Submission to the European Commission

Case DMA.100209 - SP - Alphabet - Article 6(11) DMA

Making Access to Google Search Data Work Without Inventing a Privacy Fiction

A Consultation filing on the European Commission's preliminary measures, anonymisation, utility, personal data and institutional coherence

Prepared by Dr Mark R. Leiser
Associate Professor of Law
Technology, data protection, consumer protection and platform regulation

Submission for European Commission
April 2026

Reference style note

This draft uses compact source markers in square brackets. Pinpoints are expressed as, for example, [n. 1, at paras. 20-34] or [n. 2, at pp. 1-3].

Scope of this Filing

This version is built around the actual structure of the Commission's preliminary measures. It does not merely restate the general privacy-utility argument. It identifies which elements should be retained, which should be amended, and which should be supplemented so that Article 6(11) works in practice.

Executive Summary

I support the objective of Article 6(11) DMA. A gatekeeper that benefits from scale in search data should not be able to convert that scale into a permanent insulation from contestability. The Commission is therefore right to specify measures that make Google Search data sharing operational, not merely nominal [n. 1; n. 2; n. 3].

The legal sequence matters. Article 6(11) DMA requires access to ranking, query, click and view data on FRAND terms, but it also states that any query, click and view data that constitutes personal data shall be anonymised. Recital 61 explains that personal data should be protected, including against possible re-identification risks, by appropriate means such as anonymisation, and without substantially degrading the quality or usefulness of the data. That recital does not invert the operative provision. Anonymisation is the condition for sharing personal query, click and view data. Utility matters once that condition is met; it cannot be used to lower the anonymisation threshold [n. 3].

The preliminary measures are a serious and unusually detailed attempt to make Article 6(11) DMA work. They set out eligibility rules, a broad data scope, daily API-based access, a two-layer anonymisation model, recipient obligations, independent assurance, FRAND pricing, pre-acquisition samples, and processes for refusal, suspension and termination [n. 1, at paras. 1-145]. That structure should be retained. But it should be improved because the proposed approach risks over-relying on contractual restrictions to perform too much of the legal work, following an insufficiently robust technical anonymisation layer.

The core weakness is not merely that the preliminary measures lean too heavily on a single transformed record-level dataset. It is that the measures risk treating a dataset as anonymous because direct identifiers have been removed, detectors and thresholds have been applied, and recipients are contractually prohibited from re-identifying, linking, augmenting or misusing the data. Those contractual measures are important, but they cannot be the foundation of anonymisation if realistic, lawful means of re-identification remain available to recipients or other relevant actors [n. 1, at paras. 19-43; n. 5].

That problem has consequences beyond this case. If the Commission creates a DMA-specific form of anonymisation that depends heavily on contractual prohibitions and private enforcement, it will add yet another unstable meaning of anonymisation to an already unsettled EU digital acquis. The same Union is currently debating pseudonymisation, identifiability and implementing-act criteria in the Digital Omnibus. Article 6(11) measures should not deepen that ambiguity by creating a special search-data standard that appears more flexible than the GDPR standard without clearly explaining why [n. 5; n. 6; n. 7].

The final decision should therefore convert the current design into a Commission-specified Safe Search Data Access Regime that starts from the primacy of anonymisation. The regime should preserve daily export for lower-risk, common-tail data where anonymisation is technically defensible; use privacy-preserving aggregates where aggregate signals are sufficient; route high-risk but high-value data into controlled API or clean-room for low-risk data access where appropriate; create regulator-supervised escrow for contested suppressions; and maintain suspension and termination mechanisms for misuse.

The final measures should also require an Anonymisation and Utility Impact Assessment. The AUIA should not be paperwork, nor should it become a mechanism for trading down the anonymisation requirement. It should be the evidence pack that first demonstrates how the data has been rendered anonymous in the relevant recipient context, and only then explains what utility remains for concrete online search engine tasks. That ordering preserves both Article 6(11)'s legal condition and its contestability purpose.

Text box 1 - What this filing asks the Commission to do

1. Reaffirm that Article 6(11) makes anonymisation the legal condition for sharing personal query, click and view data.

2. Read Recital 61's usefulness language as requiring preservation of utility after anonymisation has been achieved, not as a licence to dilute anonymisation.

3. Amend the measures so that contractual obligations complement, but do not substitute for, robust technical anonymisation.

4. Require a formal AUIA that tests means reasonably likely to be used, recipient influence over the dataset, auxiliary-data linkage, and utility after anonymisation.

5. Use tiered access modalities to preserve contestability value where safe export is not possible.

6. Clarify recipient obligations, DPA notification and conditional compliance reliance without creating joint-controllership or private-enforcement uncertainty.

7. Avoid creating a DMA-specific meaning of anonymisation that destabilises the GDPR, Digital Omnibus, DSA, Data Act and AI governance landscape.

1. The Consultation Object: Not a General Policy Debate, but a Draft Operative Regime

The Commission asks for feedback on the effectiveness, completeness and implementation timelines of the proposed measures [n. 1, at p. 1]. That framing matters. The task is not to rehearse a general debate about whether privacy or competition should prevail. The legislature has already created an access obligation in Article 6(11). The Commission is now designing the operational regime through which that obligation will be implemented.

For that reason, this submission engages directly with the structure of the preliminary measures: eligibility; data scope and conditions of sharing; anonymisation; FRAND pricing; pre-acquisition testing; finalisation of the Search Dataset; licensing; eligibility review; continuity of access; refusal, suspension and termination; financial penalties; and notification to the Commission and data protection supervisory authorities [n. 1, at pp. 1-29].

My position is constructive but not deferential. The measures are directionally right because they recognise that anonymisation requires technical alteration and recipient governance. They are incomplete because they need a clearer evidence layer, field-level tiering, recipient-context testing, utility benchmarks and a better treatment of high-risk but high-value search signals.

Text box 2 - The One-Line Position

The final measures should require effective access to contestability value without treating controlled access as a substitute for anonymisation. Where data can be technically anonymised for ordinary export, export. Where data has been anonymised, but ordinary download would create marginally elevated operational, security, combination or output-leakage risks, use controlled access. Where the underlying data cannot be anonymised to the Article 6(11) standard, do not move that data into a clean room[1] and call it anonymous. Suppress it, aggregate it, or permit only privacy-tested model or output access, where the artefact leaving the environment is tested against memorisation, extraction, reconstruction and re-identification risks.

2. Legal Starting Point: Article 6(11), Anonymisation and Personal Data

Article 6(11) DMA requires a gatekeeper operating an online search engine to provide third-party undertakings providing online search engines with access, on FRAND terms, to ranking, query, click and view data in relation to free and paid search generated by end users on its online search engine. The same provision then adds the decisive condition: any such query, click and view data that constitutes personal data shall be anonymised [n. 3]. That sentence is not a gloss on the obligation. It is the legal condition for sharing personal query, click and view data.

That matters because Article 6(11) is sometimes too easily described as if it were a general data-sharing duty with privacy safeguards attached. It is more precise than that. The operative provision requires access, but it also requires anonymisation where the data is personal. Recital 61 confirms the contestability rationale and refers to the need to protect personal data, including against re-identification risks, without substantially degrading the quality or usefulness of the data [n. 3]. But a recital cannot reverse the sequence created by the operative text. The first question is whether personal data is present. If it is, the data must be anonymised. Only then should the Commission ask how much utility can be preserved. Usefulness matters, but it does not lower the legal threshold of anonymisation.

This should also discipline the Commission's use of Article 8(2) DMA. Article 8(2) allows the Commission to specify measures that a gatekeeper must implement to ensure effective compliance with the DMA. It does not allow the Commission to create a DMA-specific, softer version of anonymisation. If the final measures treat search data as anonymised because a limited set of technical treatments has been applied and recipients have promised not to re-identify users, the problem is not merely technical. It risks creating yet another operational meaning of anonymisation within the EU digital rulebook.

That would be a serious legal-certainty problem. The Union is already struggling with the boundary between personal data, pseudonymised data and anonymous data in the Digital Omnibus debate, in EDPB pseudonymisation guidance, in DPA guidance, in the GDPR/DMA interface and across the wider digital acquis. If Article 6(11) proceeds on the basis that under-tested technical measures can be rescued by contract, the result will be one standard of anonymisation for the DMA, another for GDPR guidance, another for the Omnibus debate, and

[1] For avoidance of doubt, when this filing refers to a "clean room" or "controlled API", it does not mean a setting into which data that enables re-identification may be placed and then treated as compliant because access is restricted. A clean room is not a substitute for Article 6(11) anonymisation. The term is used here in a narrower sense. It refers to a controlled environment for data that has already satisfied the relevant technical anonymisation threshold in the recipient context, but where ordinary download would create marginally elevated operational, security, combination or output-leakage risks. In that setting, controlled access may add safeguards such as no local copy, output checks, query whitelisting, logging, rate limits, role controls, extraction testing and auditability. Where the data cannot be anonymised to the required legal standard at the level of the underlying training or input data, the answer is not to disclose that data through a clean room. The relevant signal should instead be suppressed, aggregated, or used only to generate a model, feature, benchmark, aggregate or other output that is itself rendered non-identifying and tested for memorisation, extraction, reconstruction and re-identification risk before any disclosure to a recipient.

potentially others for the DSA, Data Act, AI Act and data-space initiatives. The Commission should not add to that fragmentation. It should use this proceeding to strengthen legal certainty [n. 4; n. 6; n. 7].

The GDPR starting point is familiar but important. Article 4(1) GDPR defines personal data as information relating to an identified or identifiable natural person. Recital 26 GDPR requires attention to all means reasonably likely to be used, either by the controller or by another person, to identify the natural person directly or indirectly. That assessment is objective: cost, time, available technology and technological developments matter. Article 4(5), by contrast, defines pseudonymisation as processing that prevents attribution without additional information, where that additional information is kept separately and protected. Pseudonymisation is valuable. It is not anonymisation.

The final measures must therefore apply Recital 26 GDPR properly. The question is not whether a recipient says it has no intention to identify users. Nor is it whether a contract predicts good behaviour. The question is whether identification is reasonably likely in the concrete processing context, taking account of lawful and practical means available to the recipient or to another person whose means must realistically be considered. That does not require imagining unlawful dark-web databases or fanciful intrusions. It does require considering public sources, the recipient's own search data, advertising or analytics data, browser or location data, panel data, model outputs and other information lawfully available to or generated by the recipient [n. 5].

This is where the preliminary measures are most vulnerable. They propose a technical layer and a contractual layer. The technical layer removes direct identifiers, suppresses precise timestamps and certain attributes, replaces image queries, uses allowlists, length thresholds, metadata generalisation and mini-sessionisation [n. 1, at paras. 19-34]. Those measures are not irrelevant. Some are necessary. But search data is not just a table with identifiers removed. It is language, behaviour, geography, timing, intent and context. Query strings may contain names, addresses, workplaces, schools, medical details, allegations, rare events and unusual combinations of facts. A detector-based system will miss some of this. A threshold-based system will misclassify some of it. At scale, that is not an edge case. It is a foreseeable feature of the domain.

A simple example makes the point. Suppose a user enters a distinctive run of searches. One query includes "my name is X", but X is also a common noun and is not caught by a name detector. Other queries reveal a city, a university, a profession and a narrow institutional connection. A recipient with access to public professional profiles, institutional webpages or its own search logs may not need anything unlawful to connect those details. The issue is not whether the recipient promised not to do so. The issue is that ordinary, lawful means may make identification objectively feasible. In that case, the data has not been anonymised in the relevant recipient environment. It has been partially transformed, or at most pseudonymised.

A second example illustrates the active-risk problem. Suppose a third-party OSE understands that entities appear in the Search Dataset only above a particular frequency threshold. The OSE, its employees, testers, contractors or associated users can lawfully run searches on Google Search. Over time, they may be able to influence whether certain entities appear in the shared dataset, or create a reference trail that later helps them recognise records once the dataset is accessed. Again, this does not depend on hacking, theft or illegal datasets. It uses lawful search behaviour and the recipient's own knowledge. The Recital 26 question is not whether this will certainly happen. It is whether the means are reasonably likely to be used in the concrete context.

The contractual layer does not eliminate that problem. Contractual prohibitions on re-identification, record-level linking, augmentation, onward transfer, advertising or analytics use, excessive retention and circumvention are valuable safeguards [n. 1, at paras. 38-43]. They may reduce misuse. They may help define the recipient environment. They may be relevant to whether certain means are lawful or reasonably likely. But they cannot convert a technically fragile dataset into anonymous data. If a dataset remains identifying when assessed against

realistic lawful means, a promise not to identify does not cure the legal status of the data. It moves the problem from anonymisation to private enforcement.

That move is risky for both the DMA and the GDPR. If the Search Dataset remains personal data for some recipients, the final measures risk distributing pseudonymised or personal data to multiple third-party OSEs, possibly including recipients and processing environments outside the EEA, without a settled account of GDPR rights, lawful bases, transparency duties, controller responsibilities, security obligations or Chapter V transfer consequences. Article 6(11) was not designed as a route around those protections. It was designed to require anonymised sharing where personal query, click and view data is involved.

It also creates uncertainty about roles and responsibilities. Article 4(7) GDPR defines a controller by reference to who determines the purposes and means of processing. The preliminary measures seek to recognise Alphabet and recipient OSEs as independent controllers while requiring Alphabet to impose detailed contractual controls on recipient processing [n. 1, at paras. 35-37]. There may be good practical reasons for those controls. But the more the regime depends on Alphabet policing recipient behaviour by contract, the more fragile the allocation of responsibility becomes. The Commission should not design a public-law obligation whose privacy effect depends on private enforcement while the legal status of the data remains unclear.

This is why stronger technical anonymisation must remain available, even if it reduces some utility. The final measures should not force Alphabet away from better-understood technical approaches, including query-level and query-plus-metadata frequency thresholds, aggregation where thresholds are not met, and stronger suppression where semantic or contextual identifiability remains. The duty under Article 6(11) is not to maximise utility at all costs. It is to provide access on FRAND terms subject to anonymisation of personal query, click and view data. Utility matters within the set of methods that actually achieve anonymisation.

The treatment of personal data contained in query text also requires care. A query may include a name, address, workplace, school, medical condition, allegation or other personal data element. Sometimes that element relates to the person who searched. Sometimes it relates to someone else. Often no system can reliably know which is true. For Article 6(11) purposes, the safer and more coherent assumption is that personal data embedded in query text may be part of the end-user anonymisation problem. The final measures should not allow such data to be moved outside Article 6(11) merely because it might relate to another person. That would create an artificial gap in protection.

The Commission should therefore avoid presenting the problem as a flat trade-off between anonymisation and usefulness. The legal order is stricter than that. Anonymisation is the precondition. Once anonymisation is achieved, usefulness should be preserved as far as possible. If unrestricted record-level export cannot meet the anonymisation requirement, the answer is not to lower the standard. The answer is stronger technical treatment, aggregation, delayed access, controlled API access, clean-room access for low-risk data, trusted execution or regulator-supervised escrow, depending on the data slice and the residual risk.

Text box 3 - The legal sequence the final measures should follow

First, ask whether the query, click or view data constitutes personal data under Article 4(1) GDPR and Recital 26 GDPR.

Second, ask whether that data has been anonymised, not merely pseudonymised within the meaning of Article 4(5) GDPR.

Third, assess the means reasonably likely to be used by the recipient or another person by reference to objective, lawful and practical means, not predicted incentives or contractual promises alone.

Fourth, use contractual, organisational and audit measures as safeguards supporting a technically robust anonymisation design, not as substitutes for anonymisation.

Fifth, once anonymisation is achieved, preserve contestability through the least risky access modality that still provides meaningful value.

The Commission can still make Article 6(11) work. But it should not do so by adopting a fragile, contract-dependent understanding of anonymisation. Nor should it create a DMA-specific category of "anonymous enough because recipients promised not to re-identify". That would be bad for users, bad for legal certainty, and bad for the coherence of the EU digital acquis. The better approach is to reaffirm the primacy of anonymisation, permit stronger technical measures where necessary, and then preserve contestability through a tiered access architecture that does not require pretending that weak record-level disclosure has become anonymous by contract.

3. The Proposed Architecture: A Safe Search Data Access Regime

The final measures should treat data sharing as a family of access modalities. The Commission already does something similar in other sensitive data settings. Access can be mediated by eligibility, purpose limitation, documentation, secure environments, audit, confidentiality and output controls. Article 6(11) does not require the Commission to collapse access into downloadable possession where possession is not the safest or most useful way to deliver the contestability objective [n. 12; n. 13].

The governing principle should be simple, but the order matters: anonymise first, then choose the access modality. Export should be available where the data can be technically anonymised for ordinary disclosure. Controlled API or clean-room access should be available only where the data has already met the Article 6(11) anonymisation threshold, but ordinary download would create marginally elevated operational, security, combination or output-leakage risks. Where the underlying data cannot be anonymised for recipient access, Article 6(11) should not be implemented by placing non-anonymous training data into a controlled environment. The appropriate alternatives are stronger technical treatment, aggregation, suppression, regulator review, or privacy-tested model/output access.

Tier	Access modality	Safeguards	Contestability value
0. Methodology and synthetic preview	Public schema, methods, Sample A/Sample B style testing, developer documentation.	No real query exposure; synthetic memorisation testing; clear limits on representativeness.	Onboarding, testing and integration without privacy leakage.
1. Privacy-preserving aggregate layer	Differentially private or thresholded counts, trends,	Privacy budget, clipping, rate limits, composition accounting, public parameter summaries.	Trends, market entry, language and geography signals.

	query-pair frequencies and cohort signals.		
2. Exportable common-tail record API	Transformed record-level data for frequent, lower-risk queries and coarse metadata.	Technical measures, recipient certification, no-join obligations, retention and audit.	Core Article 6(11) search-quality use cases.
3. Controlled API / clean room	Anonymised data whose residual risk is marginally elevated relative to ordinary export, or privacy-tested model/output access for signals that cannot be disclosed as underlying data.	No non-anonymous raw data disclosed to recipients; no local copy; output checks; query whitelisting; rate limits; logs; role controls; extraction, memorisation and reconstruction testing; independent audit.	Preserves useful signals in narrow cases without treating a clean room as a substitute for anonymisation.
4. Regulator escrow/expert review	Confidential inspection of contested suppressions and threshold decisions.	Commission, DPA or trusted expert access; professional secrecy; reproducible reports.	Trust without exposing raw sensitive data or trade secrets.
5. Emergency suspension and breach tier	Misuse, re-identification attempt, security incident, unlawful onward transfer or AI misuse.	Expedited pause, notification, deletion certification and forensic audit.	Safety valve without a unilateral Alphabet veto.

Text box 4 - Contestability value, not Raw-Form Parity

The parity principle in paragraph 3 should not be treated as the governing legal standard. Article 6(11) requires access on FRAND terms to ranking, query, click, and view data; it does not require raw-form parity with every internal use of data by Google. The correct test is whether the recipient obtains an effective contestability value under FRAND conditions through the least risky access modality that preserves that value.

The detailed paragraph-by-paragraph engagement with the Commission's preliminary measures is set out in Annex C. The main text now turns to the cross-cutting additions that hold the proposal together: the AUIA, AI ring-fencing, implementation phasing and operative drafting.

4. Anonymisation and Utility Impact Assessment (AUIA)

The most important addition to the final measures should be an AUIA. The Commission's current technical pipeline is detailed. Still, the final measures require a mechanism to demonstrate that the pipeline first meets the anonymisation requirement and then shows the extent to which contestability utility has been preserved. That order matters: utility is a statutory reason to avoid unnecessary degradation after anonymisation has been achieved; it is not a reason to lower the anonymisation standard.

The AUIA should be prepared before the final Search Dataset is made available, updated periodically, and revised after material changes in data fields, access modalities, recipient classes, AI capabilities, attack models or evidence of misuse. A confidential version should be available to the Commission and, where appropriate, competent DPAs or trusted auditors. A non-confidential version should be available to eligible recipients.

AUIA module	Question	Output

Data inventory	Which fields exist, which are removed, generalised, binned, replaced, noised, linked or suppressed?	Field map, schema, sensitivity labels and transformation map.
Personal data and actor analysis	Does a query contain personal data, and in which recipient context could that data allow the end user who issued the query to be identified, taking into account that the system cannot reliably determine whether named or described persons are the end user or someone else?	End-user identifiability assessment, recipient-context assessment and responsibility map.
Threat model	Which attacks are reasonably likely: singling out, linkability, inference, active attack, auxiliary data, AI-assisted linkage?	Attack tree and residual-risk analysis by recipient profile.
Technical measures	Why these thresholds, bins, allowlists, location cells, suppression classes and session rules?	Parameter justification and calibration curves by language, market and query class.
Utility tests	Can recipients still improve ranking, query understanding, local search, freshness, head-and-torso learning, carefully controlled rare-query analysis, and evaluation of the OSE function?	Benchmark dashboard, degradation curves and task-level utility report.
Recipient controls	What can the recipient join, access, train, derive, store, share and delete?	Certification, architecture diagram, no-join controls, AI lineage and audit mapping.
Red team and monitoring	Can realistic adversaries identify, link or infer? Do risks change over time?	Red-team report, remediation log and reassessment triggers.

Proposed AUIA clause

Alphabet shall maintain an Anonymisation and Utility Impact Assessment (AUIA) for the Search Dataset and for each specified access modality. The AUIA shall assess, at a minimum, field-level risk, recipient-context risk, singling out, linkability, inference, personal data embedded in query text where it cannot be determined whether the data relates to the end user or another person, auxiliary data linkage, AI-assisted re-identification, utility for concrete OSE tasks after anonymisation has been achieved, threshold calibration, suppression effects, red-team results, and reassessment triggers. A non-confidential summary shall be made available to eligible recipients, and a confidential version shall be made available to the Commission and, where appropriate, competent supervisory authorities or trusted auditors.

5. AI Chatbots and Search Assistants

The preliminary measures take a broad approach to AI chatbots with OSE functionality. That is exactly why the final measures should be more precise. The relevant question should not be whether a broader AI service contains some search-like component. The relevant question is whether the requested data is necessary for a genuine OSE function under Article 6(11). If an AI service is eligible because it provides OSE functionality, access should be limited to that function and not extend to the wider service. Article 6(11) should not serve as a route to obtain search query, click, view and ranking data for general-purpose AI development. The final measures should separate permitted OSE uses from prohibited non-OSE uses.

AI use case	Recommended treatment	Reason

Retrieval/ranking evaluation	Permitted under ordinary recipient controls.	Directly tied to OSE contestability.
Evaluation of OSE retrieval or ranking outputs in a chatbot interface	Permitted only where the evaluation is limited to the OSE function and does not use Article 6(11) data for general model training, fine-tuning or grounding of a general-purpose model.	Keeps the use tied to retrieval or ranking quality, not to an AI data pipeline.
Fine-tuning a search-specific component	Not part of ordinary Article 6(11) access; at most a separate, expressly specified and audited exception for a separable OSE component.	Fine-tuning raises memorisation, deletion, and auditability problems and should not be treated as routine search data sharing.
General foundation-model training	Prohibited.	Not the OSE function, not foreseen by Article 6(11), and difficult to reverse or audit.
Advertising, analytics or identity enrichment	Prohibited.	Creates the auxiliary-data linkage risk that the measures are meant to prevent.
Onward transfer to affiliates or processors.	Only approved processors under equivalent controls.	Maintains Article 6(11) governance perimeter.

Text box 5 - AI principle

An AI chatbot should be eligible only for its online search engine function. Retrieval, indexing, ranking, query understanding, SERP interaction and evaluation of the OSE function may be legitimate. Using Article 6(11) data to train, fine-tune or ground a general-purpose AI system, or to enrich advertising, identity graphs, profiling or unrelated services, should be prohibited.

6. Implementation timeline and phasing

The Commission should not allow additional safeguards to become a delaying tactic. The answer is phasing. The final measures can preserve the Commission's proposed two-month and three-month milestones while sequencing access modalities by complexity and risk.

The first phase should include public documentation, Sample A, Sample B, licence templates, the non-confidential AUIA framework, the recipient request form, and lower-risk common-tail access. The second phase should include the privacy-preserving aggregate layer, utility dashboards, reason-code reporting and standard recipient certification. The third phase should pilot controlled access or clean-room access for low-risk data with high-value signals. The fourth phase should operationalise regulator escrow, independent red-team and periodic AUIA updates.

Timing	Deliverable	Why it matters
0-2 months after the final act	Licence templates; public webpage; Sample A; Sample B; non-confidential methodology; request forms; standard AI/no-join clauses.	Matches paras. 100-104 and reduces onboarding uncertainty.
0-3 months	Finalise core Search Dataset; launch exportable common-tail record API and initial AUIA summary.	Matches para. 99 while avoiding delay for lower-risk data.
3-6 months	Privacy-preserving aggregate API; suppression reason codes; utility dashboard; recipient certification library.	Adds evidence and challenge without blocking initial access.

6-9 months	Pilot controlled API / clean-room access for low-risk data with rare-tail, local, fresh and sequence-rich signals.	Preserves high-value data that would otherwise be suppressed.
9-12 months and annually	Independent red team; Level 2 assurance cycle; AUIA update; Commission/DPA pattern review.	Keeps anonymisation and utility current.

7. Suggested Operative Amendments

The following drafting is not intended as legislative text. It is intended to show how the final measures could incorporate the main recommendations without weakening Article 6(11).

Draft clause 1 - Access modalities

Alphabet shall provide effective access to Search Data through access modalities specified in these measures. Where unrestricted record-level export can satisfy the Article 6(11) anonymisation requirement while preserving material contestability value, Alphabet shall provide export access. Where data has satisfied the anonymisation requirement but ordinary download would create marginally elevated operational, security, combination or output-leakage risks, Alphabet may provide access through controlled API access, clean-room access, trusted execution or equivalent controlled modalities. Where the underlying data cannot be anonymised to the Article 6(11) standard for recipient access, Alphabet shall not be required to disclose that data to recipients through a clean room. In such cases, the final measures should allow stronger technical treatment, aggregation, suppression, regulator-supervised review, or privacy-tested model/output access, provided that any model, feature, benchmark, aggregate or other artefact disclosed to a recipient is itself tested against memorisation, extraction, reconstruction and re-identification risks.

Draft clause 2 - Suppression, transparency, and challenge

Alphabet shall provide eligible recipients with non-sensitive categorical reason codes and aggregate statistics for suppression, generalisation, or routing to a controlled-access modality. Recipients shall have a confidential challenge route for cases where they consider that suppression or routing materially degrades the contestability value. The Commission, and where appropriate, competent supervisory authorities or trusted experts, may review contested categories under confidentiality.

Draft clause 3 - Recipient Certification

A recipient shall certify and obtain independent assurance in relation to permitted OSE use, no re-identification, no record-level linking, no circumvention, separation from auxiliary datasets, access controls, retention, deletion, breach response, processors, AI use, model/data lineage and no onward transfer. The certification framework shall be standardised and proportionate, with pooled or Commission-approved assurance mechanisms where necessary for micro, small and medium-sized enterprises.

Draft clause 4 - Conditional Compliance Reliance

Where Alphabet has implemented the technical, contractual, organisational, audit, notification and suspension controls specified in the final measures, and has not known and could not reasonably have known that a recipient would defeat or misuse those controls, downstream unlawful use, re-identification attempts, unauthorised onward transfer or breach by that recipient shall be treated as the responsibility of the recipient, without prejudice to Alphabet's responsibility for its own processing and for any failure to implement or enforce the mandated controls.

8. Conclusion

The Commission can make Article 6(11) work. But it will not work if the final decision turns anonymisation into a label placed on a daily record-level export. It will also not work if the only way to achieve anonymity is to remove so much useful search signal that the access right becomes nominal rather than contestability-enhancing.

The better approach is a governed access regime. It keeps the Commission's enforcement ambition intact but makes the implementation more honest: privacy risk is tested; utility is measured; recipient controls are verified; AI uses are ring-fenced; suppressions are reviewable; and responsibility is allocated to the actor that controls the relevant processing.

That is not a privacy veto. It is not an Alphabet veto. It is not a data free-for-all. It is the architecture Article 6(11) needs if it is to survive contact with EU data protection law, smaller search engines and the reality of modern AI systems.

Final position
Do not weaken Article 6(11).
Do not pretend that search logs become anonymous by formula.
Do not lower anonymisation to preserve utility.
Once anonymisation is achieved, preserve contestability through a tiered, evidence-led and independently reviewable access regime.

Annex A - Concordance with the Commission's Preliminary Measures

This annex provides a condensed paragraph-by-paragraph implementation map. It is designed to enable the Commission to determine what should be retained, amended, or supplemented in the current preliminary measures.

Preliminary measure paragraphs	Retain	Amend / supplement
1-2 Eligibility	Retain eligibility for undertakings that provide an OSE.	Where a broader AI or chatbot service claims eligibility, it requires function-specific OSE certification and ring-fencing.
3 Parity principle	Do not retain parity as the governing principle.	Replace with the statutory test of effective access to contestability value on FRAND terms, using the least risky access modality that preserves that value.
4-5 Query data	Retain broad query, modification and metadata scope.	Apply field-level tiering; common-tail export, rare-tail controlled access, suppression/escrow for the highest risk.
6-7 View data	Retain URLs, visual content and result-type information.	Assess risk when paired with sensitive queries; use controlled access for sensitive vertical pairings.

8-9 Click data	Retain clicks, click-backs and interaction signals as core utility fields.	Binning, sequence limits, and controlled access for high-risk interaction traces.
10-11 Ranking data	Retain ranking and position data.	Share broadly where query passes; protect combinations with rare tail/session data.
12-13 Search sessions	Retain limited session value.	Richer sequence data should be clean-room for low-risk data or controlled API only.
14 Invalid traffic	Retain exclusion.	Provide non-confidential methodology and aggregate rates to avoid under- or over-exclusion.
15-18 Frequency, method, duration	Retain API, new data access and at least five years.	Allow tier-specific frequency and modality where specified by risk/utility criteria.
19 Two-layer anonymisation	Retain technical plus contractual structure.	Clarify that technical alteration is indispensable and that contractual controls complement it only within a verified recipient environment.
20 Daily record-level	Retain for lower-risk data.	Do not make it universal; define record export as one access modality.
21 Attribute suppression	Retain direct identifier, timestamp, image and click-back transformations.	State explicitly that baseline suppression is necessary but not sufficient.
22-27 Allowlist, thresholds, query suppression	Retain threshold logic.	Require calibration by sensitivity, language, market, seasonality, signed-out bias and utility loss; acknowledge that over-anonymisation is preferable to under-anonymisation.
28-31 Metadata generalisation	Retain cohort and location generalisation.	Add dynamic geography controls and local-search utility assessment.
32-33 Mini-sessionisation	Retain limited refinements.	Controlled access for richer session structures; no recipient re-sessionisation.
34 All other data unaltered	Do not retain as drafted.	Replace with field-level AUIA and access-tier classification.
35-37 Roles and responsibilities	Retain the need to allocate responsibility between Alphabet and recipients.	Clarify that contractual labels do not determine GDPR roles. Recipient obligations should be Commission-specified access conditions, not merely private contractual promises. The final measures should avoid creating uncertainty about joint controllership, private enforcement and the public-law status of the Search Dataset.
38-43 Recipient obligations	Retain no-link, no-re-identification, purpose limitation, retention, no onward sharing, governance and traceability.	State expressly that these obligations do not substitute for anonymisation. They are safeguards and breach triggers. Add active-attack prohibitions, dataset-influence prohibitions, AI-assisted linkage, no-general-training, no model grounding outside the OSE function, model lineage, auxiliary-dataset separation and enforceable deletion/isolation of materially derived artefacts.
44-51 Security controls	Retain encryption, access governance, no local copies, logs, erasure and breach reporting.	Including derived datasets, embeddings and model leakage where materially influenced.
52-70 Assurance	Retain Level 1 and Level 2 independent assurance.	Add pooled auditor, standard controls, SME proportionality and utility/audit annex.
71-84 FRAND pricing	Retain cost-based pricing and SME protection.	Add modality-specific pricing and prevent safe access from becoming prohibitively costly.
85-98 Samples A, B, C	Retain pre-acquisition testing.	Add synthetic privacy/memorisation tests; apply tiering to Sample C.
99-104 Finalisation and licences	Retain timelines and template transparency.	Use phased implementation and standard clauses for AI, no-join, audit and deletion.
105-112 Request and initial eligibility process	Retain public webpage, one-week response, limited requests, formal completeness review.	Add standard recipient risk/use declaration without trade secrets.

113-124 Level 2, change of control, insolvency	Retain continuity protections.	Add material changes in AI use, processors, auxiliary data and security posture.
125-139 Refusal, suspension, termination	Retain reasoned decisions, cure paths and expedited termination.	Add triggers for re-ID attempts, AI misuse, linking, onward transfer, security incidents and false assurance.
140-141 Financial penalties	Retain proportionality.	Clarify penalties for specified serious breaches that threaten anonymisation or integrity, not for ordinary processing discretion.
142-145 Notifications	Retain Commission and DPA notifications.	Use as a structured review channel for patterns and high-risk disputes.

Annex B - Minimum Recipient Certification Checklist

Area	Certification question	Evidence
Purpose	Which OSE function will the data improve?	Use-case statement mapped to crawling, indexing, ranking, retrieval, SERP interaction, query understanding or evaluation of the OSE function.
No-join	How will Article 6(11) data be separated from ads, analytics, identity, browser, mobile, location, panel, publisher and CRM datasets?	Architecture diagram, access-control policy and separation attestations.
Security	Who can access the data and under what controls?	Least privilege, MFA, encryption, logging, monitoring and incident response.
Retention and deletion	How long is data retained, and how is deletion verified?	Retention schedule, deletion certificates and processor commitments.
AI use	Will data be used for retrieval, ranking or evaluation of the OSE function, and how will general model training, fine-tuning and grounding of non-OSE systems be excluded?	Model lineage, no-general-training controls, output checks, extraction/memorisation tests.
Onward transfer	Which processors or affiliates can access data?	Processor register, equivalent safeguards and audit rights.
Audit	How can compliance be verified without exposing the recipient's trade secrets?	Independent assurance report, control mapping and Commission-viewable summary.

Annex C - Direct Engagement with the Preliminary Measures

This Annex follows the Commission's own structure. The recommendations are designed to be administrable: retain the core structure, add an evidence layer, distinguish access modalities by risk, and create reviewable rather than discretionary controls.

C.1 Eligibility: paragraphs 1-2

Paragraphs 1-2 should be retained to the extent that eligibility is anchored in the provision of an online search engine in the EU and, following EEA incorporation, the EEA. The difficult issue is not the categorical exclusion of AI chatbots, but over-inclusion. A broader AI service should not receive Article 6(11) access merely because one part of the service has OSE functionality [n. 1, at paras. 1-2; n. 2, at p. 1].

The final measures should, however, make eligibility functional and purpose specific. A recipient should not receive Article 6(11) data merely by attaching a search-like feature to a wider AI, advertising, analytics, browser,

mobile or identity ecosystem. It should identify the OSE function for which data is requested: crawling, indexing, ranking, retrieval, query understanding, SERP interaction, autocomplete, spelling, search-function evaluation.

For AI chatbots and broader AI services, the final measures should explicitly state that access is limited to the OSE function only. Search-like functionality should not convert Article 6(11) into service-wide access for the entire AI product. General foundation-model training, fine-tuning of non-OSE systems, model grounding outside the OSE function, ad-tech enrichment, identity-graph improvement, profiling, unrelated product optimisation and irreversible model absorption should not be permitted uses.

Proposed Addition after Paragraph 2

A third-party undertaking providing an AI chatbot or broader digital service shall be eligible only to the extent that the requested access is necessary for the development, improvement or optimisation of its online search engine functionality. Eligibility shall not extend to general-purpose model training, fine-tuning of non-OSE systems, model grounding outside the OSE function, advertising enrichment, identity-graph improvement, profiling, unrelated product optimisation or onward transfer outside the permitted OSE function.

C.2 Data scope and conditions of sharing: paragraphs 3-18

The parity principle in paragraph 3 should not be retained as the governing standard. Article 6(11) requires effective access on FRAND terms; it does not require raw-form parity with every internal Google use of data. The case summary confirms the Commission's broad data-scope approach. Still, that approach should be disciplined by the statutory standard of effective contestability value and by the anonymisation requirement [n. 2, at pp. 1-2].

The problem is not the breadth of statutory ambition. The problem is the assumption that a broad scope is best implemented through a single access form. Paragraphs 4-11 list fields that vary sharply in privacy risk and utility. Query text, modified queries, query modifiers, access point, input method, location, click timing, scrolling, swipes, SERP visibility and mini-session structure should not be treated as a homogeneous package.

Paragraphs 15-17 require that the frequency and method of sharing be on par with Alphabet's internal access, and that new data be delivered via an API rather than repeated full refreshes [n. 1, at paras. 15-17]. That is sensible for operational effectiveness. The final measures should add that frequency and method may vary by access tier, subject to the Commission's specified objective risk and utility criteria. Daily access can be the default for lower-risk common-tail records. Delayed, aggregated, or controlled access may be necessary for high-risk, rare-tail, or sequence-rich data.

Field/category	Search utility	Privacy and misuse risk	Recommended treatment
Query input and modified query	Very high: intent, spelling, reformulation, autocomplete, entity understanding.	Very high: semantic self-identification, personal data embedded in queries, health/legal/political distress, exact addresses.	Export common-tail; controlled access for rare, entity-rich, local or sensitive queries; suppress or escrow highest-risk categories.
Query modifiers and filters	High: reveals user intent and refinement path.	High: free text filters may contain personal or sensitive details.	Apply the same thresholds as query text, reason-code suppression, and controlled access for high-value categories.

Timestamp, language, device, access point, input method	High: freshness, market, device and modality optimisation.	Medium to high: linkability, small cohorts, AI/multimodal uniqueness.	Coarsen and calibrate by language/market; daily only where cohort risk is low.
Location/S2 cell	Very high: local search and market relevance.	Very high in small populations, rural areas, events and minority-language contexts.	Dynamic geography controls/local aggregate access where export would over-suppress.
View data: URLs and SERP visual content	High: candidate sets, verticals, SERP formats, ad/organic mix.	Medium to high: Paired with a query can expose interests and sensitive verticals.	Generally, share where query passes; consider controlled access for sensitive verticals and high-risk query pairings.
Click, click-back, hover, scroll, swipe, expansion.	Very high: relevance, satisfaction, presentation bias.	High: behavioural sequence and inference risk.	Binning and aggregation; controlled access for fine-grained interaction sequences.
Ranking and position data	Very high: de-biasing clicks, quality evaluation.	Lower in isolation, higher when paired with query and click traces.	Share broadly with transformed query records; maintain safeguards when combined with rare-tail data.
Search sessions/mini sessions	Very high: query evolution and intent continuity.	Very high: sequence re-identification and sensitive intent chains.	Limited export; richer sequence access is available only in the controlled tier, with no further sessionisation.

Proposed Addition after Paragraph 3

The final measures shall implement Article 6(11) by reference to effective access to contestability value on FRAND terms, using the least-risky access modality that preserves that value once the anonymisation requirement has been met. Where a data field or data slice cannot be made available through unrestricted record-level export without creating a residual likelihood of identification or misuse that cannot be rendered insignificant, Alphabet shall provide an alternative access modality specified in the final measures, including aggregate access, controlled API access, clean-room access for low-risk data, trusted execution or regulator-supervised escrow.

C.3 Technical Anonymisation: Paragraphs 19-34

Paragraph 19 correctly distinguishes between technical and contractual measures. That distinction should be preserved. Contracts cannot anonymise a technically weak dataset. They can, however, help define the recipient environment in which reasonably likely means of identification are assessed, provided the controls are durable, verifiable and enforceable [n. 1, at para. 19; n. 4].

Paragraph 20 requires daily, record-level sharing after technical measures [n. 1, at para. 20]. This should be amended. Daily record-level sharing should remain available for lower-risk data. It should not be the universal baseline for every field and every slice of data. The final measures should make record-level export one access modality among several, not the definition of compliance.

Paragraph 21 removes direct identifiers, precise timestamp, selected SERP layout geometry, image queries and fine click-back time [n. 1, at para. 21]. This is necessary, but not sufficient. Search data remains identifying through the content of queries, entity combinations, location, language, device/access point, URLs viewed, click behaviour and mini-session context.

Paragraphs 22-27 create an allowlist and query suppression regime based on a 13-month EEA signed-in-user reference window, a more-than-50 signed-in-user entity threshold, language-specific length thresholds and daily

checks of modified, initial and refiner queries [n. 1, at paras. 22-27]. This is a defensible starting point, but the final measures should be empirically calibrated. The number 50 should not become a talisman. It may be under-protective of sensitive entities and over-suppressive of minority languages, smaller Member States, local queries, emerging events, or seasonal topics.

Paragraphs 28-31 generalise metadata, including location as a country plus an S2 cell with at least 1,000 signed-in users and at least 3 km2, and require at least 50 signed-in users with the same language, location, and device values [n. 1, at paras. 28-31]. Again, the structure is sensible but should not be frozen without utility and privacy curves in place. Local search utility is one of the places where rivals may most need scale. If the location signal is over-generalised, a nominally shared dataset may not serve Article 6(11). If it is under-generalised, re-identification risk rises sharply.

Paragraphs 32-33 create mini-sessionisation for bounded query refinements and overwrite location and device metadata with the last record in the mini-session [n. 1, at paras. 32-33]. Mini-sessionisation is both valuable and dangerous. It preserves continuity of search intent while also creating linkage. The final measures should provide only a limited number of exportable mini-sessions and route richer session-like structures to controlled access.

Paragraph 34 is the most problematic technical sentence. It provides that after the five steps, Alphabet shall include the search records in the Search Data and include all other data as defined in section 2.1 unaltered, without applying additional measures [n. 1, at para. 34]. This should be amended. Some view, click, ranking, URL, visual and interaction data may be identifying or inferential when paired with query content and metadata. The final measures should require field-by-field testing, not an unaltered-data residual category.

Technical element	Retain	Amend or supplement
Record-level daily export	Retain for low-risk common-tail data.	Add tier-specific frequency and modality.
Attribute suppression	Retain account/IP/device/direct identifier removal and precise-time suppression.	State explicitly that this is a baseline, not sufficient anonymisation; acknowledge that over-anonymisation is preferable to under-anonymisation.
Image query placeholder	Retain image replacement.	Add multimodal risk assessment for text, images, and access-point metadata.
Click-back time bins	Retain as baseline.	Test whether bins preserve ranking/satisfaction utility and whether additional interaction sequences require controlled access.
Allowlist >50 signed-in users	Retain the principle of entity-frequency filtering.	Calibrate by language, geography, sensitivity, seasonal events, signed-out bias and entity class.
95% length threshold	Retain as one signal.	Do not treat length as a proxy for identifiability; add semantic rarity and entity-risk tests.
Metadata generalisation	Retain generalisation sequence.	Publish non-confidential utility/risk curves; dynamic controls for local queries.
Mini-sessionisation	Retain limited refinements.	Richer sequences are only available through controlled access or escrow.
All other data unaltered	Do not retain in the current form.	Replace with field-by-field AUIA and access-tier classification.

Proposed replacement for paragraph 34

Alphabet shall include in the Search Dataset only those records and fields that have passed the technical measures and the field-level risk and utility assessment specified in the Anonymisation and Utility Impact Assessment. No category of query, view, click, or ranking data shall be treated as unaltered by default where, when combined with retained fields or recipient-held auxiliary data, it materially increases the likelihood of singling out, linkability, or inference.

C.4 Contractual Measures, Roles and Independent Assurance: Paragraphs 35-70

Paragraphs 35-37 define Alphabet and third-party OSEs as independent controllers, allocate pre-sharing processing to Alphabet, allocate responsibility for personal data embedded in queries in the Search Dataset to the recipient, and limit Alphabet to contractual measures necessary to complement anonymisation [n. 1, at paras. 35-37]. This structure should be retained only if it is made legally sharper. The final measures should not give the impression that contractual labels determine GDPR roles. If data is personal data from the recipient's perspective, the recipient is a controller for its processing. Conversely, if the final measures require Alphabet to impose detailed restrictions on the purposes, means, onward use, security, derived models and deletion practices of recipients, the allocation of responsibility must be made legally certain rather than left to contractual drafting.

The contractual layer is therefore not a harmless add-on. It is one of the legally most delicate parts of the preliminary measures. The Commission should not design a regime in which the public-law condition of anonymisation depends on Alphabet privately enforcing contractual promises against recipients. That would move the problem from anonymisation to private enforcement. It would also create uncertainty about whether breach of contract is sufficient to make subsequent recipient processing unlawful, whether data subjects can meaningfully enforce the restrictions, whether DPAs can act without first resolving the contract position, and whether Alphabet is exposed to allegations of influence over recipient processing.

The final measures should say more plainly that contractual restrictions are access conditions and enforcement triggers. They may support a technically robust anonymisation design by limiting the recipient environment and reducing the means reasonably likely to be used. But they cannot transform a technically fragile dataset into anonymous data. If the technical layer does not render the data anonymous in the relevant recipient context, the contractual layer cannot cure the legal status of the data. At most, it creates obligations that may be breached after the privacy harm has already occurred.

The treatment of personal data contained in query text is especially important. Search queries may include a name, address, workplace, school, allegation, medical term, or other personal data element. In many cases, there will be no reliable way to determine whether that personal-data element relates to the end user who issued the query or to another person. For Article 6(11) purposes, the safer assumption is that personal data contained in the query may relate to the end user and must be treated as such in the anonymisation analysis. Paragraph 36(c) recognises responsibility for personal data concerning individuals other than end users, but the final measures should not allow that point to become a way of excluding embedded query data from the Article 6(11) anonymisation problem. The practical answer is stronger technical detection, suppression classes and direct recipient responsibility.

Paragraphs 38-43 prohibit record-level linking, re-identification, augmentation, circumvention, auxiliary advertising/analytics access, unrelated-purpose use, excessive retention and onward sharing, and require governance and traceability for automated processes and derived datasets or models [n. 1, at paras. 38-43]. These restrictions are necessary, but they should be recast as Commission-specified access conditions. They should not be

presented as part of the mechanism by which weak technical treatment becomes anonymisation. A prohibition on re-identification is not the same thing as the absence of means reasonably likely to identify. A prohibition on linking is not the same thing as technical unlinkability. A prohibition on augmentation is not the same thing as preventing inference.

The final measures should also address active and AI-assisted attacks expressly. Recipient restrictions should cover canary-style attacks, deliberate dataset influence, query injection, record recognition, AI-assisted entity resolution, model-mediated inference, session reconstruction, reconstruction of suppressed or generalised attributes, and attempts to infer whether a particular user or entity appears in the Search Dataset. These risks are not exotic. They arise precisely because search data is semantic, repeated, and capable of being influenced by ordinary lawful search behaviour.

For AI, governance and traceability should include model lineage, no-general-training commitments, no fine-tuning of non-OSE systems, no model grounding outside the OSE function, no irreversible absorption into general-purpose models, extraction and memorisation testing, and deletion or isolation logic for materially derived artefacts. Without this, Article 6(11) risks becoming a route for general AI development under the language of search contestability.

Paragraphs 44-51 impose integrity, confidentiality, encryption, access governance, no local copying, physical security, logging, erasure and breach reporting controls [n. 1, at paras. 44-51]. These should be retained, but the final measures should make clear that these are safeguards against misuse and leakage, not proof of anonymisation. The security layer should include derived datasets, embeddings, model outputs and other materially influenced artefacts where those artefacts could leak, reconstruct or encode Search Dataset information.

Paragraphs 52-70 establish a detailed independent assurance mechanism, including Level 1 pre-access assurance and Level 2 annual assurance under ISAE 3000 or an equivalent standard, with four assurance objectives [n. 1, at paras. 52-70]. This is a strong feature of the proposal, but it should be more substantive. Assurance should not become a formal paper exercise in which a recipient confirms that policies exist. It should test whether the recipient can actually separate Article 6(11) data from auxiliary datasets, prevent linking, prevent AI absorption, detect prohibited access, delete or isolate derived artefacts, and withstand realistic re-identification and extraction attempts. Where smaller OSEs cannot bear the cost of bespoke assurance, the solution is a pooled auditor, a standard control library and Commission-approved certification, not weaker controls.

Text box C1 - Contracts cannot anonymise, but verified context controls matter

Contracts can police a recipient. They cannot anonymise a dataset that remains identifiable by lawful and reasonably likely means. If the technical layer fails, the contractual layer cannot rescue the legal status of the data. It merely creates a private enforcement problem, role uncertainty and the possibility that misuse will be addressed only after the privacy harm has occurred. The final measures should therefore treat contractual obligations as access conditions, assurance requirements and breach triggers, not as substitutes for anonymisation.

Contractual / assurance area	Retain	Amend or supplement
Recipient responsibility	Retain independent-controller allocation.	Clarify responsibility for personal data embedded in queries and recipient misuse.

No re-identification / no linking	Retain and strengthen.	Include AI-assisted entity resolution, canary attacks, session reconstruction and model-mediated inference.
Purpose limitation	Retain the OSE optimisation purpose.	Define permitted OSE functions and prohibit general model training, ads enrichment, identity graphs and unrelated optimisation.
Access separation	Retain ads/analytics separation.	Extend to browser, mobile, location, CRM, publisher, panel and identity datasets where relevant.
Governance and traceability	Retain.	Require model/data lineage for materially derived systems and AI outputs.
Assurance reports	Retain Level 1/Level 2.	Add the standard control library, the pooled auditor option, SME proportionality, and utility/audit annexes.
Audit scope	Retain organisational and technical scope.	Ensure derived models and high-risk AI uses are in scope where materially influenced by Search Data.

C.5 FRAND pricing: paragraphs 71-84

Paragraphs 71-84 set out a cost-based FRAND pricing model: incremental costs of making data available plus a reasonable return, generally capped by Alphabet's WACC, with exceptions for certain large beneficiaries and gatekeepers, SME protections, five-year pricing application, eligible cost categories, cost allocation, transparency and dispute routes [n. 1, at paras. 71-84]. This structure should be retained.

The missing issue is pricing by access modality. A common-tail record API, a privacy-preserving aggregate layer, a controlled API, a clean-room for low-risk data environment, Sample B synthetic data, Sample C representative data and regulator escrow do not have the same cost structure. The final measures should expressly state that FRAND pricing may reflect access-tier costs, if this does not make safer access prohibitively expensive for the smaller recipients. Article 6(11) is meant to help.

The Commission should also prevent assurance costs from becoming a hidden access price. A pooled auditor, a standard control library, Commission-approved certification templates, and proportionate reporting for SMEs would reduce the risk of privacy compliance becoming a new barrier to entry.

Proposed addition to the FRAND pricing section

Where the final measures specify different access modalities, compensation shall be calculated by reference to the efficiently incurred incremental costs of the relevant modality. The pricing method shall not create a disincentive for the use of safer access modalities or place a disproportionate burden on micro, small or medium-sized recipients. Standardised assurance, pooled audit and shared certification mechanisms should be available where necessary to preserve FRAND access.

C.6 Pre-Acquisition testing, Search Dataset finalisation and licensing: paragraphs 85-104

Paragraphs 85-98 provide for three test samples: Sample A, a 1,000-row sample; Sample B, a downloadable synthetic dataset of up to 10 million representative artificial queries and metadata; and Sample C, a downloadable 5% sample of the final Search Dataset, with Member State, language and time-period selections [n. 1, at paras. 85-98]. The sampling architecture is useful and should be retained with amendments.

Sample B should be treated as an integration and schema tool, not as proof that the real Search Dataset is safe or useful. Synthetic data may leak memorised training examples or distort rare-tail distributions if not tested. The final

measures should require Alphabet to publish a non-confidential synthetic generation method, memorisation/leakage tests and limits on what Sample B can prove.

Sample C is more sensitive. A 5% representative sample, selected over a period of one month to one year and downloadable to recipient servers, may itself constitute a substantial exposure of real behavioural data. The final measures should ensure that Sample C is subject to the same field-level tiering and AUIA logic as the full Search Dataset. For high-risk slices, Sample C should be provided through controlled access rather than unrestricted download.

Paragraphs 99-104 require the Search Dataset to be finalised within three months of the implementing act, licence templates within two months, templates to be shared within one working day, and FRAND conditions not to impede effectiveness [n. 1, at paras. 99-104]. These timelines are ambitious but workable if the final measures provide a phased implementation model rather than requiring every high-risk access tier to be fully mature at the same moment.

Process element	Retain	Amend or supplement
Sample A	Retain a free small sample.	Use for schema demonstration only; do not treat as evidence of utility or safety.
Sample B	Retain a synthetic downloadable sample.	Add synthetic-data privacy tests, memorisation tests and methodology transparency.
Sample C	Retain the representative testing concept.	Do not require unrestricted download for high-risk slices; apply AUIA/tier logic.
Three-month Search Dataset finalisation	Retain with phased access.	Common-tail export and aggregate layer can launch first; rare-tail controlled access can be piloted.
Licence templates	Retain one-day transparency after finalisation.	Templates should include a standard AI clause, a no-join schedule, an audit schedule and a deletion certificate.

C.7 Eligibility Review, Continuity, Refusal, Suspension, Termination, Penalties and Notification: paragraphs 105-145

Paragraphs 105-112 require Alphabet to engage with requesting third parties, maintain a public webpage, respond within one week, provide request forms, describe datasets and samples, explain eligibility and assurance requirements, limit requested information to what is strictly necessary, restrict Alphabet's review of Level 1 reports to formal completeness, provide access within two weeks after receiving the Level 1 report and signed licence, and inform the Commission of requests [n. 1, at paras. 105-112]. This is pro-contestability and should be retained.

The final measures should add a recipient risk and use-case declaration to the request form. This should not ask for trade secrets. It should ask for the requested OSE function, auxiliary data separation, AI use, processors, access environment, and proposed data tier. The information should be limited, standardised and auditable so that Google cannot use it as a fishing expedition.

Paragraphs 113-124 address Level 2 assurance, change of control and insolvency [n. 1, at paras. 113-124]. These are useful continuity provisions. They should be supplemented by rules governing material changes in AI use, data architecture, processors, auxiliary data access, security posture, and corporate group data sharing. A change in how Article 6(11) data is absorbed into an AI system can be as important as a change of control.

Paragraphs 125-139 provide mechanisms for refusal, suspension, restoration, termination, and expedited termination, including expedited termination where there is an urgent risk of serious and irreparable damage to the anonymisation of end users' personal data [n. 1, at paras. 125-139]. The structure is sound. The trigger list should

be expanded to include attempted re-identification, unlawful record-level linking, unauthorised onward transfer, breach of no-general-training restrictions, material security incident, false assurance statements, refusal to permit audit and failure to delete or isolate materially derived datasets.

Paragraphs 140-141 require proportionate financial penalties for certain failures but state that Alphabet shall not impose financial penalties in relation to the recipient's processing of the Search Dataset [n. 1, at paras. 140-141]. That distinction should be clarified. Alphabet should not have open-ended policing powers over ordinary recipient processing. But serious breaches that threaten anonymisation or the integrity of the regime - re-identification attempts, unauthorised linking, onward transfer, failure to delete, materially false assurance, or prohibited AI absorption - should be capable of contractual sanction under Commission-specified terms.

Paragraphs 142-145 require notification to the Commission and lead supervisory authority of refusal, suspension, termination, restoration and expedited termination decisions [n. 1, at paras. 142-145]. This is important and should be retained. It should become part of a structured review channel: the Commission, relevant DPAs and, where appropriate, trusted experts should be able to review patterns of suppression, suspension, breach and contested access without turning each event into a bilateral dispute.

The suspension and termination provisions are especially important because contractual obligations will otherwise operate too late. If a recipient attempts re-identification, links the Search Dataset to auxiliary data, transfers it onward, absorbs it into a prohibited AI system, or gives false assurance, the issue is not merely breach of a private bargain with Alphabet. It is a threat to the integrity of the Article 6(11) regime and to the premise that the Search Dataset is accessed only under conditions that preserve anonymisation. The final measures should therefore treat serious contractual breaches as regulatory events requiring rapid suspension, Commission notification and, where relevant, DPA notification.

Proposed addition to sections 5.5 and 5.6

Suspension or expedited termination should also be available where there is credible evidence of attempted re-identification, unlawful record-level linking, unauthorised onward transfer, breach of AI-use restrictions, material security incident, materially false assurance, refusal to permit audit, or failure to delete or isolate Search Dataset material and materially derived datasets after termination.

Annex D - Source marker register

Marker	Source	Use in submission
[n. 1]	European Commission, Case DMA.100209 - SP - Alphabet - Article 6(11), Preliminary Measures, 16 April 2026.	Primary source for paragraph references, technical measures, contractual measures, FRAND pricing and processes.
[n. 2]	European Commission, Case DMA.100209 - SP - Alphabet - Article 6(11), Case Summary, 16 April 2026.	Summary of eligibility, data scope, anonymisation, pricing and process.
[n. 3]	Regulation (EU) 2022/1925, Digital Markets Act, Article 6(11), Article 8(2) and Recital 61.	Legal basis for access to ranking, query, click and view data; mandatory anonymisation of personal query, click and view data; Commission specification powers and the relationship between usefulness and anonymisation.
[n. 4]	European Commission / European Data Protection Board, Joint Guidelines on the interplay between the DMA and the GDPR, public consultation version.	Relationship between Article 6(11), usefulness and anonymisation; role of technical, organisational and contractual safeguards.

[n. 5]	Regulation (EU) 2016/679, GDPR, Article 4(1), Article 4(7), Recital 26 and relevant Court of Justice materials, including Breyer, EDPS v SRB and recipient-context case law.	Definition of personal data; controller concept; actor-relative/contextual identifiability; means reasonably likely to be used; contractual labels not being dispositive.
[n. 6]	European Commission Digital Omnibus materials, including proposed clarification mechanisms concerning pseudonymisation and identifiability, and the EDPB/EDPS Joint Opinion on Digital Omnibus.	Current debate about pseudonymisation, personal data, implementing-act criteria and the risk of divergent operational meanings of anonymisation across the EU digital acquis.
[n. 7]	EDPB Guidelines 01/2025 on Pseudonymisation and relevant DPA guidance, including Berlin DPA/DSK materials.	Pseudonymisation domains, context controls and practical pseudonymisation/anonymisation guidance.
[n. 8]	Article 29 Working Party Opinion 05/2014 on Anonymisation Techniques; CNIL and ICO anonymisation guidance.	Singling out, linkability, inference and context-based anonymisation assessment.
[n. 9]	EDPS/AEPD, 10 misunderstandings related to anonymisation.	Limits of simplistic anonymisation assumptions.
[n. 10]	Academic work on search-query privacy, differential privacy, high-dimensional de-anonymisation and search-log release.	Technical background on query-log risk and utility.
[n. 11]	NIST SP 800-226 and differential privacy literature.	Use of differential privacy for aggregate layers and repeated query interfaces.
[n. 12]	EDPS, Towards a Digital Clearinghouse 2.0; DMA High-Level Group materials.	Cross-regulatory coherence for DMA, GDPR, DSA, AI and consumer protection overlap.
[n. 13]	DSA Article 40 data-access materials and Five Safes / trusted research environment practice.	Safe people, projects, settings, data and outputs as access-governance vocabulary.