactions with SERPs. Ranking data refers to the position of URLs within SERPs.

Regarding access conditions, Alphabet must provide search session data at a frequency equivalent to its own internal access, through an Application Programming Interface (API), and for a duration that allows third-party OSE providers to optimise their services effectively. Alphabet must also exclude invalid traffic that does not originate from human users or reflect genuine user interest.

Critical Analysis

The proposed parity principle strengthens contestability in OSE services by granting third-party providers access to data comparable to that used by Alphabet. However, it

risks distorting competition in data collection. By requiring full data sharing, the principle reduces Alphabet's ability to appropriate returns on its forward-looking investments in maintaining and improving data collection, while enabling third-party OSEs to free-ride on Alphabet's data rather than developing their own. This dynamic weakens incentives for both sides to invest and innovate.

This concern is particularly acute in the context of AI chatbots with OSEs, which invest significantly in data collection to improve the relevance of their answers. These investments include building proprietary indexes and collecting conversational user data, both of which could generate answers comparable in relevance to Google's search results.

The proposed scope and access conditions also raise material privacy concerns. The inclusion of extensive metadata, such as location information, poses a risk of user re-identification. In addition, near-real-time access to search session data may enable third-party OSEs to cross-reference queries across Google's and their own OSEs, further increasing the likelihood of re-identification.

Policy Recommendations

The Commission should replace the current parity principle with a qualified parity principle that preserves incentives to invest in data collection while limiting free-riding. Under this approach, third-party OSEs should demonstrate, as part of the second assurance objective in their reasonable assurance report, that they maintain meaningful and ongoing investment in data collection.

The Commission should also, in cooperation with the EDPB, narrow the scope of the Search Dataset and adjust access conditions to better protect user privacy. In particular, it should exclude data and metadata that could reasonably enable re-identification, such as location data. It should also introduce safeguards in access conditions, including measures to reduce temporal traceability (e.g., limiting the chronological sequencing of search sessions) and to mitigate real-time cross-use risks (e.g., introducing latency in data access).

Anonymisation

The Commission proposes a combination of technical and contractual measures to ensure the anonymisation of end users' personal data. The technical measures involve modifying the Search Dataset through techniques such as attribute suppression (removing or replacing specific attributes, such as IP addresses), allowlist creation (identifying permitted queries by splitting them into entities after detecting personal data), length-based thresholding (flagging unique queries), query suppression (excluding queries that do not meet predefined conditions), metadata generalisation (reducing the granularity of metadata, such as location), and mini-sessionisation (grouping user records). Together, these techniques aim to reduce the risk of re-identifying individual users.

The contractual measures complement these technical safeguards by imposing obligations on third-party OSEs. These include provisions on roles and responsibilities, as well as independent auditing requirements, to further mitigate re-identification risks.

Critical Analysis

The proposed technical measures reduce re-identification risks by altering user-level data. However, they largely rely on approaches conceptually similar to the k-anonymity framework, which focuses on modifying or suppressing identifiable attributes. This approach remains vulnerable to well-known risks, including singling out (isolating an individual), linkability (connecting records across datasets), and inference (deriving sensitive information).

For instance, attribute suppression removes or replaces identifiers such as IP addresses but leaves query content intact. Queries may still contain sensitive information, such as rare medical symptoms, which can enable re-identification. Similarly, the allowlist creation process filters queries based on the detection of personal data and includes them if they meet a threshold (at least 50 signed-in users in the past 13 months in the European Economic Area). However, if the detection mechanism fails to classify certain sensitive terms, such as rare medical symptoms, as personal data, those queries may enter the allowlist and be shared. Third-party OSEs could then re-identify users by linking this data

to their own datasets, such as conversational data generated during a user's interaction with an AI chatbot⁵.

The contractual measures further mitigate these risks by imposing compliance obligations on third-party OSEs. In particular, independent audits provide an important external check. However, audits occur periodically and assess compliance retrospectively; they do not ensure continuous monitoring or enable timely internal remediation of potential breaches.

Policy Recommendations

The Commission should, in cooperation with the EDPB, assess and incorporate additional anonymisation techniques that complement k-anonymity and more effectively mitigate re-identification risks. In particular, it should evaluate how these techniques perform when AI chatbots with OSE functionalities access the Search Dataset, as such systems may amplify re-identification risks. Conversational data collected by AI chatbots often contains more extensive and sensitive personal information than the data typically generated through traditional OSE use, increasing the likelihood of re-identification.

The Commission should also refine the allowlist creation process to reduce the likelihood that sensitive or rare queries that are reasonably linked to personal data are included in the dataset. This could include strengthening detection mechanisms for personal data and increasing the threshold required for inclusion in the allowlist.

Finally, the Commission should strengthen the contractual framework by requiring third-party OSEs to establish independent compliance functions. These should operate alongside existing data protection roles, such as data protection officers under the GDPR, and take responsibility for monitoring adherence to both technical and contractual obligations. Compliance officers should provide regular reports to Alphabet, the independent auditor, and the Commission, ensuring continuous, accountable oversight.

⁵ For an analysis of the technical measures, see Lukasz Olejnik, *The European Commission is Turning Google Search into a Privacy and National-Security Risk*, 26 April 2026 (accessed 27 April 2026). Available at: <https://techletters.substack.com/p/the-european-commission-is-turning>

FRAND Pricing

The Commission proposes a cost-based FRAND pricing methodology, applicable for five years from the date data sharing begins. In principle, Alphabet may charge compensation that reflects only the incremental costs of making the Search Dataset available, plus a reasonable return on the capital employed for that purpose, capped at Alphabet's weighted average cost of capital (WACC).

By exception, Alphabet may charge an additional margin where it can demonstrate that it cannot recover the costs of collecting the relevant data through its own commercial use of the Search Dataset, or where an eligible beneficiary operates at a very large scale. This margin may not exceed Alphabet's operating margin.

The level of FRAND compensation also varies by beneficiary category. Micro-enterprises and small and medium-sized enterprises (SMEs), as defined in Commission Recommendation 2003/361/EC, are exempted from the margin exception above and remain subject only to the incremental cost and capped-WACC return, even where Alphabet would otherwise be unable to recover the costs of collecting the relevant data through its own commercial use. Conversely, for beneficiaries that are themselves gatekeepers designated in respect of an online search engine core platform service, Alphabet may depart from the methodology through good-faith negotiations, provided that the resulting terms remain FRAND.

Critical Analysis

The proposed cost-based methodology provides a structured economic framework but raises important fairness concerns. In its current form, it does not adequately reflect investment in data collection, nor does it preserve incentives for such investment.

First, the methodology limits returns to incremental costs and a capped cost of capital, without explicitly accounting for the forward-looking investments required to maintain and improve data collection. This approach also risks weakening Alphabet's incentives to invest and innovate in data collection.

Second, the recoupment exception for micro-enterprises and SMEs prevents Alphabet from charging an additional margin that would cover the costs of collecting the relevant data. The result is an implicit cross-subsidy from Alphabet to SMEs.

In addition, FRAND pricing determinations typically rely on bilateral negotiations, which can delay implementation. The complexity of the proposed methodology increases the likelihood of pricing disputes between Alphabet and third-party OSEs. Although parties may refer such disputes to the Commission, resolution is likely to involve lengthy and resource-intensive proceedings. This dynamic risks delaying effective access to the Search Dataset and, in practice, may position the Commission as a price regulator.

Policy Recommendations

The Commission should retain a cost-based methodology as a foundation for FRAND pricing but refine it to better reflect fairness considerations. In particular, the methodology should incorporate a reasonable return on investments in data collection, including forward-looking improvements.

The Commission should also reconsider the recoupment exception for micro-enterprises and SMEs, so that Alphabet can apply the additional margin when the conditions are met, regardless of beneficiary category. This would avoid an implicit cross-subsidy.

Finally, to improve speed and legal certainty, the Commission should complement the methodology with a more transparent and standardised pricing framework. This could include predefined pricing benchmarks based on objective criteria such as data scope and usage volume, thereby reducing the scope for disputes and facilitating faster access to the Search Dataset.

Process for Search Dataset Acquisition and Pre-acquisition Data Testing

The Commission proposes a set of procedural and practical arrangements governing access to the Search Dataset. These include a framework for pre-acquisition data testing based on three types of samples with varying scope and conditions. Sample A consists of a small dataset of 1,000 rows provided free of charge. Sample B comprises a synthetic dataset of up to 10 million representative, artificially generated queries and associated

metadata, made available at a FRAND price. Sample C consists of a dataset corresponding to 5% of the final Search Dataset, also subject to a FRAND price.

In addition, Alphabet must finalise the Search Dataset within three months of the adoption of the Commission's implementing act and prepare template licensing agreements for both dataset acquisition and pre-acquisition testing within two months. Access to the dataset is conditional on the submission of a Level 1 reasonable assurance report for initial access and a Level 2 report for continued access. The framework further requires Alphabet to issue reasoned decisions when refusing, suspending, restoring, or terminating access, including in expedited cases where there is an urgent risk of serious and irreparable harm to the anonymisation of personal data. Alphabet may also impose strictly necessary and proportionate financial penalties for non-compliance or misrepresentation, and must notify both the Commission and the relevant data protection authorities of its decisions.

Critical Analysis

This framework raises concerns regarding governance and oversight. It assigns Alphabet a central operational role that includes decision-making authority over access, compliance, and enforcement. Although procedural safeguards limit this discretion, Alphabet retains a degree of unilateral decision-making that may create issues in the event of disagreements with third-party OSEs.

The pre-acquisition testing framework also raises concerns related to information asymmetry. In particular, Sample B, the synthetic dataset, plays a central role in enabling third-party OSEs to assess the value and usability of the Search Dataset. However, third-party OSEs cannot independently verify whether this dataset accurately reflects the underlying real Search Dataset. While Alphabet must disclose the methodology used to generate the dataset, this requirement does not enable independent verification of the synthetic dataset.

The regime governing financial penalties further raises issues of predictability and proportionality. Although the requirement that penalties remain strictly necessary and proportionate provides a general constraint, the absence of a defined methodology or explicit caps introduces legal uncertainty.

Policy Recommendations

The Commission should strengthen institutional oversight by establishing itself, or an independent adjudicatory mechanism, as the primary forum for resolving disputes between Alphabet and third-party OSEs.

To address information asymmetries with Sample B, the Commission should require independent verification of the synthetic dataset. An external expert or auditor should assess whether Sample B faithfully represents the Search Dataset, thereby enabling third-party OSEs to make informed decisions.

Finally, the Commission should introduce a clear methodology and explicit caps for financial penalties. This framework should specify how penalties are calculated to ensure predictability and reduce the risk of discretionary enforcement.

About

Digital Competition

Digital Competition (<https://www.digital-competition.com/>) is dedicated to promoting digital and competition policies that foster innovation. We bring together legal, economic, and policy expertise to deliver cutting-edge research, training, strategic advice, and high-impact regulatory intelligence.

Led by Dr Christophe Carugati, with a decade of deep expertise in digital and AI markets, we address the most pressing challenges in the rapidly evolving digital and competition policy landscape, impartially, rigorously, and with a commitment to actionable recommendations.

Google provided financial support for this submission. The views, analyses, and recommendations are solely those of the author, not those of his clients, which also include Apple and Amazon.

This paper is part of our Digital Competition Regime Hub (<https://www.digital-competition.com/digitalcompetitionregime>). We conduct research on the design, implementation, and enforcement of digital competition regimes worldwide, from the EU DMA to the UK DMCCA.

Contact us for service, membership, training, or press inquiries.

Dr. Christophe Carugati



Dr. Christophe Carugati (christophe.carugati@digital-competition.com) is the founder of Digital Competition. He is a renowned and passionate expert on digital and competition issues with a strong reputation for doing impartial, high-quality research. After his PhD in law and economics on Big Data and Competition Law, he is an ex-affiliate fellow at the economic think tank Bruegel and an ex-lecturer in competition law and economics at Lille University.