AI Visibility in South East Queensland Legal Services: An Exploratory Observational Study

Methodology Document — Version 2.4 (Pilot Complete; Capture Rules Finalised; Registration Version)

Prepared by: Jane Cluff, Felicity Jane Digital **Date:** August 2026 **Status:** Version 2.4 — registration version. The pilot specified in this document was completed on 3 August 2026; this version incorporates the operational capture rules arising from it (Revision Record, v2.4) with no change to study design, scope, queries, run counts, or analysis plan. Version 1 was externally reviewed by Dr Karen Sutherland (University of the Sunshine Coast, July 2026); this document incorporates the revisions made in response to that review. Dr Sutherland has not reviewed this consolidated document; her endorsement, if given, is gated on review of the final report. **AI assistance disclosure:** This document was drafted with AI assistance (Claude, Anthropic) and reviewed and edited by Jane Cluff, Felicity Jane Digital. All methodological decisions, research propositions, and content remain the sole responsibility of the researcher. Claude (the AI assistant) had no role in data collection, and Claude (the platform under study) is treated identically to the other three platforms in the data collection protocol. **Conflict of interest disclosure:** This research was conducted by Felicity Jane Digital, a specialist legal SEO agency serving Queensland law firms. The findings may inform FJD's commercial service offerings. The researcher has endeavoured to apply objective methodology and transparent disclosure throughout. Readers should consider this context when evaluating the findings.

Revision Record

Version 1 (July 2026): Draft submitted to Dr Karen Sutherland for external methodology review.

Version 2 (July 2026): Revised in response to that review. Changes:

1. **Run count increased** from 3 runs per query per platform to 5 runs per query per platform, repeated in two waves — 10 total runs per query per platform. The reviewer advised that three repetitions is not enough to produce reliable results given the high variability of AI responses, and recommended a minimum of five to ten runs per query per platform in two or three waves.
2. **Confidence intervals added.** Mention frequencies are now reported as proportions with 95% Wilson score confidence intervals rather than single percentages, per the reviewer's recommendation.
3. **Pre-registration added.** The full prompt set and analysis plan will be registered on the Open Science Framework (OSF) before any data is collected, per the reviewer's recommendation. A mirror PDF is hosted on the FJD website and linked from the published findings.
4. **Per-run documentation formalised.** Model versions, dates, account settings, and search-mode conditions are logged for every run using the template at Appendix B, per the reviewer's recommendation.
5. **Correlation framing strengthened.** The signal analysis is explicitly framed as correlation, not cause, throughout — the reviewer noted that established firms will score well on everything, so shared signals cannot be read as drivers of citation.
6. **Scope reduced from four locations to three.** Ipswich removed to fund the increased run count within solo-researcher capacity. Disclosed as a limitation (Section 11).

Version 2.1 (July 2026): Finalisation and source-verification pass prior to OSF pre-registration. Changes:

1. **Query set finalised (Appendix A).** All nine query strings are now locked in exact wording following keyword-data-informed term selection (Step 1). The placeholder table is replaced with the final strings. The term-selection rule, the sequence disclosure, and the commercial-litigation term decision are recorded in the companion document *fjd-research-step1-query-lock*, which is registered alongside this methodology.
2. **Inter-wave gap corrected to a minimum of three weeks** throughout (previously stated as two weeks in places), consistent with the pre-registered design.
3. **Sources verified against primary sources.** Every third-party claim was checked before registration. Three corrections resulted: (a) the Claude B2B-referral figure was re-attributed from "Quratic" to its actual source, the Goodie 2026 AI Search Traffic Report; (b) the repeated-measurement study previously cited as "St. Gallen (Schulte, Bleeker and Kaufmann)" was corrected to Schulte et al., *Don't Measure Once: Measuring Visibility in AI Search (GEO)*, arXiv 2604.07585; (c) the Sielinski citation was corrected to the verified paper (arXiv 2603.08924) and the previously cited "33–94 answers to stabilise" figure — which derives from secondary coverage of a different Sielinski paper — was removed.
4. **Authoritas citation removed.** The unverifiable "92% in eight weeks" claim was deleted and the temporal-instability

rationale rebuilt on the verified Sielinski and Schulte et al. findings.

5. **Commercial-litigation limitation added** (Section 12) covering the thin coast-market cells.
6. **Reliability control added** (Section 7): intra-rater re-coding of a 10% sample.
7. **Framing strengthened** to state affirmatively that repeated sampling with confidence intervals is the method the peer-reviewed literature establishes as defensible, which this study applies.

Version 2.2 (July 2026): Second source-verification and consistency pass prior to OSF pre-registration. Changes:

1. **SparkToro platform attribution corrected (Section 4).** SparkToro tested ChatGPT, Claude, and Google AI Overviews/AI Mode. Perplexity — not Claude — is this study's addition to the SparkToro platform set. The platform-selection rationale is rewritten accordingly.
2. **Ahrefs citation corrected (Section 10, Proposition 1).** The December 2025 Ahrefs study compared AI Mode against AI Overviews (13.7% citation overlap between the two AI surfaces), not AI citations against organic results. The AI-versus-organic reference point is re-based on Ahrefs' March 2026 analysis (top-10 organic share of AI Overview citations fell from 76% to 38% in eight months; 62% of citations now from outside the top 10).
3. **COSEAL citation corrected** to Eimer et al., *Best Practices For Empirical Meta-Algorithmic Research: Guidelines from the COSEAL Research Network* (arXiv 2512.16491, December 2025), with its cross-domain character noted and "require" softened to "recommend."
4. **Goodie citation reframed (Sections 2 and 4)** to reflect that Wave 1 (1.4%; 41-brand panel, May–August 2025) and Wave 2 (18.5%; fresh anonymised panel, cross-source brand-share average, March–April 2026) used different panels and methods — a directional shift, not a same-panel longitudinal measure.
5. **Aggarwal et al. (GEO, KDD 2024) claim softened (Section 5).** The specific finding previously attributed to the paper could not be verified against it; the query-type rationale now rests on the SparkToro precedent and internal-consistency grounds.
6. **Aggregation example corrected (Section 11)** to a form supportable by the data actually collected under Section 9.
7. **Population capture procedure specified (Section 8):** written, reproducible Google Maps enumeration procedure added.
8. **AI Overview no-trigger rule added (Sections 7, 11; Appendix B).**
9. **Account conditions fixed per platform in advance (Section 7)** rather than logged per run as an open variable; Section 5's "logged out" phrasing amended to match. (The companion query-lock document's protocol summary should be updated to the same per-platform phrasing before registration; its state-anchor rationale is unaffected.)
10. **Signal F directory set revised (Section 9):** LegalVision and Lawpath removed as service providers/marketplaces rather than open third-party directories; score rebased to 0–2.
11. **Sole-practitioner naming rule added (Section 7).**
12. **CI independence assumption disclosed (Section 11).**
13. **Intra-rater re-coding extended to signal scoring (Section 7).**
14. **Unsourced "fastest-growing" market claim softened (Section 6).**
15. **IP-location bias limitation added (Section 12).**

Version 2.3 (July 2026): Final pre-registration revision, incorporating an independent second methodology review (AI-assisted; assessed and selectively adopted by the researcher). Changes:

1. **Schema detection method corrected (Section 9, Signal B; Section 10, Proposition 2).** Google's Rich Results Test detects rich-result eligibility, not schema validity in general, and would systematically under-detect the schema types under study (notably FAQPage, restricted for rich-result eligibility since 2023). Primary detection is now the Schema.org validator (validator.schema.org), with the Rich Results Test retained as a supplementary Google-eligibility check. Proposition 2 reworded from "valid schema markup" to "detectable structured data."
2. **Confidence interval labelling clarified (Section 11):** intervals are described as "approximate 95% Wilson confidence intervals under an independence assumption" at first mention. Analysis unchanged.
3. **Denominator reproducibility language strengthened (Section 8):** the population is reproducible procedurally rather than identically, and this is stated.
4. **Google AI Overviews coupling acknowledged (Section 4; Section 10, Proposition 1):** Google's AI layer is partially coupled to Google's search ecosystem, so the AIO arm of the organic comparison is not fully independent; ChatGPT, Claude, and Perplexity provide the independent comparisons.
5. **Gemini exclusion justified explicitly (Section 4)** rather than left implicit.
6. **Practice-area and platform framing softened (Sections 4, 6):** "primary AI surfaces" → "among the principal consumer AI platforms"; practice areas described as purposively selected rather than representative.
7. **"Best" note extended (Section 5):** one sentence added stating the study measures AI recommendation behaviour, not legal quality, and that platforms may operationalise "best" differently.

8. **Family-law review exclusion reframed (Section 9, Signal E)** as a measurement-reliability rationale rather than a bare confidentiality assertion.
9. **Sunshine Coast boundary rule added (Section 8):** the Sunshine Coast population definition includes Noosa Shire alongside Sunshine Coast Regional Council, matching common usage of "Sunshine Coast"; disclosed.
10. **Terminology tightened throughout:** "AI recommendation appearance" used where recommendation behaviour is meant, reserving "citation" for footnoted-source behaviour and the measurement literature; "Google AI Overviews" used consistently for the study platform (with "AI Mode" named only where a cited study tested it); Signal A relabelled "Appears in Google's top-10 organic results (Yes/No)."
11. **External reviewer attribution:** affiliation wording unchanged pending Dr Sutherland's confirmation of her preferred attribution (personal academic capacity vs institutional identification), which is added to the sign-off request already outstanding.

Version 2.4 (3 August 2026): Pilot-derived operational amendments. The pilot (one query across all four platforms, one run each, timed, data discarded) was completed on 3 August 2026. All four platforms returned recommendation-type responses containing firm names under study conditions; web search operated as specified on all platforms; the AI Overview triggered on the pilot query. The amendments below are confined to operational capture rules in Section 7 and Appendix B; no changes to study design, scope, queries, run counts, analysis plan, or thresholds. Changes:

1. **Session isolation rule added (Section 7):** one dedicated incognito window per run, closed after logging; the run's AI query and its contemporaneous Google check share that window in separate tabs.
2. **AI Overview expansion step added (Section 7):** the Overview is expanded ("Show more") before recording, as a standard step; for AIO runs the same results page serves both the AI response record and the organic top-10 record.
3. **Firms-recorded inclusion/exclusion rules specified (Section 7; Appendix B):** directories and ranking guides cited as sources excluded; sponsored placements excluded; citation-panel-only appearances excluded; firms attached to excluded named individuals included; entity-type ambiguity resolved from the response's own framing.
4. **Named-individuals field format fixed (Appendix B):** count only, never names — "named individuals — excluded (n)".
5. **Organic top-10 capture rules specified (Section 7; Appendix B):** source name as displayed (fallback: domain); blue links only; first appearance per domain; actual count recorded where fewer than ten; AI Overviews appearing on comparison searches ignored and noted.
6. **Model/version strings fixed per platform (Appendix B),** reflecting what each interface displays under study conditions.
7. **Duration field added (Appendix B),** timed from query entry to completion of logging; pilot benchmark approximately 4–5 minutes per run.
8. **Evidence capture formalised (Section 7):** full-page capture of each AI response and comparison results page, filed by Run ID, retained with the run log.

---

1. Research Question

Which observable digital signals are shared by South East Queensland law firms that appear consistently in AI-generated legal service recommendations across ChatGPT, Perplexity, Google AI Overviews, and Claude?

---

2. Study Design

This is an exploratory observational study. It does not test causal hypotheses. It observes patterns in publicly available data and generates propositions for future confirmatory research.

This design was selected for three reasons:

First, the AI recommendation landscape is sufficiently novel in the Australian legal market that no prior Queensland-specific data exists. Exploratory research is the appropriate design where there is, as yet, not enough information to form a well-thought-out hypothesis (Eimer et al., *Best Practices For Empirical Meta-Algorithmic Research: Guidelines from the COSEAL Research Network*, arXiv 2512.16491, December 2025). The COSEAL guidelines originate in the computer-science meta-algorithmics community; they are cited here as a general methodological reference for the exploratory/confirmatory distinction, which they draw from the broader empirical-research literature (Schwab and Held, 2020), not as domain-specific authority for marketing research.

Second, the peer-reviewed and preprint literature has converged on repeated sampling with explicit uncertainty quantification as the defensible method for measuring AI visibility, and this study applies that method rather than the discredited single-run approach. Fishkin and O'Donnell (SparkToro, January 2026) found that AI recommendation lists

rarely repeat — across 2,961 prompt runs (60–100 runs per prompt per platform, 600 volunteers) the chance of the same brand list appearing twice was under 1 in 100. Two independent empirical studies corroborate this instability through repeated measurement and argue directly for confidence-interval reporting: Schulte et al. (*Don't Measure Once: Measuring Visibility in AI Search (GEO)*, arXiv 2604.07585, 2026) measured citation and brand-mention stability across four engines over a 45–46 day window and found day-to-day source overlap averaging only 0.34–0.42 (Jaccard), concluding that a single reading is unreliable and engines must be sampled repeatedly; and Sielinski (*Quantifying Uncertainty in AI Visibility*, arXiv 2603.08924, 2026) demonstrated that citation-visibility metrics are sample estimators of an underlying distribution, that many apparent differences between domains fall within the measurement noise floor, and that visibility must therefore be reported with uncertainty estimates rather than point values. (Sielinski is affiliated with IQRush, which sells AI-visibility measurement software; the finding is treated as methodological support, not as endorsement of any product.) This body of work establishes that repeated sampling with confidence intervals is the appropriate reporting format at any achievable scale. This study adopts it: it makes no ranking claims and reports appearance proportions with Wilson confidence intervals. A confirmatory study claiming statistical significance at the run counts feasible for a solo researcher would overstate its precision; an exploratory study that quantifies and discloses its uncertainty is the honest and defensible design.

Third, the AI recommendation landscape is changing rapidly. The referral-source mix alone shifted materially within months — the Goodie 2026 AI Search Traffic Report reports Claude's share of measurable B2B AI referrals rising from 1.4% (Wave 1: a 41-brand panel, May–August 2025) to 18.5% (Wave 2: a fresh anonymised brand panel, reported as a cross-source average of brand shares and triangulated against SimilarWeb and SensorTower data, March–April 2026). The two waves used different panels and methods, so this is a directional shift across the market rather than a same-panel longitudinal measure; the direction is corroborated by independent trackers — and the citation-instability studies above document meaningful drift over windows as short as several weeks. A point-in-time exploratory observation, with dates disclosed, therefore provides more value than a delayed, over-engineered confirmatory study whose conditions would have changed before completion.

What this study can claim: Directional patterns observed across a defined study period. Observable signals shared by firms that appeared in AI recommendations during that period, reported as correlations with confidence intervals. Propositions for future confirmatory research.

What this study cannot claim: Causal relationships between observable signals and AI recommendation appearance. Statistical significance in the confirmatory sense. Findings that generalise beyond the study period or study scope. Rankings or positions of individual firms.

Correlation, not cause: All signal analysis in this study is correlational. Firms that appear in AI recommendations are likely to be established practices that score well across most or all observable signals simultaneously — organic ranking, schema, GBP completeness, content recency, directory presence. This study cannot separate the contribution of any individual signal from the underlying effect of being an established, well-resourced firm. No finding in this study should be read as "signal X causes AI recommendation appearance."

---

3. Pre-Registration

Before any data is collected, the following materials are registered on the Open Science Framework (osf.io) as a timestamped registration:

1. The full query list (all nine query combinations in their exact wording). The term-selection rule, the sequence disclosure (keyword-volume data were viewed during term selection, which does not affect which firms appear in AI responses), and the commercial-litigation term decision are documented in the companion document *fjd-research-step1-query-lock*, registered alongside this methodology.
2. The complete analysis plan (Section 11 of this document), including the definition of a mention, the consistent-appearance threshold, the confidence interval method, and the signal comparison approach.
3. The per-run logging template (Appendix B).

OSF was selected over self-hosted publication because OSF registrations are timestamped by an independent third party and cannot be edited after registration — the date and content cannot be questioned. A mirror PDF of the registered materials is hosted on the Felicity Jane Digital website and linked, together with the OSF registration URL, from the published findings.

The registration date precedes the first data collection run and is stated in the published findings. No changes are made to the query list or analysis plan after registration. If any deviation becomes unavoidable during data collection, it is documented and disclosed in the published findings as a protocol deviation.

The purpose of pre-registration is to prevent — and to be seen to prevent — post-hoc adjustment of queries, thresholds, or analysis choices to favour a particular result. This is of particular importance given the researcher's disclosed commercial

interest in the findings.

---

4. Platforms

Four platforms were selected: ChatGPT (OpenAI), Perplexity AI, Google AI Overviews, and Claude (Anthropic).

These four were selected because they are among the principal consumer AI platforms through which Australian legal clients are likely to encounter AI-generated recommendations in 2026. Three of the four — ChatGPT, Claude, and Google AI Overviews — are platforms tested by Fishkin and O'Donnell (SparkToro, January 2026; their Google arm used AI Overviews with AI Mode as fallback), providing direct methodological precedent for repeated-run measurement on those surfaces. Perplexity AI was added to that set for two reasons: its native footnoted-URL citation behaviour anchors the schema analysis (see citation-mechanism note below), and it holds a material share of AI referral traffic. Claude's retention is further supported by its growth into the second-largest source of measurable B2B AI referrals — 18.5% in Goodie's Wave 2 (2026 AI Search Traffic Report; cross-source brand-share average across a fresh anonymised panel, March–April 2026), up from 1.4% in Goodie's Wave 1 41-brand panel eight months earlier, noting the two waves used different panels and methods.

Gemini exclusion (disclosed): Google Gemini is not included, despite holding a measurable share of B2B AI referrals (10.6% in Goodie's Wave 2, above Perplexity's 7.3%). Three reasons: Google's AI ecosystem is already represented by AI Overviews, and including two Google surfaces would double-weight one provider in a four-platform design; the SparkToro precedent covers AI Overviews as the Google arm; and solo-researcher capacity caps the design at four platforms at the pre-registered run count (Section 6). A Gemini extension is noted as a future research direction (Section 16).

Platform-specific note — Google coupling: Google AI Overviews is a feature of Google Search, not an independent platform. Its retrieval draws on Google's search index and ranking systems, so the comparison between AI Overview appearances and Google organic rankings (Proposition 1) is partially coupled by construction. ChatGPT, Claude, and Perplexity are independent of Google's ranking systems and provide the uncoupled arm of that comparison. Platform-level results are reported separately (Section 11) partly for this reason.

Platform-specific note — retrieval mechanism: Google AI Overviews and Perplexity AI use retrieval-augmented generation (RAG) — they retrieve and cite current web content at query time. ChatGPT and Claude also use RAG when web search is enabled. For this study, web search was enabled on both ChatGPT and Claude for all queries. This decision was made to ensure all four platforms operate as real-time retrieval systems, making cross-platform comparison valid and ensuring all three research propositions are testable against current digital signals.

Both parametric (web search disabled) and retrieval-augmented (web search enabled) modes were considered for ChatGPT and Claude. Parametric mode tests what AI systems learned during training — reflecting historical web presence months or years old. This would make content recency and GBP signals untestable and would produce results incomparable to Perplexity and Google AI Overviews. Web search enabled was therefore selected. A separate parametric study examining training-data-based citation behaviour is recommended as a direction for future research.

Platform-specific note — citation mechanism: Perplexity AI cites specific URLs as footnoted sources. ChatGPT, Claude, and Google AI Overviews mention firms by name without necessarily linking to specific pages. This means schema correlation analysis is most reliably conducted against Perplexity citations where the specific cited URL can be examined. Schema analysis across ChatGPT, Claude, and Google AI Overviews is directional only.

---

5. Query Design

Query type selected: Transactional/recommendation — Type 1.

Three query types were considered:

- Type 1 — Transactional/recommendation: "Who are the best [practice area] lawyers in [location] Queensland?"
- Type 2 — Research/informational: "How do I find a good [practice area] lawyer in [location]?"
- Type 3 — Problem-specific: "I need a lawyer for [specific legal problem] in [location]."

Type 1 was selected for three reasons:

First, internal consistency. Mixing query types in an exploratory study with a small sample produces a blended dataset that represents none of the three types accurately. The COSEAL guidelines (Eimer et al., 2025) recommend focused, consistent methodology for interpretable directional findings.

Second, methodological precedent. Fishkin and O'Donnell (SparkToro, January 2026) used recommendation-intent prompts exclusively because these are the queries that determine commercial consideration set membership — whether a

firm is presented to prospective clients as a service provider to engage. FJD's research question is about this commercial outcome.

Third, commercial relevance. For an exploratory study examining which Queensland law firms are in the AI consideration set for prospective clients, recommendation-intent queries are the direct measure of the commercial outcome relevant to law firm principals. The generative engine optimisation literature (Aggarwal et al., *GEO: Generative Engine Optimization*, KDD 2024) establishes visibility in generated answers as the relevant commercial outcome variable for this class of study; the specific choice of recommendation intent rests on the consideration-set logic above rather than on any query-type finding attributed to that paper.

Limitation — practice area variation: Type 1 transactional queries may not accurately reflect search behaviour in personal injury and criminal defence practice areas. Clients in these areas tend to search with urgency and proximity intent rather than recommendation intent — Type 3 problem-specific queries are more representative of their actual search behaviour. Findings for personal injury should be interpreted with this additional caveat. Future research should test Type 3 queries for these practice areas.

Query template (finalised): "Who are the best [practice area] lawyers [location phrase], Queensland?" where the location phrase is a fixed per-location token — "in Brisbane", "on the Gold Coast", or "on the Sunshine Coast" — reflecting natural Australian usage. The state anchor "Queensland" is retained on every query because the protocol runs incognito with location services disabled and with no personalised account history (per-platform account conditions, Section 7), so the query text must carry geographic disambiguation itself.

Term selection (outcome). Both "lawyers" and "solicitors" were considered, alongside practice-area descriptor variants, against Google Australia keyword-volume data (Mangools KWFinder, July 2026). The noun "lawyers" was selected for all three practice areas: the "solicitors" variants returned zero measurable volume in one or more target locations and so could not be applied consistently across the nine-cell design. Descriptors were selected for construct fidelity to the practice area as defined in Section 6 (for example, "family lawyers" rather than the higher-volume but narrower "divorce lawyers"). The full four-constraint selection rule, the evidence tables, and the disclosure that keyword volumes were viewed during selection are recorded in the companion document *fjd-research-step1-query-lock*.

Note on "best". The word "best" is used as a recommendation-elicitation device — it prompts the platform to return a consideration set of firms it associates with the query intent. It is not a quality claim endorsed by the researcher, and no finding treats an AI's use of "best" as an assessment of firm quality. The study intentionally measures AI recommendation behaviour, not objective legal quality; different platforms may operationalise "best" differently (reputation, reviews, authority signals, citation frequency), and that variation is part of the behaviour under observation rather than a confound to be removed.

Note on the keyword data. Mangools KWFinder measures Google *search* volume, not how users phrase questions to AI assistants. The keyword data legitimately informs the noun and descriptor inside the fixed question; it does not validate the question structure, and a low or absent Google volume for a phrasing does not establish that the corresponding AI prompt is implausible.

The exact wording of all nine query combinations is fixed in the pre-registered prompt set (Section 3, Appendix A) before data collection begins.

---

6. Scope

Practice areas (3): Family law, personal injury, commercial litigation.

These three were purposively selected to span the spectrum of commercial relevance and search behaviour across Queensland legal services — family law (research-phase, high volume), personal injury (urgency-driven, compliance-restricted), and commercial litigation (B2B research-phase, lower volume). Purposive selection is the appropriate sampling logic for an exploratory observational design; the three areas are chosen for contrast and observability of digital signal variation, not as a statistically representative sample of Queensland legal practice.

Locations (3): Brisbane, Gold Coast, Sunshine Coast.

These three were purposively selected to span the geographic spread of the SEQ legal market — Brisbane (largest, most competitive), Gold Coast (second-largest LGA in Australia by population), and Sunshine Coast (a rapidly growing regional market, reflecting the region's sustained population growth). Sunshine Coast provides regional contrast against the two metropolitan markets.

Scope change from Version 1: Ipswich was included in Version 1 as a fourth location. It was removed in Version 2 to fund the increased run count recommended by the external methodology reviewer, within the data collection capacity of a solo researcher. The trade-off was assessed as follows: run count addresses the primary validity concern in this study

(response variability), while a fourth location adds geographic breadth to a dimension already partially covered by the Sunshine Coast's regional character. The reduced urban/regional contrast is disclosed as a limitation (Section 11).

Total queries: 3 practice areas × 3 locations \= 9 unique query combinations.

Runs per query: 5 runs per query per platform per wave, across two waves separated by a minimum of three weeks — 10 total runs per query per platform. Within each wave, runs of the same query on the same platform are conducted on separate days, with a minimum 24-hour gap between runs.

Total data collection: 9 queries × 4 platforms × 10 runs \= 360 recorded runs, each with a contemporaneous Google organic top-10 check.

Rationale — run count: The external methodology reviewer advised that three repetitions per query is not enough to produce reliable results, given the high variability of AI responses, and recommended a minimum of five to ten runs per query per platform, ideally repeated in two or three waves a couple of weeks apart, to enable reporting of confidence intervals rather than single percentages. Ten runs per query per platform meets that recommendation. It remains below the 60–100 runs used by Fishkin and O'Donnell (SparkToro, January 2026) — findings remain exploratory, and this is disclosed as a limitation.

Rationale — wave structure: The two-wave design captures temporal variability that single-window collection cannot. Repeated-measurement studies document meaningful drift in AI citation and recommendation behaviour over short windows — Schulte et al. (arXiv 2604.07585) observed low day-to-day source overlap (Jaccard 0.34–0.42) across a 45–46 day window, and Sielinski (arXiv 2603.08924) documented rank drift across a nine-day window. A two-wave design separated by a minimum of three weeks allows within-study observation of whether the same firms appear across both waves and prevents a short-lived platform state from dominating the dataset. Wave-level results are compared in the analysis (Section 10).

---

7. Data Collection Protocol

Session protocol: Each query is run in a new session with no prior context — a new conversation for ChatGPT and Claude, a new search for Perplexity, a new incognito window for Google AI Overviews. This prevents session context contamination where earlier queries influence later responses within the same session.

Session isolation (per run): Each run is conducted in its own dedicated incognito window, opened fresh for the run and closed after the run is logged. The AI platform query and the contemporaneous Google organic check for that run are conducted in separate tabs of the same run window. No windows or tabs from a previous run remain open when a new run begins. Claude runs: the dedicated study account is logged into within the run's incognito window (via emailed sign-in link); closing the window ends the session, and no sign-out step is required.

AI Overview capture: For Google AI Overviews runs, the AI Overview is expanded ("Show more") before recording, as a standard step in every run; firms are recorded from the fully expanded Overview. For AIO runs, the same results page serves both the AI response record and the contemporaneous organic top-10 record — no separate Google search is conducted.

Browser and location controls: All queries run in Chrome incognito mode. Location services disabled. VPN not used — queries run from the researcher's Sunshine Coast Queensland IP address to ensure geographic consistency. The same device is used for all queries. The residual IP-location signal this creates is disclosed as a limitation (Section 12).

Account conditions (fixed in advance, per platform): Account state is a pre-registered condition, not a per-run variable. Claude requires an account, so Claude runs are conducted logged in to a dedicated, clean study account with no prior conversation history, default settings, and memory/personalisation features disabled where available. ChatGPT runs are conducted logged out where web search is available in the logged-out state; if a platform change during the study period makes logged-out search unavailable, runs switch to a dedicated clean study account under the same conditions as Claude, and the change is recorded as a protocol deviation. Perplexity runs are conducted logged out. Google AI Overviews runs use incognito Google Search with no account signed in. The Appendix B account-state field confirms compliance with these fixed conditions; it does not select them.

AI Overview trigger rule: Google AI Overviews do not appear for every query on every occasion. If no AI Overview is returned for a run, the run is recorded as "AIO not triggered," is excluded from the mention-frequency denominator for that combination, and is replaced by an additional run on a subsequent day within the wave window where feasible. The AI Overview trigger rate per query combination is itself recorded and reported as a finding.

Per-run documentation: Every run is logged at the time it is conducted using the logging template at Appendix B. The log records, for each run: date and time (AEST), wave number, run number within wave, platform, model or platform version string as displayed at time of run, account state (logged in / logged out), web search mode (enabled / disabled /

native), browser mode, query text as entered, and the firm names recorded from the response. If a model update occurs during the data collection period, pre-update and post-update runs are distinguished in the analysis.

Timing: Wave 1 is completed within a two-week collection window. Wave 2 begins a minimum of three weeks after Wave 1 ends and is completed within a two-week collection window. The start and end dates of each wave are recorded and disclosed in the published findings.

Recording: Firm names appearing in AI responses are recorded. AI response text is not reproduced verbatim — firm names only are recorded. Each firm is recorded once per run, in order of first appearance, and firm-name strings are recorded with consistent spelling across runs (the first-recorded form of a firm's name is reused verbatim in subsequent runs). Individual lawyer names are excluded from the dataset; the log records a count only, in the fixed format "named individuals — excluded (n)", where individuals named without a surname count toward n, and "none" where no individuals are named. This protects individual privacy under the Australian Privacy Act 1988. **Sole-practitioner naming rule:** many firms trade under a personal name (e.g. "Jane Smith Family Law"). Where a name presented by the AI is the registered trading or business name of a firm, it is recorded as a firm; where the AI names an individual practitioner as a person (rather than as a firm's trading name), it is recorded within the excluded-individuals count. Where the distinction is unclear from the response, the name is checked against the firm's GBP listing and website at signal-recording time and classified accordingly. **Entity-type framing rule:** whether a name denotes a firm or an individual is determined in the first instance from the response's own framing (e.g. an entity described as containing practitioners is a firm); genuinely ambiguous cases are recorded as firms and flagged in the Anomalies field for post-collection review.

Firms-recorded inclusion and exclusion rules: Included: firms presented in body text, list entries, or firm profile cards as providers to contact; firms identified as the employer or practice of an excluded named individual (the firm is recorded; the individual is not). Excluded: directories, ranking guides, and comparison platforms named as sources of the rankings (e.g. Doyle's Guide, Best Lawyers, LawConnect) — these are recorded only where they appear as organic search results in the comparison field; firms appearing solely in citation panels, source cards, or inline citation chips; sponsored or advertising placements displayed with or alongside the response; firms mentioned only as ownership or backing context and not presented as the provider to contact (noted in Anomalies on first occurrence).

What counts as a mention: A mention is recorded when a firm's name or website URL appears in the AI response in a context where the AI is presenting the firm as a legal service provider to contact — not where the firm's content is cited as an information source. Recommendation appearances only are recorded.

Intra-rater reliability control: Because all coding is performed by a single researcher, the recommendation-versus-source-citation judgement (above) is a potential reliability weakness. Two controls are applied. First, every AI response is saved (screenshot or full-text capture) at time of collection, so every coding decision is auditable against the original response. Second, after a minimum two-week gap following the completion of data collection, a random 10% sample of saved responses is re-coded blind to the original coding, and the intra-rater agreement rate is computed and reported in the findings. The same re-coding check is applied to the discretionary signal judgements: a random 10% sample of GBP completeness scores (Signal C) and schema classifications (Signal B) is re-scored blind against the saved screenshots and test outputs from the original observation, and agreement rates are reported alongside the mention-coding rate. This does not substitute for independent inter-rater coding, which is not feasible for a solo researcher, but it provides a disclosed, quantified check on coding consistency.

Google organic ranking check: Immediately after recording the AI response for each run, the same query is entered into Google Search in a new tab of the same run window (incognito, same session). The organic blue-link results are recorded in ranked order, up to ten, by source name as displayed (the site name shown above the URL); where no source name is displayed, the domain is recorded. Each domain is recorded once, at its first position. Excluded: sponsored results, map pack / Places, featured snippets, People Also Ask boxes, Discussions and forums modules, and all other non-organic modules. AI Overviews appearing on comparison searches are ignored and their appearance noted in Anomalies (AI Overview content is recorded only on designated AIO runs). Where fewer than ten organic results appear on page one, the actual count is recorded in Anomalies and page two is not consulted. This check is conducted in the same session as the AI query to ensure contemporaneous comparison.

Evidence capture: A full-page capture (PDF or screenshot) of each AI response and each comparison search results page is saved at the time of the run, filed by Run ID, and retained with the run log for the study's data-retention period. The run log remains the primary record; captures are supporting evidence and the audit basis for the intra-rater re-coding check above.

---

8. Study Population and Denominator

Population definition: Any law firm with a current, active Google Business Profile listing in the nominated practice area and location at the time of data collection, as identified by a Google Maps search using the practice-area term matching the locked query for that combination ("family lawyer", "personal injury lawyer", or "commercial litigation lawyer") plus

the location, conducted immediately before Wave 1 data collection for that combination begins. The same population list is used for both waves. Firms first appearing in AI responses during Wave 2 that were not in the Wave 1 population list are recorded and disclosed but analysed separately.

Rationale: The QLS Register of Solicitors was considered as the primary population source. It was not adopted for three reasons: it was under scheduled maintenance at the time of study design; it does not distinguish active local practices from relocated solicitors; and it does not reflect current digital presence, which is the relevant criterion for this study. GBP presence is the most current, accessible, and observable proxy for active local legal practice available without regulatory access.

Denominator recording: The Google Maps firm count for each practice area and location combination is recorded before queries are run. This count becomes the denominator for all percentage calculations. Multi-office firms are counted once per location where they have a distinct GBP listing.

Population capture procedure (reproducible): Google Maps results are ranked, viewport-dependent, and not a fixed list, so the capture is proceduralised as follows. (1) Desktop Google Maps in Chrome incognito, location services disabled, same device as the study runs. (2) Exact search string: the practice-area term matching the locked query plus the location name (e.g. "family lawyer Sunshine Coast"). (3) The map viewport is set by selecting the location's own Maps entity (which frames the LGA area) before searching, and is not manually panned or zoomed during capture. (4) The results list is scrolled to exhaustion — until Google returns no further results — and every listing is recorded. (5) Inclusion rule: a firm is included if its GBP address falls within the relevant local government area — City of Brisbane, City of Gold Coast, or, for the Sunshine Coast, either Sunshine Coast Regional Council or Noosa Shire Council. Noosa Shire is included in the Sunshine Coast definition because common usage of "Sunshine Coast" — and therefore the plausible referent of the query — spans both LGAs; this boundary decision is fixed before data collection and disclosed. Listings surfaced from outside the defined area are recorded but excluded from the denominator. (6) Date, time, search string, and a full-scroll screen capture are retained for each capture, so the population list is auditable. The capture for each combination is completed immediately before Wave 1 collection for that combination begins, per the population definition above.

Reproducibility of the denominator (disclosed): Google Maps results are affected by indexing updates, business suspensions, duplicate handling, and ranking changes over time, so the denominator is reproducible procedurally rather than identically — another researcher applying this procedure at a different time may obtain a slightly different count. The retained captures make this study's specific population list fully auditable, which is the relevant standard for an observational study of a moving platform.

Population threshold vs signal variable: GBP presence is the population threshold — firms without a GBP listing are not in the study population. GBP completeness score (the 5-point rubric defined in Section 9) is a signal variable. These are distinct and not circular — every firm in the population has a GBP listing, but firms vary significantly in how complete that listing is.

---

9. Observable Signal Variables

For each firm appearing in an AI recommendation response, the following signals are recorded during the wave in which the firm first appears. All signals are observable from publicly available sources without accessing the firm's backend systems. Signals for firms appearing in both waves are recorded once per wave, so that signal change between waves can be observed.

Signal A — Google organic ranking (Research Proposition 1): Does the firm appear in Google's top-10 organic results for the same query, checked immediately after the AI response? Recorded as: Appears in Google's top-10 organic results (Yes/No). Position within the top 10 is also recorded (1-10).

Signal B — Structured data presence (Research Proposition 2): Does the firm's website expose detectable structured data for FAQPage, LocalBusiness/LegalService, or Person schema types? Primary detection: the Schema.org validator (validator.schema.org), which tests markup validity independently of rich-result eligibility. Supplementary check: Google's Rich Results Test, recorded separately as a Google rich-result-eligibility indicator — noting that the Rich Results Test detects only schema types Google considers eligible for rich results (FAQPage rich-result eligibility, for example, has been restricted since 2023), so it cannot serve as a general validity test. Recorded as: Detected valid / Detected with errors / Not detected (Schema.org validator), plus Rich-result eligible (Y/N). Analysis of structured-data association is conducted separately for Google AI Overviews (where structured data is a confirmed Google ecosystem signal) and ChatGPT/Perplexity/Claude (where the relationship is less established).

Signal C — GBP completeness score (Research Proposition 3): Scored 0-5 on five observable elements:

1. Business description present (Y/N)
2. Primary category set (Y/N)

3. Services section populated with at least one service (Y/N)
4. Practice-uploaded photos present — at least one photo showing people, premises, or work context, excluding logo-only images and Google-generated images (Y/N)
5. GBP post visible within 90 days at time of observation (Y/N)

Note on virtual firms: Law firms without physical premises are scored 0 on element 4 with a notation "virtual practice — premises photos not applicable." This score is excluded from photo-related analysis to avoid penalising a practice model rather than a digital signal gap.

Signal D — Website content recency (Research Proposition 3): Date of most recent blog, news, or journal post on the firm's own website at time of observation. Recorded as: actual date. Not categorised as a binary threshold — the full dataset is analysed for patterns. Firms with no published content are recorded as "no content published." Firms with content that has not been updated are recorded with the date of the most recent post.

Signal E — Google review count (Exploratory observation — restricted practice areas): Total Google review count as shown on GBP at time of observation. Recorded for commercial litigation only within this study's practice areas. Not recorded for personal injury (PIPA testimonial restrictions make review count an unreliable signal for compliance-driven firms) or family law (review volumes are likely depressed relative to other practice areas by client confidentiality sensitivities — many clients will not publicly review a family lawyer — making cross-firm comparison unreliable as a measurement, independent of any individual firm's review presence).

Signal F — Legal directory presence (Exploratory observation): Presence recorded across three directories, each verified against the qualifying criterion on 28 July 2026: QLS Find a Solicitor (opt-in referral service, hosted on the QLS You & The Law site), LawConnect (public Australian lawyer/firm directory, opt-in/claim-based listing), and FindLaw Australia (open national legal directory, 10,000+ listings). Scored 0-3. Directories qualify only if they host publicly accessible, crawlable listings of third-party firms that any Queensland firm can in principle appear on.

Recording rules for QLS Find a Solicitor: the service lists individual solicitors rather than firms, so a firm is recorded as present if any solicitor at the firm appears for the relevant location; and because members may nominate a maximum of three practice areas, presence is recorded at location level, not practice-area level — a firm genuinely practising an area may not have nominated it, so practice-area matching within this directory would under-count.

Exclusions (recorded): LegalVision and Lawpath, included in an earlier draft, were removed because LegalVision is itself a legal services provider and Lawpath is a legal services marketplace, so most Queensland firms cannot appear on them regardless of digital signal quality, which would have imposed a structural ceiling on the score. Doyle's Guide and Best Lawyers Australia were considered and excluded on the same shared ground: both are selection-based (peer nomination / annual peer review), so firms cannot list themselves, failing the any-firm-can-appear criterion.

---

10. Research Propositions

These are observational propositions for an exploratory study — not confirmatory hypotheses. They do not claim causal relationships. Any association observed between a signal and AI recommendation appearance is a correlation; it may reflect the underlying characteristics of established firms rather than the effect of the signal itself.

Proposition 1: Queensland law firms appearing consistently in AI recommendations across four platforms also tend to appear in Google's top 10 organic results for the same query — examining whether AI recommendation appearance and organic ranking are correlated or independent in the Australian legal niche, and whether the pattern of increasingly independent AI-citation and organic-ranking behaviour reported in the broader literature holds in a tight geographic niche. The reference point: Ahrefs' March 2026 analysis found the share of Google AI Overview citations drawn from top-10 organic pages fell from 76% (July 2025) to 38%, with 62% of citations now coming from pages outside the top 10 — attributed in part to query fan-out and in part to improved citation parsing, so the two readings are not directly comparable. Relatedly, Ahrefs' December 2025 study of AI Mode versus AI Overviews found only 13.7% citation overlap between the two Google AI surfaces themselves, underscoring how weakly coupled AI citation is to any single result set. This comparison is exploratory; the independence pattern is treated as a reference point to test against, not an established fact assumed by the study. **Coupling caveat:** because Google AI Overviews is itself a feature of Google Search (Section 4), its arm of this comparison is partially coupled by construction; the ChatGPT, Claude, and Perplexity arms provide the independent test.

Proposition 2: Queensland law firms appearing consistently in AI recommendations tend to exhibit detectable structured data (FAQPage, LocalBusiness/LegalService, Person) via the Schema.org validator — examining whether structured data presence is an observable shared characteristic of firms with consistent AI recommendation appearances in the Queensland legal market.

Proposition 3: Queensland law firms appearing consistently in AI recommendations tend to have more complete GBP

profiles and more recently published website content than firms not appearing — examining whether active digital maintenance is an observable shared characteristic of firms with consistent AI recommendation appearances.

Exploratory observations: Review count (commercial litigation only), GBP post recency (Google AI Overviews only), and multi-directory presence are recorded but not framed as propositions. Patterns observed are reported descriptively.

---

11. Analysis Approach

This analysis plan is fixed by pre-registration (Section 3) before data collection begins.

Mention frequency: For each query combination, a firm is recorded as appearing in n/10 runs per platform (0–10), with wave-level sub-counts (n/5 per wave). Firms appearing in 7 or more of 10 runs on a platform are described as "consistent appearances" for that combination on that platform. This threshold is set before data collection, is arbitrary at the exploratory stage, and is disclosed as such.

Threshold behaviour in sparse cells: In query combinations where very few firms appear at all — anticipated for commercial litigation on the Gold Coast and Sunshine Coast (Section 12) — the 7-of-10 threshold may be unreachable or, conversely, met by the only firm that appears. In such cells the threshold is uninformative, and this is reported as an observation in its own right (that AI platforms do not reliably surface a stable set of firms for that combination) rather than smoothed over or excluded.

Confidence intervals: Mention frequencies are reported as proportions with approximate 95% Wilson score confidence intervals under an independence assumption (see below). The Wilson interval is used because it performs reliably for small samples and proportions near 0 or 1, where the standard Wald interval is known to fail (Brown, Cai and DasGupta, Statistical Science, 2001). Example reporting format: "Firm-level mention frequency 7/10 (70%; 95% CI 40%–89%)." Intervals are reported in full in the findings appendix; body text uses plain-language proportions with the interval stated at first mention of each headline figure.

Interpretation of intervals: At 10 runs, confidence intervals will be wide. This is expected and is itself a finding — it communicates honestly how variable AI recommendation behaviour is at this scale. Wide intervals are not smoothed, narrowed, or omitted from reporting.

Independence assumption (disclosed): Wilson intervals assume runs are independent Bernoulli trials. Runs of the same query on the same platform, even on separate days, may be correlated through persistent platform state, so nominal 95% coverage is approximate. The 24-hour minimum gap and the two-wave separation reduce, but do not eliminate, this dependence. This is consistent with Sielinski's (arXiv 2603.08924) treatment of visibility metrics as sample estimators, and intervals are interpreted as disclosed approximations rather than exact coverage guarantees.

Wave comparison: Wave 1 and Wave 2 mention counts are reported side by side for each consistent-appearing firm. Firms appearing consistently in one wave but not the other are reported as such. No statistical test of wave difference is conducted — the comparison is descriptive.

Signal comparison: The observable signals recorded for consistent-appearing firms are compared descriptively against the observable signals recorded for non-appearing firms from the same practice area and location combination. Non-appearing firms are drawn from the GBP-based population for each combination — a random sample of five non-appearing firms per combination is selected for signal recording to provide a directional comparison group. All comparisons are reported as correlations. No comparison in this study supports a causal claim, and the published findings state this in the body text, not only in the limitations section.

Aggregation: Findings are reported at practice area and location level. No individual firm is identified in the published findings. Aggregated statements are limited to forms the collected data can support: mention-frequency proportions over the run denominator (e.g. "X% of runs for Brisbane family law returned at least one firm from the top 10 organic results, 95% CI: Y%–Z%"), and signal comparisons between appearing firms and the sampled non-appearing comparison group (e.g. "N of M consistently appearing Brisbane family law firms had a GBP completeness score of 4 or 5, versus N of 5 sampled non-appearing firms"). Population-wide signal percentages are not reported, because signals are recorded only for appearing firms and the comparison sample (Section 9), not for the full population.

Platform-specific analysis: Where platform differences are observed — particularly between Google AI Overviews and ChatGPT/Claude/Perplexity — these are reported separately. Cross-platform aggregation is secondary to platform-specific analysis.

---

12. Limitations

The following limitations apply to this study and must be considered when interpreting findings:

1. **Run count:** Ten runs per query per platform meets the external reviewer's recommended range and permits confidence interval reporting, but remains below the 60–100 runs per prompt used by Fishkin and O'Donnell (SparkToro, January 2026) and below the query volumes Sielinski (arXiv 2603.08924) indicates are needed for narrow confidence intervals on some platforms. This study does not claim rankings and does not require ranking stability; confidence intervals will be wide, and findings are directional observations with quantified uncertainty, not confirmatory frequency measurements.

2. **Two-wave snapshot:** This study captures AI recommendation behaviour across two windows separated by a minimum of three weeks. Repeated-measurement studies (Schulte et al., arXiv 2604.07585; Sielinski, arXiv 2603.08924) show that AI citation behaviour drifts over comparable and longer windows, so the two-wave design observes short-term stability only, and findings may not represent the landscape at time of publication or in subsequent months.

3. **RAG vs parametric distinction:** Even with web search enabled, ChatGPT and Claude draw on a combination of training data and real-time retrieval. The relative weighting of these two mechanisms is not publicly disclosed by the platform providers. Content recency and GBP signals may have different influence on parametric-heavy responses than on retrieval-heavy responses.

4. **Comparison group:** Non-appearing firms are compared directionally using a random sample of five firms per combination. This is insufficient for statistical correlation analysis. Findings describe shared characteristics of appearing firms — they do not establish which signals distinguish appearing from non-appearing firms with statistical confidence.

5. **Correlation only — established-firm confound:** Firms appearing in AI recommendations are likely to be established practices scoring well on all observable signals simultaneously. No observed association between an individual signal and AI recommendation appearance can be attributed to that signal. This confound cannot be resolved within an observational design; a controlled intervention study is required (Section 16).

6. **Observer effects:** All data was collected by a single researcher with a commercial interest in the findings. Bias mitigation measures (pre-registration, new session per query, incognito mode, location services disabled, consistent device, per-run logging) were applied but cannot fully eliminate observer effects.

7. **Black box limitation:** The selection mechanisms behind AI recommendations and citations are not publicly disclosed by any platform. This study can only observe correlations between digital signals and recommendation appearances. It cannot determine which signals cause selection or explain the mechanism.

8. **Geographic bias:** ChatGPT, Perplexity, and Claude are US-developed platforms trained primarily on English-language content with a US geographic weighting. Recommendation behaviour for Queensland legal queries may be influenced by training data that over-represents US legal content and under-represents Australian legal content. Firms with international content signals may receive disproportionately high recommendation frequency independent of their Queensland-specific digital signals.

9. **Practice area query type variation:** Type 1 transactional queries may not reflect the urgency-driven, proximity-sensitive search behaviour typical of personal injury clients. Findings for this practice area should be interpreted with additional caution.

10. **GBP population proxy:** GBP presence is used as a proxy for active local legal practice. Firms without GBP listings are excluded from the study population. This may exclude practices operating without a GBP listing — which is itself a signal characteristic but one that cannot be captured within this population definition.

11. **Training data timing:** Observable signals (schema, content, reviews) reflect the state of a firm's digital presence at time of data collection. AI training data has a cutoff that may predate current signals. Observed correlations between current signals and AI recommendation frequency may reflect historical rather than current signal states.

12. **Reduced geographic scope:** Ipswich was removed from the study scope in Version 2 to fund the increased run count. The study's urban/regional contrast rests on the Sunshine Coast alone. Findings do not describe smaller regional legal markets, and the urban/regional comparison is weaker than in the Version 1 design.

13. **IP-location bias:** All runs originate from a Sunshine Coast IP address with location services disabled. Platforms may still infer approximate location from IP, so Brisbane and Gold Coast queries carry a possible systematic pull toward Sunshine Coast results — one of the three studied markets. The explicit location anchor in every query text mitigates but does not eliminate this. The bias is constant across all runs (same device, same connection), so it is a consistent, disclosed condition rather than a variable one.

14. **Commercial-litigation coast cells — thin market:** The two litigation-specific search terms for commercial litigation return no measurable Google search volume on the Gold Coast or the Sunshine Coast, while broader terms ("commercial lawyers", "business lawyers") denote the whole commercial-law field rather than disputes. The study

retains "commercial litigation lawyers" across all three locations to hold the practice area constant (the term-selection reasoning is documented in the companion query-lock document), accepting that the two coast cells (queries 8 and 9) may return thinner firm populations than the family-law cells and correspondingly wider or, in the sparsest case, uncomputable Wilson confidence intervals. A firm population does exist in these markets — the disqualified search terms still return populated local result sets — so the cells are expected to yield usable but sparse data. Sparsity in these cells is treated as a reportable observation, not as a reason to exclude them, to avoid biasing the sample toward populated combinations.

---

13. Ethical Considerations

Individual privacy: Individual lawyer names appearing in AI responses are excluded from the dataset. Only firm-level data is recorded and reported.

Firm-level privacy: Law firm entities are not protected under the Australian Privacy Act 1988. Recording that a named firm appeared or did not appear in AI recommendations for a given query does not constitute a privacy breach for the firm entity.

Data retention: All raw data collection records, including per-run logs (Appendix B), platform response records, and signal observation records, are retained for a minimum of three years from date of publication. This enables verification of findings if challenged.

Defamation: Findings are aggregated at practice area and location level. Individual firms are not identified as "not appearing" or as having weak digital signals. The research is descriptive, not evaluative. It does not constitute a quality assessment of any Queensland law firm.

Copyright: AI response text is not reproduced verbatim in the published findings. Firm names only are recorded from AI responses. All secondary sources cited in the methodology and published findings have been directly accessed and verified before citation.

ACL and ACCC AI-washing: The findings of this study will be reviewed against ACCC guidance on AI-related claims before any use in FJD's commercial marketing materials. The ACCC's 2026 AI-washing enforcement priorities require that claims about AI system behaviour be accurate, not hide important limitations, and use clear language. FJD commits to applying these standards to any subsequent use of study findings.

---

14. External Validation

Methodology review — completed. Version 1 of this methodology was reviewed by Dr Karen Sutherland, Senior Lecturer in Public Relations at the University of the Sunshine Coast and author of Artificial Intelligence for Strategic Communication (Palgrave Macmillan, 2025), in July 2026. Her review addressed research design rigour and transparency from an AI research ethics standpoint. The revisions made in response are documented in the Revision Record.

Dr Sutherland's review is of the methodology only — not an endorsement of the findings, which had not been collected at the time of her review. Any attribution or quotation from Dr Sutherland in the published findings is subject to her review and written approval of the final report before publication. Her preferred form of attribution — personal academic capacity or institutional identification — will be confirmed with her as part of that approval, and the published wording will follow her confirmation; her affiliation is stated in this document for identification, not as institutional endorsement.

Findings review — outstanding. A second reviewer will be identified to review the findings before publication. This reviewer will provide a face validity check on the analysis and a challenge function on the interpretation of results.

---

15. Publication Plan

1. Pre-registration (prompt set, analysis plan, logging template) registered on OSF before Wave 1 begins, with a mirror PDF hosted on felicityjane.com.au. Neither is publicised before the findings are published.
2. Findings published on the Felicity Jane Digital journal (felicityjane.com.au/blog) on the date of first publication — with full methodology, data collection dates, wave dates, confidence intervals, and limitations section.
3. Trade media pitch (QLS Proctor, Lawyers Weekly) sent same day or next day with the live published URL. No embargoed findings distributed before publication.
4. LinkedIn post series follows journal publication — summarising key directional findings with explicit reference to the full methodology.

---

16. Future Research Directions

This exploratory study generates the following directions for future confirmatory research:

1. A higher-volume study (minimum 30 runs per query per platform) to establish statistically valid mention frequency rates for Queensland law firms across AI platforms.
2. A longitudinal study tracking the same query combinations across 12 months to capture AI recommendation concentration trends in the Queensland legal market.
3. A Type 3 problem-specific query study for personal injury and criminal defence practice areas — reflecting the urgency-driven search behaviour of clients in those practice areas.
4. A parametric knowledge study — running ChatGPT and Claude without web search enabled — to distinguish training-data-based recommendation behaviour from retrieval-based recommendation behaviour for Queensland law firms.
5. A controlled intervention study — measuring AI recommendation frequency before and after targeted schema implementation, GBP completion, or content publishing — to establish whether the observed correlations hold under controlled conditions.
6. A regional-market extension covering Ipswich, Toowoomba, and Moreton Bay, restoring the urban/regional contrast reduced in this study's Version 2 scope.
7. A platform extension adding Google Gemini, to test whether Google's standalone assistant behaves like AI Overviews (shared ecosystem) or like the independent platforms.

Appendix A — Pre-Registered Query List (FINAL)

The nine query combinations below are locked in exact wording. No changes are made after registration. The term-selection rule and evidence are recorded in the companion document *fjd-research-step1-query-lock*.

#	Practice area	Location	Exact query string
1	Family law	Brisbane	Who are the best family lawyers in Brisbane, Queensland?
2	Family law	Gold Coast	Who are the best family lawyers on the Gold Coast, Queensland?
3	Family law	Sunshine Coast	Who are the best family lawyers on the Sunshine Coast, Queensland?
4	Personal injury	Brisbane	Who are the best personal injury lawyers in Brisbane, Queensland?
5	Personal injury	Gold Coast	Who are the best personal injury lawyers on the Gold Coast, Queensland?
6	Personal injury	Sunshine Coast	Who are the best personal injury lawyers on the Sunshine Coast, Queensland?
7	Commercial litigation	Brisbane	Who are the best commercial litigation lawyers in Brisbane, Queensland?
8	Commercial litigation	Gold Coast	Who are the best commercial litigation lawyers on the Gold Coast, Queensland?
9	Commercial litigation	Sunshine Coast	Who are the best commercial litigation lawyers on the Sunshine Coast, Queensland?

Appendix B — Per-Run Logging Template

Every run is logged at the time it is conducted. One record per run. Fields:

Field	Description
Run ID	Sequential identifier: [Query #]-[Platform]-[Wave]-[Run], e.g. 04-CGPT-W1-R3
Date	Date of run (AEST)
Time	Time run commenced (AEST)
Wave	1 or 2
Run number within wave	1–5
Platform	ChatGPT / Perplexity / Google AI Overviews / Claude
Model / platform	Version string as displayed by the platform at time of run. Fixed entries per platform under study conditions: ChatGPT — "not displayed (logged out)"; Perplexity — "default (not selectable, logged out)" (model selector present

version	but locked to logged-out users); Google AI Overviews — "not displayed"; Claude — version string as displayed in the composer. If a platform's display changes during the study, the new string is recorded and the change noted in Anomalies
Account state	Confirms the fixed per-platform condition (Section 7): Claude — dedicated study account, logged in; ChatGPT — logged out (or dedicated study account if protocol deviation recorded); Perplexity — logged out; AI Overviews — no account signed in
Web search mode	Enabled / native (Perplexity, AI Overviews)
Browser mode	Chrome incognito (confirm)
Location services	Disabled (confirm)
Query text	Exact query as entered
Firms recorded	Firm names the response presents as service providers to contact, once per run each, in order of first appearance, per the Section 7 inclusion and exclusion rules (source directories, citation-only appearances, and sponsored placements excluded; firms attached to excluded individuals included)
Named individuals	Count only, never names: "named individuals — excluded (n)"; "none" where no individuals are named
Organic top 10	Organic blue-link results from the contemporaneous Google check, in ranked order by source name as displayed (fallback: domain), one entry per domain at first position, per the Section 7 capture rules; actual count noted in Anomalies where fewer than ten
AIO triggered	Google AI Overviews runs only: Yes / No. If No, run recorded as "AIO not triggered", excluded from mention denominator, replacement run scheduled (Section 7)
Duration	Elapsed time from query entry to completion of data logging for the run (pilot benchmark: approximately 4–5 minutes per run including logging)
Anomalies	Model update notices, refusals, errors, deviations from protocol; AI Overview appearances on comparison searches; sub-ten organic counts; ownership/backing-context firm mentions; entity-type ambiguity flags

---

Document version: 2.4 — pilot complete (3 August 2026); pilot-derived capture rules incorporated; registration version Prepared: August 2026 Researcher: Jane Cluff, Felicity Jane Digital Next actions: (1) Register on OSF (this document + companion query-lock document + run-log template), embargoed to publication. (2) Capture population lists immediately before Wave 1. (3) Begin Wave 1.

FJD Research Project — Step 1 Query-Lock Record

AI Visibility in South East Queensland Legal Services

Purpose: Authoritative record of the exact query wording for all nine combinations. Feeds Appendix A of the OSF pre-registration. Once a practice area is locked here, the wording does not change. **Created:** 28 July 2026 **Updated:** 28 July 2026 — aligned to methodology v2.3 (see revision note at end; no query wording changed) **Status:** ALL NINE QUERY STRINGS FINAL. Location phrasing locked 28 July 2026. One item remains before OSF (Sutherland sign-off on CL cells — validity, not wording). **Owner:** Jane Cluff (AI-assisted drafting disclosed in methodology) **Related:** methodology v2.3 Appendix A · master reference v1.1 Part 4 Action #9

Fixed template (finalised 28 July 2026)

Who are the best [practice area] lawyers [location phrase], Queensland?

Type 1 transactional/recommendation. “Best” and the question form are locked from methodology v2.3 Section 5. The bracketed practice-area term is set by the term-selection rule below. The [location phrase] is a fixed per-location token — “in Brisbane” / “on the Gold Coast” / “on the Sunshine Coast” — reflecting natural Australian usage (locations do not share a preposition). The state anchor “Queensland” is retained via comma on every query because the protocol runs incognito, with location services disabled and no personalised account history (per-platform account conditions, methodology v2.3 Section 7), so the query text must carry geographic disambiguation itself. Locations are fixed: Brisbane, Gold Coast, Sunshine Coast.

Location-phrasing rationale (disclosed): An earlier draft used a uniform “in [location] Queensland” frame. This was revised before pre-registration because “in Sunshine Coast” / “in Gold Coast” is not natural Australian phrasing, and the study’s ecological validity depends on prompts matching how a real AI-using client would phrase the question. The location phrase is held constant across the three practice areas within each location, so the preposition never varies inside a practice-area comparison — it is part of the location reference, not a manipulated variable, and introduces no confound. This is finalisation before lock, not a post-registration protocol deviation.

Term-selection rule (as applied)

For each practice area, the noun (lawyer/solicitor) and descriptor are chosen on Google Australia seed search volume in Mangools KWFinder, read across the three target locations, subject to the following constraints applied in strict priority order:

1. **Cross-location consistency (Constraint 1).** The chosen term must return usable volume in all three locations. A term returning zero volume in any location is disqualified — it cannot be applied consistently across the nine-cell design.
2. **Construct validity (Constraint 2).** The chosen term must denote the practice area as defined in the study scope — not a subset of it and not a superset broader than the defined practice area.
3. **Preference tiebreak (Constraint 3).** Only among terms already judged construct-equivalent and consistency-valid does the shorter, more common phrasing win a volume tie (terms within 10% of each other).
4. **Last-resort rule (Constraint 4) — when Constraints 1 and 2 cannot both be satisfied.** Where no candidate term is both construct-valid and consistent across all three locations, construct fidelity (Constraint 2) prevails over cross-location consistency (Constraint 1). The consistency gap is documented as a disclosed limitation, and the affected cells are flagged for thin-population / wide-confidence-interval risk in the pre-registered limitations. Rationale: a term that measures the wrong practice area invalidates the construct in every cell; a consistency gap weakens specific cells while leaving the construct intact and the gap itself observable and reportable.

Sequence disclosure: Volumes were viewed before terms were selected. This is documented honestly and disclosed in the method. Term choice does not affect which firms appear in AI responses, so there is no result to game; the disclosure protects the pre-registration’s integrity claim.

Data-source caveat (for the method): KWFinder measures Google search keyword volume, not how people phrase questions to AI assistants. The keyword data legitimately drives the noun and descriptor

choice inside the fixed question — it does not validate the question structure itself, and a zero or absent Google volume does not establish that the corresponding AI prompt is implausible or that no firm population exists.

Note on rule evolution (disclosed): An earlier draft of this rule made “shorter/more colloquial wins ties” a primary clause. It was overridden on construct grounds in both practice areas where a tie arose (family law, personal injury). The rule was therefore rewritten before commercial litigation was locked, to demote that preference to Constraint 3 and to state construct-fidelity primacy explicitly (Constraints 2 and 4). The written rule now matches the decisions actually made.

Locked queries

Family law — LOCKED 28 July 2026

Term: family lawyers [location]

Evidence (Google Australia, seed search volume):

Term	Brisbane	Gold Coast	Sunshine Coast	3-loc avg
family lawyers [loc]	3,900	690	1,200	1,930
divorce lawyers [loc]	3,900	950	1,200	2,017
family solicitors [loc]	3,900	950	0	disqualified (C1)
family law [loc]	3,900	260	280	1,480

Reasoning: Noun = “lawyers” (“family solicitors sunshine coast” returns 0 — fails Constraint 1). Descriptor = “family”: “divorce” leads on 3-location average by ~4% (a tie-band margin) but is overridden under Constraint 2 — divorce is a subset of family law (excludes de facto, property settlement, children’s matters). Brisbane returns a flat 3,900 across all four terms (Google buckets them as synonyms), so separation exists only on the Coast.

Final strings: 1. Who are the best family lawyers in Brisbane, Queensland? 2. Who are the best family lawyers on the Gold Coast, Queensland? 3. Who are the best family lawyers on the Sunshine Coast, Queensland?

Personal injury — LOCKED 28 July 2026

Term: personal injury lawyers [location]

Evidence (Google Australia, seed search volume):

Term	Brisbane	Gold Coast	Sunshine Coast	3-loc avg
personal injury lawyers [loc]	710	460	160	443
injury lawyers [loc]	710	460	160	443
compensation lawyers [loc]	370	100	80	183
personal injury solicitors [loc]	710	0	0	disqualified (C1)

Reasoning: Noun = “lawyers” (“personal injury solicitors” returns 0 in Gold Coast and Sunshine Coast — fails Constraint 1). Descriptor = “personal injury”: the pre-stated hypothesis that “compensation” (a parallel umbrella, not a subset) might win outright did not hold — it is materially lower in all three locations (370/100/80 vs 710/460/160), so there is no basis to move off the personal-injury descriptor. “injury lawyers” is volume-tied (identical 710/460/160, Google synonym-bucketing) but truncates the construct named throughout the study; retained “personal injury” under Constraint 2 over the Constraint 3 shorter-form preference.

Final strings: 4. Who are the best personal injury lawyers in Brisbane, Queensland? 5. Who are the best personal injury lawyers on the Gold Coast, Queensland? 6. Who are the best personal injury lawyers on the Sunshine Coast, Queensland?

Commercial litigation — LOCKED 28 July 2026 (via Constraint 4)

Term: commercial litigation lawyers [location]

Evidence (Google Australia, seed search volume):

Term	Brisbane	Gold Coast	Sunshine Coast	3-loc avg
commercial litigation lawyers [loc]	110	0	0	construct-valid; fails C1
commercial lawyers [loc]	260	150	70	160 (superset)
business lawyers [loc]	140	100	30	90 (superset)
commercial litigation solicitors [loc]	10	0	0	fails C1 and construct-noun

Reasoning — Constraint 4 applied. No candidate satisfies both Constraint 1 and Constraint 2. The two litigation-specific terms return zero coast volume (fail C1); the two consistent terms ("commercial lawyers", "business lawyers") name the whole commercial-law field — advisory, contracts, corporate — not disputes, and are supersets of the defined practice area (fail C2). Under Constraint 4, construct fidelity prevails: "commercial litigation lawyers" retained. This also preserves the "[practice area] lawyers" pattern across all three areas.

Why the coast zeros do not block retention: Commercial litigation is B2B, low-volume, and Brisbane-concentrated (Brisbane itself is only 110 vs family law's 3,900). The coast zeros reflect Google search demand below KWFinder's reporting floor in a thin market, not an absence of firms and not different client vocabulary. The disqualified term still returns populated coast SERPs — Gold Coast: marinolaw, nystlegal, simontaylorsolicitor, roselitigation, tdllaw; Sunshine Coast: axialawyers, clutchlegal, awsunshinecoastlawyers, argonlaw, stonegatelegal — confirming a real firm population exists for the AI to name. Per the data-source caveat, an AI prompt at zero Google volume remains plausible.

Disclosed limitation (must appear in pre-registered limitations): The two coast commercial-litigation cells (query 8 × all platforms, query 9 × all platforms) may return thinner firm counts than the family-law cells, producing wider or, in the sparsest case, uncomputable Wilson confidence intervals. A sparse or empty result is itself a reportable observation (AI may not reliably surface coast-level commercial-litigation specialists) and is retained rather than removed, to avoid biasing the sample toward populated cells. Flagged for Dr Sutherland's review before OSF lock as the one Step 1 cell where the two integrity principles (consistency vs construct fidelity) directly collide.

Final strings: 7. Who are the best commercial litigation lawyers in Brisbane, Queensland? 8. Who are the best commercial litigation lawyers on the Gold Coast, Queensland? 9. Who are the best commercial litigation lawyers on the Sunshine Coast, Queensland?

Remaining before OSF pre-registration

1. **Location phrasing — RESOLVED 28 July 2026.** Natural per-location phrasing locked ("in Brisbane" / "on the Gold Coast" / "on the Sunshine Coast"), state anchor "Queensland" retained by comma. See fixed-template block above. All nine strings now final.
2. **Dr Sutherland review point — OUTSTANDING.** The commercial-litigation Constraint-4 decision and its thin-cell limitation should be put to Dr Sutherland before OSF lock (see CL limitation note above). Not a wording change — a validity sign-off. This is the only item standing between Step 1 and pre-registration.

Lock log

Date	Practice area	Term locked	Governing constraint
28 Jul 2026	Family law	family lawyers [location]	Noun C1; descriptor C2
28 Jul 2026	Personal injury	personal injury lawyers [location]	Noun C1; descriptor C2 over C3
28 Jul 2026	Commercial litigation	commercial litigation lawyers [location]	C4 (construct fidelity over consistency)
28 Jul 2026	Location phrasing (all 9)	"in Brisbane" / "on the Gold Coast" / "on the Sunshine Coast", ", Queensland" anchor	Ecological validity; state anchor required by incognito protocol (per-platform account conditions, methodology v2.3 Section 7)

Revision notes

28 July 2026 — alignment to methodology v2.2. No query wording, term selection, constraint, or lock decision was changed. Three consistency updates only:

1. **Protocol phrasing:** “incognito, logged out, with location services disabled” replaced with “incognito, with location services disabled and no personalised account history (per-platform account conditions, methodology Section 7)” — Claude requires a logged-in account, so account state is fixed per platform in the methodology rather than described as uniformly logged out. The state-anchor rationale is unaffected.
2. **Owner line:** “Jane Cluff / Claude” replaced with “Jane Cluff (AI-assisted drafting disclosed in methodology)” for consistency with the methodology’s AI assistance disclosure, under which all methodological decisions remain the sole responsibility of the researcher.
3. **Cross-references:** “methodology v2” updated to “methodology v2.2” throughout.

28 July 2026 — alignment to methodology v2.3. Cross-references updated from v2.2 to v2.3. No other change: nothing in methodology v2.3 touches query wording, term selection, or the Section 7 account conditions this document relies on.