AI Visibility in South East Queensland Legal Services

Findings report

Jane Cluff, Felicity Jane Digital. Published 4 November 2026. Version 1.0.

An exploratory observational study. Methodology version 2.4, registered on the Open Science Framework (OSF) on 3 August 2026: osf.io/wcvbk. Post-registration notes, with firms replaced by study codes: osf.io/c63dk.

Dr Karen Sutherland reviewed Version 1 of the methodology in July 2026. Her review was of the methodology only. The study is an exploratory, observational study designed to generate hypotheses rather than establish causal or confirmatory conclusions.

Contents

1. Executive summary
2. About this research
3. Introduction
4. Method
5. Findings: Part 1, consistent appearance (5.1 to 5.8); Part 2, Signals A, D, E and F (5.9 to 5.14); Part 3, Signals B and C (5.15 to 5.19)
6. Limitations
7. Conclusions
8. Changes from the registered method
9. Conflict of interest
10. Acknowledgements
11. Data availability and retention
12. References

Appendices: A. The nine registered queries; B. Signal definitions; C. Run conditions; D. Glossary.

Terms used in this report are defined in the glossary (Appendix D).

1. Executive summary

This study was run by Felicity Jane Digital, a marketing agency that provides search and AI-visibility services to law firms. Several of the signals it measures, including structured data, Google Business Profile completeness and website content, are areas in which the agency sells services. Section 9 sets out this interest.

What the study found. Firms that AI platforms named consistently were more likely than comparison firms to show four observable signals: a listing in the QLS Find a Solicitor directory, structured data on their website, a more complete Google Business Profile, and a more recent website post. The QLS directory gap was the largest, at 76% against 40%, and it held in all nine practice area and location combinations. Review count, recorded for commercial litigation only, did not follow this pattern. About two in five firm entries (one firm for one query on one platform) appeared in one run only (335 of 810). Agreement between platforms was low: just over half of the consistent results (38 of 74) reached the threshold on one platform only. These are associations in a small exploratory study. They do not show that any signal causes a firm to be named. Section 7 gives the result for each registered proposition.

The study is exploratory. It ran nine fixed queries, for three practice areas in three South East Queensland locations, 10 times each on four AI platforms in August and September 2026: 361 recorded runs. It then compared observable signals of the firms the platforms named consistently with those of a random sample of firms the platforms did not name for the same query.

Every comparison below is a correlation between small, unevenly balanced groups, observed after the runs. None shows that a signal causes a firm to be named. Established firms are likely to score well on many of these signals at once, and the study cannot separate any one signal from that (section 6.1). The two groups are also the two ends of the range: consistent firms were named in at least 7 of 10 runs on a platform, and comparison firms in none of that query's runs. Firms named in 1 to 6 runs on every platform are in neither group, and the differences may be wider than they would be across all firms (section 6.1). Every figure below is set out, with its conditions, in the section cited.

- **Firms named and consistent appearance.** Counting each firm once per query and platform gave 810 entries, and about two in five appeared in one run only. In total, 137 firm, platform and query combinations reached the threshold of 7 of 10 runs. They involve 74 consistent rows (one firm for one query) and 63 distinct firms (sections 5.2 and 5.3).
- **Agreement between platforms.** Of the 74 consistent rows, 38 reached the threshold on one platform only, and 7 on all four (section 5.4).
- **Organic ranking.** About half of the consistent rows (39 of 74) ranked in Google's organic top 10 for the same query. The comparison figure is withheld under reporting rules that were set after it had been seen (sections 4.10 and 6.2). These figures cannot show whether organic ranking and consistent appearance are related (section 5.10).
- **Website content recency.** Among those with a dated post, the median time since the last post was 15 days across consistent rows (27 days counting each firm once), against 101 days across comparison firms (section 5.11).
- **QLS directory.** 76% of consistent rows were listed in the Queensland Law Society's Find a Solicitor directory (95% interval 65% to 84%), against 40% of comparison firms (27% to 55%), in the same direction in all nine practice area and location combinations. LawConnect showed little difference (section 5.13).
- **Review count, commercial litigation only.** The median was lower for the 24 consistent firms than for the 15 comparison firms: 34.5 reviews against 60, an exploratory observation (section 5.12).
- **Structured data.** Detected for 62% of consistent rows (95% interval 51% to 72%), against 31% of comparison firms (20% to 46%) (re-check: reliable, alpha 1.000) (section 5.16).
- **Google Business Profile completeness.** Listings scoring 4 or 5 of 5: 43% of consistent rows (95% interval 33% to 55%), against 24% of comparison firms (14% to 39%) (re-check: reliable, alpha 0.905) (section 5.17).
- **Google Business Profile post recency, Google AI Overviews only.** A post within 90 days was on the listing for 33% of rows consistent on Google AI Overviews (95% interval 19% to 51%), against 18% of comparison firms (9% to 31%), an exploratory observation (re-check: reliable, alpha 0.905) (section 5.18).

Key limitations

- One researcher collected and coded all the data, with no second coder. The two re-checks were made by the same person, on samples of 37 runs and 13 records, with help from an AI assistant, which departed from the method set for them (sections 4.9 and 6.2). Signals A, D, E and F were not re-checked.
- An AI assistant from the same provider as one of the platforms studied (Claude) transcribed the data and entered corrections. Of the 57 AI Overview corrections it entered, the researcher checked a random 6 (sections 4.11 and 6.2).
- Several rules were set after results were seen. The name-merging rule moved the headline count from 125 to 148 consistent combinations, and the count became 137 after a later AI Overview counting rule, a recount that was not blind, and other corrections (section 6.2).
- The reporting rules, and several decisions on what to publish, were set with the figures known (sections 4.10 and 6.2).
- Signal log values were corrected after the re-check lock, and one set of changes made on 19 October 2026 has no record of who made it (section 6.2).
- Section 8 summarises the 54 post-registration amendments that record these changes.

2. About this research

About this research

This research is descriptive, not evaluative. It does not constitute a quality assessment of any Queensland law firm. A firm's presence, absence or frequency in the AI-generated outputs sampled should not be read as an indication of legal competence, service quality, client satisfaction or professional standing.

Who conducted this research

This research was conducted and is published by Felicity Jane Digital, a marketing agency that provides AI-visibility and search optimisation services to businesses, including law firms of the type examined in this study. Felicity Jane Digital is not an independent third party in relation to the market this research examines, and readers should take this relationship into account when considering the findings.

Methodology and scope

Findings are based on a sample of outputs generated by ChatGPT (OpenAI), Perplexity AI, Google AI Overviews and Claude (Anthropic) in response to defined queries relating to legal services across the Brisbane, Gold Coast and Sunshine Coast regions of Queensland. Queries were run between 10 August 2026 and 22 September 2026. Google results are the AI Overview as first displayed, and Perplexity was queried through a signed-in study account. Results were compared against a population of 805 Google Business Profile listings for law firms in the defined practice areas and locations. A firm with more than one qualifying office or practice area may appear more than once in that count. Firms without a disclosed street address were excluded, so the population is not a census of Queensland law firms. The registered methodology, sampling parameters and prompt design are available at osf.io/wcvbk, with post-registration notes in the associated project at osf.io/c63dk.

How firms are reported

Apart from ten firms listed by website domain in the registration as examples of organic search results, no firm is identified. This publication does not name those ten firms. Figures that describe firms' characteristics rest on at least 10 distinct firms, and comparisons are pooled by practice area. For those figures, where 90% or more, or 10% or fewer, of a group share one value, the figure is combined with others or not reported. Counts that would point to a single firm are not published, with two exceptions: one difference of one, disclosed beside the figures it concerns, and the counts for the adviser's firm in the conflict-of-interest paragraph below.

A snapshot, not a current or ongoing status

AI-generated outputs are not static or deterministic. The same or similar queries may produce different results on a different date, with different wording, or on a different platform. The findings in this publication reflect a snapshot as at the sampling dates stated above and do not represent an ongoing, current or guaranteed position for any firm or category. Website and Google Business Profile signals were observed between 28 September and 5 October 2026.

Use of AI tools in this research

An AI assistant, Claude (Anthropic), was used to help transcribe the saved screen captures into the study's data logs. The researcher checked and entered every transcribed entry. Corrections to 57 Google AI Overview records, made on 13 October 2026, were entered by the assistant from the researcher's recorded decisions, and the researcher checked a random 6 of them. Corrections to the website, listing and directory records, made from 19 to 21 October 2026, were also entered by the assistant, each with the researcher's approval. One further set of changes to those records, made on 19 October 2026, has no record of who made it or of any approval; the researcher reviewed and decided each change on 21 October 2026. The assistant also drew the re-check samples, ran comparisons, calculated figures and helped draft study documents, under the researcher's direction. Where Claude (Anthropic) is one of the platforms studied, its responses were collected through a separate study account.

Two re-checks were made by the researcher, coding a second time from the saved records, without seeing the original coding and at least two weeks after it. They are consistency checks by the same person, not independent reviews.

- Firm mentions: 37 of 361 runs, drawn at random, re-coded on 8 October 2026. Positive specific agreement 0.876 (95% interval 0.805 to 0.937), against a reference point of 0.800. That reference point was designed for a measure that corrects for chance, so it is used here as an approximation.
- Website structured data and Google Business Profile completeness: 13 of 123 firm records (one record per firm per practice area and location), drawn at random, re-scored on 19 October 2026. Structured data: 13 of 13 records agreed (95% interval 0.772 to 1.000; alpha 1.000). Google Business Profile completeness: 62 of 65 element judgements agreed (95% interval 0.862 to 1.000; alpha 0.905). Reference point 0.800.

The samples are small, so these figures are approximate.

In both re-checks the researcher did all the coding and gave every value. At her request, the AI assistant explained the process for the first records and, before each sheet was locked, checked it for empty cells. When the researcher raised a coding question, it pointed her to the written rule or rules that applied, without saying how they applied. In the second re-check that is disclosed as arguably advice on a scoring judgement. In the second re-check it also read out two records' dates, and after the lock it read three evidence files to confirm the researcher's reading of the disagreements, all in one record. These depart from the method set for the re-checks and are disclosed in the post-registration notes.

Reviewer conflict of interest

A legal adviser was engaged to review the study's methodology and publication materials: defamation and identifiability risk, data retention, the adequacy of the study's disclaimer under the Australian Consumer Law, and claims about AI. The adviser's own firm appeared in the AI-generated results for one of the study's nine queries, in 4 of 10 runs on each of two platforms. That is below the threshold of 7 of 10 runs used to define consistent appearance. The firm was captured under the same protocol as every other firm, and the adviser took no part in collecting, coding or analysing the data. The adviser declined to advise on the wording of this disclosure, citing Rule 12 of the Australian Solicitors' Conduct Rules, and no lawyer has reviewed it. Neither the adviser nor the firm is named.

No endorsement or affiliation

This research is not affiliated with, endorsed by, sponsored by, or reviewed by OpenAI, Perplexity AI, Google or Anthropic.

Verification

The registered methodology is publicly available for independent review at osf.io/wcvbk. The post-registration notes, with firms replaced by study codes, are available at osf.io/c63dk. Apart from the registration, records that name firms are kept privately. For each named record with a published coded version, a SHA-256 fingerprint of the named original is published beside it, so the private record can be matched to it without identifying any firm.

Scope of this disclaimer

Nothing in this disclaimer is intended to exclude or limit any liability Felicity Jane Digital may have under the Australian Consumer Law or otherwise for the contents of this publication. This disclaimer is intended only to clarify and qualify the nature, scope and limitations of the findings presented in this document.

3. Introduction

3.1 Why this study was done

When a person asks an AI platform to recommend a lawyer, the platform answers with a short list of firms. Those firms enter the person's consideration set. When the study was designed, the researcher found no Queensland-specific data on which law firms the main AI platforms name (methodology v2.4, Section 2).

Earlier studies show that AI recommendation lists change from one run to the next. Fishkin and O'Donnell (SparkToro, January 2026) made 2,961 prompt runs and found the chance of the same brand list appearing twice was under 1 in 100. Schulte et al. (arXiv 2604.07585, 2026) and Sielinski (arXiv 2603.08924, 2026) found the same instability and argued that AI visibility must be measured repeatedly and reported with uncertainty. This study follows that method. Each query was run 10 times on each platform, across two waves, and frequencies are reported with confidence intervals.

3.2 Research question

The registered research question is:

> 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?

3.3 Research propositions

The study registered three observational propositions and three exploratory observations (methodology v2.4, Section 10). They are quoted here as registered.

> 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.

The quotations keep the registered wording, including its dashes. Both Ahrefs figures in Proposition 1 were checked at source on 16 October 2026: 38% of AI Overview citations from top-10 pages in March 2026, against about 76% in July 2025 (Linehan, 2026), and 13.7% citation overlap between AI Mode and AI Overviews (Gavoyannis, 2025). The registered text attributes the fall in part to improved citation parsing. The Ahrefs article gives query fan-out as its main explanation and notes the improved parsing without presenting it as a cause.

3.4 What kind of study this is

This is an exploratory observational study. It observes patterns in publicly available data and generates propositions for later confirmatory research. The registered methodology sets out what it 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.

Every association in this report is a correlation. Firms that AI platforms name consistently are likely to be established practices that score well on many signals at once, and this study cannot separate the effect of any one signal from that (methodology v2.4, Sections 2 and 12).

3.5 Who ran the study

The study was carried out by Jane Cluff of Felicity Jane Digital, a marketing agency that provides search and AI-visibility services to businesses, including Queensland law firms. The findings may inform its services. Section 9 sets out this interest, and the disclaimer in section 2 applies to every finding.

3.6 How this report is organised

Section 3.7 summarises earlier studies. Section 4 describes the method as it was carried out. Section 5 gives the findings in three parts: consistent appearance (Part 1), Signals A, D, E and F (Part 2), and Signals B and C (Part 3). Section 6 lists the limitations. Section 7 sets out the conclusions: the answer to the research question, the result for each registered proposition, agreement between platforms, propositions for later confirmatory research, and directions for further research. Section 8 summarises every change from the registered method. Sections 9 to 12 cover the conflict of interest, acknowledgements, data availability and references.

3.7 What earlier studies show

This review was conducted as a semi-systematic, concept-centric review (Snyder, 2019; Webster & Watson, 2002). Sources were identified through databases including arXiv, the ACM Digital Library, and the ACL Anthology, with Google Scholar used for forward and backward citation chasing and named-vendor publications consulted for industry studies. The search covered work from November 2023, when the generative engine was first formalised, through October 2026, with the date a system was tested recorded separately from its publication date. Searches were first run in August 2026 and updated in October 2026; the October update searched arXiv and the open web on 11 and 12 October, and the ACL Anthology and the ACM Digital Library on 12 October. Studies were included where they measured AI answer-engine visibility or contributed measurement methodology; critical appraisals of the field were also included. Undated vendor claims without a disclosed method were excluded. Each source was graded by evidentiary weight, from peer-reviewed work through preprints to industry studies with a disclosed method, and every source was accessed and verified at its primary origin before inclusion.

3.7.1 The generative-engine shift

Search is no longer only a ranked list of links. Aggarwal et al. (2024) formalised the "generative engine": a system that retrieves documents and uses a large language model to synthesise them into a single grounded answer. In this setting, visibility is whether, and how prominently, a provider is named or cited within that synthesised answer. Citations are dispersed through the response, and a provider's prominence depends on placement inside it.

This shift removes a control that content providers once held. The retrieval-and-synthesis pipeline is opaque and fast-changing, so a provider has limited insight into when or how it is surfaced. Working against a benchmark of roughly 10,000 queries across nine datasets, Aggarwal et al. (2024) showed that content can still be deliberately optimised for generative-engine visibility: the largest gains came from adding verifiable statistics and citing credible sources, on the order of 30 to 40% on a position-adjusted word-count metric. Legacy tactics such as keyword stuffing produced negligible or negative effects.

That demonstration defines the problem without settling it. Martinez's (2026) later critical survey places the finding in context: the reported gains hold within their experimental setting; they remain conditional on a source already being present in a fixed retrieval context, and they establish neither organic discoverability nor durable downstream traffic. In-context optimisation is possible. The prior question, whether a given firm is retrieved and recommended at all under live conditions, stays open, and it is the behaviour this study observes directly.

Google's AI Overviews show how far this surface departs from ranked results. They appear selectively and draw on sources beyond the ranked list. Xu et al. (2026) issued 55,393 trending queries from a United States setting between 13 March and 21 April 2026. An AI Overview appeared for 13.7% of queries overall and 64.7% of queries phrased as questions. On average, 29.8% of the domains an AI Overview cited did

not appear anywhere on the corresponding first page of results.

3.7.2 AI recommendation is non-deterministic

The same question, asked twice, rarely returns the same answer. Fishkin and O'Donnell (2026) ran 12 recommendation prompts about 2,961 times across ChatGPT, Claude, and Google's AI, drawing on 600 volunteers over late 2025. Asked for brand recommendations a hundred times, ChatGPT and Google's AI returned the same list on fewer than one occasion in a hundred, and Claude only slightly more often; the same list in the same order was rarer still, close to one in a thousand. The number of items returned also varied between runs. One layer proved steadier. In some categories, a core set of well-established names recurred in a majority of responses even as their order churned, so appearance in the consideration set was more stable than any single position within it.

This instability has since been characterised in statistical terms. Sielinski (2026) treats a visibility metric as a sample estimator of an underlying response distribution and shows that many apparent differences between domains fall inside the measurement noise floor, where the gap cannot be separated from sampling variation at typical run counts. Schulte et al. (2026) reach a compatible conclusion through repeated measurement. Across four engines (ChatGPT, Gemini, Google AI Mode and Perplexity), tested on Swiss-German queries in four consumer and commercial categories over a 45 to 46 day window in early 2026, the day-to-day overlap of cited sources averaged a Jaccard similarity of 0.34 to 0.42. Roughly two-thirds of cited sources changed from one day to the next. They describe this day-to-day instability as a persistent characteristic of AI search. A further preprint found almost no overlap in the URLs four engines cited for the same prompt, and 67% mean turnover within three of those engines between two consecutive days, which the author treats as a diagnostic, on 15 prompts about AI-visibility software, with the cross-engine comparison drawn from a single day (Tannenbaum, 2026).

Google's own surfaces show the same pattern. A peer-reviewed study of 11,500 queries compared Google's organic results, AI Overviews and Gemini (Grossman et al., 2026). The sources each returned for the same query overlapped little, with average Jaccard similarity below 0.2, and in a 100-query consistency test AI Overviews were less consistent than organic results across two runs of the same query (Jaccard 0.66 against 0.78). The main data were collected in December 2025 in a United States setting.

A second peer-reviewed study found the same pattern across runs about two months apart. Kirsten et al. (2026) compared Google's organic results with five generative search systems, including AI Overviews. The sessions were logged out, in a United States setting, and the queries were run in 2025. Between the two runs, the links AI Overviews retrieved overlapped 18%, against 45% for organic results.

Four of the sources in this section have commercial ties. Fishkin and O'Donnell's co-author had joined Gumshoe.ai, an AI-visibility tracking company, which the report itself flags as a possible conflict of interest; Schulte et al.'s first author is affiliated with Aurora Intelligence, whose data export supplied the daily dataset; Sielinski is affiliated with IQRush, which sells AI-visibility measurement software; and Tannenbaum founded Aiso Boost, which develops AI-visibility software and collected the data. The peer-reviewed studies by Grossman et al. and Kirsten et al. have no stated commercial ties. Each of the four is used as methodological support only, with no endorsement of any product.

The methodological implication is shared across these studies. A single reading cannot represent a distribution, so visibility must be sampled repeatedly and reported with its uncertainty. Schulte et al. (2026) call for repeated measurement and for treating visibility as a distribution; Sielinski (2026) frames the requirement as reporting an interval around each estimate. A peer-reviewed study formalises the same approach. Mehta et al. (2026) define visibility as the probability that an entity is cited. They estimate it from ten independent runs per query and report it with 95% Wilson score intervals, the interval this study uses. Their data come from the Indian market. Their main contribution is an optimisation method, which this review does not rely on, and the paper has no stated commercial ties. This study adopts repeated sampling with interval reporting, and makes no claim about a firm's position within an AI answer. Three limits carry forward. These precedents measured other query domains in overseas markets, and this study applies their method to a different domain, Queensland legal services. The platform sets also differ. Two of Schulte et al.'s four engines, ChatGPT and Perplexity, are studied here; it did not test Claude, and it excluded the Google AI Overviews data it collected, which this study includes (recorded as first displayed; section 4.2). The largest precedent used 60 to 100 runs per prompt per platform (Fishkin & O'Donnell, 2026), above what a single researcher can sustain, so the present study's smaller run count keeps its findings exploratory.

3.7.3 An exploratory design for an emerging field

This review is semi-systematic and concept-centric. It organises the field around concepts, and it works with a body of research that different groups have defined and measured in different ways. That is the condition a semi-systematic approach is built for (Snyder, 2019; Webster & Watson, 2002). The same condition shapes the study the review introduces. Exploratory research works without a set hypothesis. It develops the questions that later confirmatory research can test, and it is the appropriate design where a field is too new for a well-formed hypothesis (Schwab & Held, 2020). Applied empirical guidance adopts the same distinction (Eimer et al., 2025). AI recommendation of Queensland legal services is such a field: no prior study has reported which Queensland firms AI platforms recommend, and the behaviour itself is only a few years old.

The instability documented above sets the measurement requirement. A single reading is one draw from a distribution, so a defensible estimate must come from repeated observation reported with its uncertainty (Schulte et al., 2026; Sielinski, 2026). This study meets that requirement within the limits of a solo researcher. It runs each query multiple times across two waves and reports appearance with approximate 95% Wilson confidence intervals.

The exploratory frame is a deliberate choice about what can honestly be claimed. A confirmatory study asserting statistical significance at the run counts one researcher can sustain would overstate its precision. An exploratory study that quantifies its uncertainty and discloses it states only what the data support. The distinction between exploratory and confirmatory work is a general principle of empirical method (Schwab & Held, 2020); here it sets the ceiling on the study's claims. The findings are directional patterns and propositions for later

confirmatory research. Appearance figures are reported with an approximate interval that shows their uncertainty.

3.7.4 The gap this study addresses

The research on optimising for generative engines has grown quickly, and a recent survey has weighed how much of it holds up. Martinez (2026) reviewed 45 studies published between late 2023 and mid-2026, grading each by evidentiary weight, and found the evidence base heterogeneous: terminology and standards vary from study to study. Its central finding is that no reviewed technique shows a stable, cross-platform causal effect on organic discoverability. This is the state of knowledge the present study enters.

That state suits an observational design. A study that records which signals co-occur with appearance, without claiming any signal causes it, produces propositions a later controlled study can test. The literature this review draws on is itself young, resting substantially on preprints and industry studies.

The field-level gap takes a specific form in this market. Industry studies have begun to measure the legal sector. A vendor study of 540 lawyer-hiring prompts across 30 US metros on Perplexity's Sonar model, each run three times, found that a legal directory was the first source cited in 77.8% of answers (InterCore Technologies, 2026). An Australian study by Integral Media, which sells SEO and AI-visibility services to law firms and discloses that interest, asked 25 legal-services questions on four search and AI surfaces (Morrison, 2026). Each question ran once on each surface, in a single session on 21 September 2026. Twenty of the 25 questions named one of five cities, including Brisbane. Twelve of 73 usable answers met the study's safe-and-actionable standard. Claims about named firms were not independently verified and were excluded from its public accuracy findings. This review found no study of which firms AI platforms recommend for legal services in an Australian market, and none that combines a defined population of local firms, repeated runs across two waves reported with confidence intervals, and a comparison of observable signals between consistently appearing firms and a random sample of firms not named for the same query. South East Queensland legal services are a defined, localised market, and what AI platforms return when asked to recommend firms there is unmeasured.

This study addresses that gap. It asks 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. It observes nine query combinations (three practice areas across three locations), each sampled repeatedly on each platform. It reports the signals associated with consistent appearance as descriptive correlations, with intervals on each group's share and no statistical test (Amendment 47). The method that follows sets out how each term in that question is operationalised.

Table 1. Concept matrix: reviewed sources against the concepts organising this review.

SOURCE	GENERATIVE-ENGINE SHIFT	INSTABILITY / NON-DETERMINISM	REPEATED-SAMPLING MEASUREMENT	EVIDENCE GRADING / NO STABLE CAUSAL EFFECT	LOCAL LEGAL MARKET
Aggarwal et al. (2024)	✓				
Fishkin & O'Donnell (2026)		✓	✓		
Sielinski (2026)		✓	✓		
Schulte et al. (2026)		✓	✓		
Tannenbaum (2026)		✓			
Grossman et al. (2026)		✓			
Kirsten et al. (2026)		✓			
Xu et al. (2026)	✓				
Mehta et al. (2026)			✓		
InterCore Technologies (2026)			✓		✓
Morrison (2026)					✓
Martinez (2026)	✓			✓	

Note. Rows are the empirical sources reviewed; columns are the concepts the review is organised around. The methodology-of-review references (Snyder, 2019; Webster & Watson, 2002; Schwab & Held, 2020; Eimer et al., 2025) are excluded by design: they are the method by which this review was conducted, not contributions to the field's concepts. Two reviewed sources address local legal markets. The United States study measures cited sources without a defined firm population; the Australian study measures answer quality from single runs and excludes named-firm claims. No reviewed source measures which firms appear in AI recommendations for an Australian legal market. That absence is the gap this study addresses.

4. Method

This section describes the method as carried out. Where it differs from the registered methodology, the post-registration amendment is cited. Section 8 lists every change.

4.1 Practice areas, locations and queries

The study covers three practice areas (family law, personal injury and commercial litigation) in three locations (Brisbane, the Gold Coast and the Sunshine Coast). The practice areas and locations were chosen to give contrast in search behaviour and market size. They are a purposive selection, and they do not represent Queensland legal practice as a whole (methodology v2.4, Section 6).

Each of the nine combinations has one fixed query, locked before registration (Appendix A). Every query follows one form: "Who are the best [practice area] lawyers [in or on] [location], Queensland?" The word "best" is a device to prompt a list of firms. No finding treats a platform's use of it as an assessment of a firm (methodology v2.4, Section 5).

All 361 recorded runs used the query text exactly as registered (Amendment 15).

4.2 Platforms and run conditions

Four platforms were studied: ChatGPT (OpenAI), Perplexity AI, Google AI Overviews and Claude (Anthropic). Web search was enabled on ChatGPT and Claude, so that all four platforms were set to retrieve current web content at the time of each query. ChatGPT and Claude still draw on training data as well, in proportions their providers do not disclose (methodology v2.4, Sections 4 and 12).

Every run used a new session in its own Chrome incognito window, on one device and one Sunshine Coast internet connection, with location services off and no VPN. The Google organic check for each run was made in the same window, straight after the AI response. For AI Overview runs, the organic results come from the same Google search (methodology v2.4, Section 7).

Account conditions as run:

- **ChatGPT:** logged out, as registered. No deviation was recorded.
- **Claude:** a dedicated study account with no prior history, and memory and personalisation disabled where available, as registered.
- **Perplexity:** registered as logged out. Perplexity began to require sign-in during the first study session, so 87 of 90 Perplexity runs were made signed in to a dedicated study account with memory inactive. Earlier study threads were visible in the account's sidebar at one run, so some carry-over of account context cannot be ruled out (Amendment 15, Change 1; Amendment 37).
- **Google AI Overviews:** incognito Google Search with no account signed in, as registered. The registered method and the registered run-log template said to expand the Overview before recording (Amendment 17). The researcher decided before the first run to record the Overview as first displayed (Amendment 15, Change 2). Some captures show a full Overview, and whether it was expanded or displayed in full by Google cannot be told from the capture, so consistent application cannot be confirmed (Amendment 37). Under a recording rule set after collection, firms were counted from the Overview as displayed in the main capture only (Amendment 29).

4.3 Runs and waves

Each query was run 5 times on each platform in each of two waves, giving 10 runs per query per platform. Runs of the same query on the same platform were made on separate days, at least 24 hours apart.

- Wave 1: 10 to 18 August 2026.
- Wave 2: 14 to 22 September 2026, 27 days after Wave 1 closed (the registered minimum was 21 days).

That is 360 runs. One more run replaced a Google AI Overviews run that showed no Overview, for 361 recorded runs. The replacement was made on 22 September 2026, outside the Wave 1 window, and is counted in Wave 1 as logged (Amendment 15, Change 4). A second run with no Overview, in Query 03, could not be replaced, so that query's AI Overview figures are out of 9 runs (Amendment 33).

4.4 What was recorded at each run

A firm was recorded when the platform presented it as a legal service provider to contact, in body text, a list or a firm card. Directories and ranking guides cited as sources, firms that appeared only in citation panels, and sponsored placements were left out. Out-of-area firms and barristers' briefing entries were also left out, under a rule formalised during collection and not applied consistently (Amendments 11 and 37). Each firm was recorded once per run, in order of first appearance.

The registered rule was to record a count of named individuals, never their names. Where a response identified a named lawyer's firm as the lawyer's practice or employer, the firm was recorded (methodology v2.4, Section 7). Names were in fact recorded in notes or organic entries for 63 runs, and in five firm entries. All have since been removed from the study records, and the named originals are held privately for the retention period (Amendments 15, 38 and 44). Names recorded in the signal log's notes were removed on 19 October 2026 (Amendment 49, Change 5).

For each run, the Google organic top 10 was recorded from the same window, by source name as displayed, with sponsored results, the map pack and other modules left out.

A full-page capture of every AI response and every Google results page was saved at the time of the run. Coding decisions can be checked against those captures, apart from a few partial captures listed in Amendment 29.

4.5 Consistent appearance

A firm is a "consistent appearance" for a query on a platform when it is named in 7 or more of the 10 runs. This threshold was set before data collection. It is a working definition for an exploratory study (methodology v2.4, Section 11). For Query 03's AI Overviews, out of 9 runs, the threshold is 7 of 9 (Amendment 33).

Frequencies are reported with approximate 95% Wilson score intervals. At 10 runs these intervals are wide, and that width is part of the finding. The intervals assume runs are independent, which repeated runs of one query on one platform may not be, so their coverage is approximate (methodology v2.4, Section 11). Shares in the signal comparisons carry the same intervals, for each group separately (Amendment 47).

4.6 Population

The population is every law firm with an active Google Business Profile (GBP) listing for the practice area and location, found by a fixed Google Maps search using the practice-area term from the query plus the location. Listings had to fall within the local government area: City of Brisbane, City of Gold Coast, or Sunshine Coast Regional Council and Noosa Shire Council for the Sunshine Coast (methodology v2.4, Section 8).

Before capture, filters were added for practice-area relevance and entity type, and listings with no disclosed street address were excluded (Amendments 1 to 6). The population was captured on 6 August 2026 and verified on 7 August, before Wave 1, when ten listings with no disclosed address and other doubtful rows were corrected (Amendment 7). The corrected population is 805 listings (Amendment 37). A firm with more than one qualifying office or practice area can appear more than once in that count. Because of the address rule, the population is not a census of Queensland law firms.

4.7 Signals

Six signals were recorded for each consistent firm and for each comparison firm (section 4.8). Signal A comes from the organic checks made at each run. Signals B to F were observed once for each firm in each practice area and location, after both waves (Amendment 10), between 28 September and 5 October 2026 (Amendments 19 to 26).

- **Signal A, Google organic ranking:** whether the firm appeared in Google's organic top 10 for the same query, from the checks recorded at each run. The rule for combining a row's checks is in Amendment 12 and section 5.10 (Proposition 1).
- **Signal B, structured data:** whether the firm's website exposes detectable FAQPage, LocalBusiness or LegalService, or Person structured data, recorded as detected valid, detected with errors, or not detected in the Schema.org validator, on three fixed pages: the homepage, the matching practice-area page and the first lawyer profile (Proposition 2; Amendment 16).
- **Signal C, GBP completeness:** a score of 0 to 5 for a business description, a primary category, a services section, the practice's own photos, and a GBP post within 90 days (Proposition 3).
- **Signal D, website content recency:** the date of the most recent blog, news or journal post on the firm's own website (Proposition 3).
- **Signal E, Google review count:** recorded for commercial litigation only. At registration, personal injury was left out because of the restrictions on personal injury advertising under the Personal Injuries Proceedings Act 2002 (Qld), which limit advertising to a practice's name, contact details and areas of practice (s 66), and family law because client confidentiality is likely to depress review volumes (exploratory observation).
- **Signal F, legal directory presence:** listing in the QLS Find a Solicitor service and in LawConnect. The third registered directory, FindLaw Australia, had ceased publication in December 2022, so Signal F is scored 0 to 2. The QLS search method was changed during capture (exploratory observation; Amendments 19, 20, 27, 37 and 39).

The detailed recording rules set during capture are in Amendments 19 to 26.

4.8 Comparison sample

For each query, 5 firms were drawn at random from the population listings that no platform named in that query's runs, giving 45 comparison firms (methodology v2.4, Section 11). Six drawn firms had in fact been named in their own query's runs and were replaced by the same random method (Amendment 12, Change 4; Amendment 13, Change 1). Comparison firms that were consistent in another practice area or location were kept and disclosed. Four Gold Coast comparison firms had changed their GBP category since the population capture, and were kept and disclosed (Amendment 20).

4.9 Consistency re-checks

The registered method requires a blind re-code of a random 10% sample after a gap of at least two weeks (methodology v2.4, Section 7). Two re-checks were made by the researcher, coding a second time from the saved records, without seeing the original coding and at least two weeks after it (Amendments 15, 28, 30 and 53):

- **Firm mentions:** 37 of the 361 runs, re-coded on 8 October 2026 (Amendment 28). Positive specific agreement 0.876 (95% interval

0.805 to 0.937), against a reference point of 0.800 (Amendment 28; section 5.8). That reference point was designed for a measure that corrects for chance, so it is used here as an approximation. The two-week gap is measured as set out in Amendment 28; section 6.2 notes how it reads when measured from the end of all data collection.

- **Structured data and GBP completeness:** 13 of the 123 firm records, re-scored on 19 October 2026 (Amendments 30 and 48). Structured data: 13 of 13 records agreed (95% interval 0.772 to 1.000; alpha 1.000). GBP completeness: 62 of 65 element judgements agreed (95% interval 0.862 to 1.000; alpha 0.905). Both signals are banded Reliable (Amendment 48, Change 5). All three disagreements were in one record, and were an error in the re-check; the original stands (Amendment 48, Change 6).

In both re-checks the researcher did all the coding and gave every value. At her request, the AI assistant explained the process for the first records and, before each sheet was locked, checked it for empty cells. When the researcher raised a coding question, it pointed her to the written rule or rules that applied, without saying how they applied. In the second re-check that is disclosed as arguably advice on a scoring judgement. In the second re-check it also read out two records' dates, and after the lock it read three evidence files to confirm the researcher's reading of the disagreements, all in one record. These depart from the method set for the re-checks and are disclosed in the post-registration notes (Amendment 48, Changes 2 and 6; Amendment 52, Change 4; Amendment 53).

The measures and the reporting rules for both were fixed before each re-check (Amendments 28, 30 and 40). They are consistency checks by one person. Independent coding by a second person was not feasible for a single researcher (methodology v2.4, Section 7). Signals A, D, E and F, and Google's Rich Results Test, were not re-checked.

4.10 How figures are reported

Apart from ten firms listed by website domain in the registration as examples of organic results, no firm or individual is identified in this report. This report does not name those ten firms. Findings are reported at practice-area and location level. Comparisons are pooled by practice area, with a whole-study figure, and each figure about firms' characteristics rests on at least 10 distinct firms (Amendment 41). Where 90% or more, or 10% or fewer, of a group share one value, the figure is combined with others in an order fixed in advance, or not reported (Amendment 43). Counts that would point to a single firm are not published, with two exceptions: one difference of one, disclosed beside the figures it concerns, and the counts for the adviser's firm in the conflict-of-interest paragraph in section 9 (Amendment 45, Change 4(e); Amendment 52, Change 6). These rules were set after the Part 1 and Part 2 figures had been calculated, and before the Signal B and C re-check (Amendments 43 and 45). The second exception was added on 22 October 2026, after the re-check and with all figures known (Amendment 52, Change 6).

Three applications of these rules were decided with the figures known. Medians are published: the researcher reads the single-firm rule as covering a value a reader can attribute to one firm, as a minimum or maximum can be, and not a median of an unnamed group. Four published medians each equal one record's value, and each rests on at least 15 firms (Amendment 51, Change 9(e)). Read the other way, publishing them departs from Amendment 45, Change 4(d). Google's Rich Results Test is not reported for either group (Amendment 52, Change 2(c)). The photo element of GBP completeness (C4) is published for both groups. The rule that would hold it back exists to stop a smaller base revealing how many firms lack premises. No firm in either group is recorded as without premises, and the C4 bases equal the group sizes. Publishing it departs from that rule for this case (Amendment 52, Change 3).

4.11 Use of AI assistance

An AI assistant, Claude (Anthropic), helped transcribe the saved captures into draft log entries. The researcher reviewed, corrected and entered every entry, apart from 57 Google AI Overview corrections made on 13 October 2026, which the assistant entered from the researcher's recorded decisions and of which the researcher checked a random 6 (Amendments 15 and 32). The assistant also transcribed the website and GBP signal values from the saved evidence, for the researcher to check and enter; drew the re-check samples; ran comparisons, calculations and record audits; and helped draft study documents, under the researcher's direction. It made the corrections to the signal log through a spreadsheet connection, each approved by the researcher (Amendments 49 to 51). On the evening of 19 October 2026, 15 cells were changed with no record of who made the changes or of any approval; each was checked at source and decided by the researcher on 21 October (Amendment 51, Change 1). Separate instances checked the signal log before the Part 3 figures, under a standing rule set on 19 October for each later analysis or publication step (Amendments 50 and 51). Two separate instances calculated every Part 3 figure, each with its own code and without seeing the other's work, and a third recounted every figure (Amendment 52, Change 1). It also gave the help in both re-checks set out in section 4.9. On 21 October 2026 it suggested a wrong reason for one photo value, which the researcher accepted and withdrew the same day. Where a capture was too large for its tools, the researcher's reading is the only reading (Amendment 51). Before the first run, the assistant wrongly told the researcher that expanding the AI Overview was optional, which contributed to the AI Overview deviation in section 4.2 (Amendment 15, Change 2). Claude as a platform under study was queried only through its separate study account.

4.12 Changes after registration

The study recorded 54 post-registration amendments between 6 August and 28 October 2026, plus two entries of 10 August restored later (R9 and R10, Amendment 18). Each amendment records its trigger, the change and its timing relative to data collection. Section 8 summarises them. The full notes are on osf.io/c63dk, with firms replaced by study codes.

5. Findings

The disclaimer in section 2 applies to every finding in this section. A reference to "Section N", capitalised, means a section of methodology v2.4.

Part 1 of 3: Consistent appearance

All figures below come from mention-frequency workbook v6 (13 October 2026), built from the corrected run log. The later coded run log treats three further name spellings as the same firms as existing entries. On those merges, Table 2's 810 and 335 would each be about five lower. The one-run share stays at 41%, and the 137 consistent combinations are unchanged. No firm or individual is named. Counts of how many firms met the threshold, and of how often the platforms named firms, may be reported at any size, subject to the rule on counts that point to a single firm. Figures about firms' characteristics need at least 10 distinct firms (Amendments 41, 43 and 45).

5.1 What was measured

Nine queries were run: three practice areas (family law, personal injury, commercial litigation) in three locations (Brisbane, Gold Coast, Sunshine Coast). Each query was run 10 times on each of four platforms (Google AI Overviews, Perplexity, ChatGPT, Claude) across two waves: Wave 1 ran 10 to 18 August 2026 and Wave 2 ran 14 to 22 September 2026. That is 360 runs, plus one replacement run, for 361 recorded runs.

A firm counts as appearing in a run when the platform presented it as a provider to contact. Directories, platforms and barristers' chambers are not counted as firms. Workbook v6 lists 15 rows for them, 6 of which at least one platform named. These rows are excluded from every count below. A firm is a "consistent appearance" for a query on a platform when it appears in 7 or more of the 10 runs. This threshold was set before data collection. It is a working definition for an exploratory study, not a standard from the literature.

Four conditions qualify every figure in this section:

- **Google AI Overviews** are recorded as first displayed, without expanding the Overview, and counted from the main capture only, excluding faded text, under a recording rule set after collection (Amendment 29). The registered method said to expand the Overview, so AI Overview counts may be lower than the registered method would have produced.
- **Perplexity** was run signed in to a dedicated study account with memory inactive, in 87 of 90 runs. The registered condition was signed out. Earlier study threads were visible in the account at one run, so some carry-over of account context cannot be ruled out.
- **Two AI Overview cells had a run with no Overview.** Every query showed an AI Overview in 10 of 10 runs except two. Query 03 (family law, Sunshine Coast) showed one in 9 of 10 runs; no replacement was possible, so its threshold is 7 of 9. Query 02 (family law, Gold Coast) showed one in 10 of 11 runs: the missing run was replaced by a run made on 22 September 2026, 35 days after Wave 1 closed and on the last day of Wave 2, and counted in Wave 1. Figures with and without that replacement run are in section 5.6.
- **All runs** were made from one Sunshine Coast internet connection on one device, with location services off, which may pull results towards the Sunshine Coast.

5.2 Each platform names many firms, and about two in five are named only once

Counting each firm once per query and platform, the four platforms named firms 810 times across the nine queries:

Table 2. Firm and query entries, entries named once and consistent combinations, by platform.

PLATFORM	FIRMS NAMED AT LEAST ONCE (SUMMED OVER QUERIES)	NAMED IN ONLY 1 RUN	CONSISTENT (7 OR MORE OF 10)
Google AI Overviews (as first displayed)	107	38 (36%)	30
Perplexity (study account, signed in)	267	122 (46%)	38
ChatGPT	205	83 (40%)	36
Claude	231	92 (40%)	33
All four	**810**	**335 (41%)**	**137**

Across the four platforms, 22 firm, platform and query combinations were named in all 10 runs. This count covers every query and platform with 10 runs. Query 03's AI Overviews had 9 runs and are not included in it.

About two in five firms a platform named for a query appeared in one run and not again. Perplexity named the most firms and the highest share of one-off names. AI Overview counts come from the Overview as first displayed in the main capture, which can show fewer firms than the expanded Overview (Amendments 15 and 29).

5.3 Consistent appearance: 137 firm, platform and query combinations reached the threshold

137 firm, platform and query combinations reached 7 or more of 10 runs. They involve 74 firm-and-query rows and 63 distinct firms: some firms are consistent for more than one query.

At 10 runs, every count carries a wide interval. A firm named in 7 of 10 runs has a 95% interval of about 40% to 89%; a firm named in all 10 has an interval of about 72% to 100%. For Query 03 AI Overviews, 7 of 9 is about 45% to 94%. These are Wilson score intervals, approximate because runs may not be independent. The threshold is a cut-off for firms named often. It does not show that a platform will name a firm on any given search.

Consistent rows by query. Each cell counts the firm-and-query rows that met the threshold on at least one platform. The cells add to 74. A firm consistent for two queries appears in two cells.

Table 3. Consistent rows by query.

PRACTICE AREA	BRISBANE	GOLD COAST	SUNSHINE COAST
Family law	8	7	10
Personal injury	5	11	9
Commercial litigation	9	7	8

5.4 Just over half of consistent rows reach the threshold on one platform only

Of the 74 firm-and-query rows that reached the threshold:

Table 4. Consistent rows by number of platforms.

CONSISTENT ON	ROWS
1 platform	38
2 platforms	16
3 platforms	13
All 4 platforms	7

Just over half of the consistent rows (38 of 74) reach the threshold on one platform only. Seven reach it on all four. The platforms overlap in part and differ in part, which supports reporting each platform separately, as the registered method requires (Section 11).

5.5 The coast commercial litigation cells were not sparse, and one platform showed no stable set

The methodology expected the commercial litigation queries for the Gold Coast and Sunshine Coast to return very few firms. They did not. Both reached the threshold (7 and 8 firms, section 5.3), and Perplexity named 42 different firms for commercial litigation on the Gold Coast, the most of any query and platform. None of those 42 appeared in 7 or more runs. In these 10 runs, Perplexity did not return a stable set of firms for that query.

Counts for each query on each platform are kept in the private analysis record (Amendment 45).

5.6 The headline count under each sensitivity check

Table 5. Consistent combinations under each sensitivity check.

CHECK	CONSISTENT COMBINATIONS
Primary result	137
Query 03 AI Overviews counted out of 10 (the run with no Overview counted as a run with no firms)	137
Replacement run excluded (Query 02 AI Overviews out of 9)	136

Each check is shown for the whole study, as the number of consistent combinations. Without the replacement run, one combination in Query 02's AI Overviews falls below the threshold. This difference is reported so that the result can be read with and without that run. Their other values, and a check excluding the judgement items (ten firm counts from eight runs where the coding called for a judgement, such as a combined entry naming two firms or a partly obscured name), are kept in the private analysis record (Amendment 45).

The run log was corrected before this result was calculated. The blind re-check (Amendment 28) found that some AI Overview runs had been recorded with firms from outside the Overview itself: the map pack, sponsored listings, organic results and source cards. After the re-check, firm records in 8 runs across all four platforms were corrected on 11 October 2026. Every AI Overview run was then audited against its capture, and 57 AI Overview runs were corrected on 13 October 2026 (Amendments 29 and 32). Before those corrections the count was 148 combinations, 35 at 10 of 10. All of the change in consistent combinations is on AI Overviews (41 to 30). The consistent counts on the other three platforms are unchanged; six of their counts changed, none crossing the threshold.

Table 6. The 57 AI Overview corrections of 13 October 2026, by type.

TYPE OF CORRECTION	RUNS
Run set to "not triggered" (no Overview in the main capture)	1
Firms removed	54
Firms added	8

Six runs had firms both removed and added. Where firms were removed, the dated note names where they had come from (a run can name several):

Table 7. Where removed firms had come from.

SOURCE OF A REMOVED FIRM	RUNS NAMING IT
Organic results	28
The other capture of the same search	27
Map pack	19
Source cards or citation chips	14
Sponsored listings	7
Below "Show more" or faded	6 (3 stated, 3 implied by the note)
Not specified (not in the main capture's Overview)	4

The 8 runs with firms added had missed a firm shown inside the Overview: named as the practice of an individual (3), in a community recommendation (2), on a card at a page break (1), or with no reason given (2). The types were coded from the dated notes twice, separately, with the same counts; the only difference, one run, is reflected above.

5.7 Most consistent combinations were named in at least 3 runs of each wave

For the 137 consistent combinations, runs named in Wave 1 and Wave 2:

Table 8. Consistent combinations by runs named in each wave.

RUNS NAMED, WAVE 1 AND WAVE 2	COMBINATIONS
5 and 5	22
5 and 4, or 4 and 5	36
4 and 4	20
5 and 3, or 3 and 5	21
4 and 3, or 3 and 4	26
5 and 2, or 2 and 5	12

Each wave had 5 runs per query and platform, with two exceptions. Query 03 AI Overviews had 4 runs in Wave 2, so their wave counts are out of 5 and 4. Query 02 AI Overviews' Wave 1 count includes the replacement run made on 22 September 2026.

In 125 of 137 combinations the firm was named in at least 3 runs in each wave. In the other 12, the firm was named in every run of one wave and 2 of the other. No statistical test of the wave difference is made; the comparison is descriptive, as registered.

5.8 What these figures do and do not show

- They show how often each platform named firms for nine fixed queries, in two windows in August and September 2026, under the stated conditions.
- They do not show what any person sees: platforms personalise and change, and these runs were made from one connection and device.
- They do not rank firms or assess legal services. No firm is identified, and a firm's absence says nothing about its quality.
- Every association reported in Parts 2 and 3 is a correlation. No comparison in this study supports a causal claim.
- Firm names were transcribed from the saved captures with the help of an AI assistant from the same provider as one of the platforms studied (Claude). The researcher checked and entered the values, except the 57 AI Overview corrections of 13 October 2026, which the assistant entered from her recorded decisions; she checked a random 6 of them (Amendment 32).
- The same researcher re-coded 37 of the 361 runs, chosen at random, without seeing the original coding. This is a consistency check, not an independent review, and the sample is small. Raw positive specific agreement was 0.876 (95% interval 0.805 to 0.937), calculated before the corrections it prompted, against a reference point of 0.800 used here as an approximation. The researcher coded every run. The AI assistant explained the process for the first two runs, answered process questions, pointed her to the written rule on one coding question, and checked three columns for completeness before the lock. She decided the coding question. After the check she filled one empty count and removed an individual's name from one note. The assistant viewed no capture and did not advise on coding judgements. This departs from the method set for the re-check (Amendment 28, point 4; Amendment 52, Change 4).

Part 2 of 3: Signals A, D, E and F

All figures below come from the Appendix C signal log (columns A to S and AF to AK; no Signal B or C value was read) and the consistent group in workbook v6. No firm or individual is named.

5.9 How the groups are compared

Each consistent firm-and-query row (74 rows, 63 distinct firms) is compared with the comparison sample: 5 firms per query, drawn at random from the population listings that no platform named in that query's runs (45 rows, 45 distinct firms). Under Amendment 41, comparisons are pooled by practice area, with a whole-study figure as well, and each figure rests on at least 10 distinct firms with a recorded value for that figure (Amendment 41). A figure is also withheld where it is near-uniform: where 10% or fewer, or 90% or more, of a group share the value it reports. A near-uniform category is first combined with others, in an order fixed in advance, and a figure that cannot be combined is withheld (Amendment 43). Withheld figures read "not reported". The minimum and maximum of a group are not published, because each gives one firm's value. The four rows that left the consistent group after the 13 October corrections are not reported as a group.

These points apply to every figure in this part:

- **Rows and firms.** A firm consistent in more than one location of a practice area is one row per location. In personal injury, 25 consistent rows come from 17 firms, and such a firm carries the same Signal D value in each of its rows (Amendments 21 and 22). The number of distinct firms is given beside each figure.
- **Group balance.** Each comparison group has 5 rows per location; the consistent groups do not (for example, personal injury: 5 rows in Brisbane and 11 on the Gold Coast). Pooled figures weight the locations differently in the two groups.
- **Overlap.** Some firms that are consistent in one practice area and location were drawn as comparison firms in another, so they sit on both sides of the family law and whole-study figures (Amendments 13 and 41). A comparison firm can also be linked to a consistent firm through shared premises or a principal (Amendment 50). Both may narrow the differences on Signal D, and on Signal B in Part 3.
- **Population frame.** Comparison firms were drawn only from the population listings. 11 of the 74 consistent rows are firms with no listing in the population for that practice area and location.
- **Timing.** Signals D, E and F were observed between 28 September and 5 October 2026, after both waves of runs. Signal A comes from the organic checks recorded at each run.
- **Correlation only.** These are descriptive comparisons between small groups. No statistical test is made, and none of these differences shows that a signal causes a platform to name a firm. Established firms may differ from others in many ways at once.
- **Registered comparison kept.** The comparison was also run for each of the nine practice area and location combinations, as registered, and kept in the private analysis record (Amendment 41, Change 4). The number of combinations that point the same way as the pooled result is given with each signal.
- **Intervals.** Shares carry approximate 95% Wilson score intervals, the method Section 11 registers for mention frequencies, extended to these shares by Amendment 47. Each is given at the first mention of a headline figure, and all are listed in the appendix to this part. An interval describes one group's share. It is not an interval on the difference between the groups, and overlapping or separate intervals are not a test. Intervals assume each row is independent. Consistent rows from the same firm are not (personal injury: 25 rows from 17 firms), and some firms sit in both groups, so the intervals are approximate. A comparison group's interval describes that random sample.

5.10 Signal A: about half of consistent rows rank in Google's top 10

Signal A records whether a firm appears in Google's organic top 10 for the same query, pooled across the organic checks recorded for that query: 40 per query, 41 for Query 02 (majority rule, Amendment 12).

Table 9. Signal A: consistent rows in the organic top 10, by practice area.

PRACTICE AREA	CONSISTENT ROWS IN THE TOP 10	COMPARISON FIRMS IN THE TOP 10
Family law	13 of 25 (25 firms)	not reported
Personal injury	12 of 25 (17 firms)	not reported
Commercial litigation	14 of 24 (24 firms)	not reported
Whole study	**39 of 74 (63 firms)**	**not reported**

About half of the consistent rows rank in the organic top 10 (39 of 74, 53%; 95% interval 41% to 64%).

By platform, for rows consistent on that platform (a row can be consistent on more than one). The last row counts only the rows not consistent on AI Overviews:

Table 10. Signal A: consistent rows in the organic top 10, by platform.

PLATFORM	CONSISTENT ROWS IN THE TOP 10
Google AI Overviews	24 of 30 (28 firms)
Perplexity	20 of 38 (32 firms)
ChatGPT	20 of 36 (33 firms)
Claude	18 of 33 (27 firms)
Consistent only on Perplexity, ChatGPT or Claude (not on AI Overviews)	15 of 44 (39 firms)

Four in five rows consistent on AI Overviews rank in the organic top 10 (24 of 30, 80%; 95% interval 63% to 90%). About half do on each of the other platforms, and one in three of the rows consistent only on Perplexity, ChatGPT or Claude (15 of 44, 34%; 95% interval 22% to 49%). AI Overviews are part of Google Search, so their link to organic ranking is partly built in (the coupling caveat in Section 10 of the methodology). The 44 rows consistent only on Perplexity, ChatGPT or Claude are the rows least affected by that coupling.

Proposition 1 is addressed at firm level by Signal A (Amendment 46). The comparison firms' figures are not reported, so these figures show how many consistent rows rank in the organic top 10. They cannot show whether organic ranking and consistent AI appearance are correlated or independent. No run-level figure is reported (Amendment 46).

5.11 Signal D: consistent firms had published content more recently

Signal D is the date of the most recent dated post on the firm's own website, counted in days before the firm's observation date (28 September to 5 October 2026, after the runs). Firms with no content, content without dates, or content behind a personal injury advertising gate are counted together as "no dated post".

Median days since the last dated post:

Table 11. Signal D: median days since the last dated post.

PRACTICE AREA	CONSISTENT	COMPARISON
Family law	29 days (20 rows, 20 firms)	81 days (10 firms)
Personal injury	12 days (18 rows, 12 firms)	not reported
Commercial litigation	12 days (21 rows, 21 firms)	120 days (12 firms)
Whole study	**15 days (59 rows, 50 firms)**	**101 days (31 firms)**

Counting each consistent firm once, the whole-study median is 27 days (50 firms). The direction is the same in eight of the nine combinations.

Across the whole study, 46 of 59 consistent rows with a dated post had posted within 90 days (37 of 50 firms, 74%; 95% interval 60% to 84%), against 15 of 31 comparison firms (48%; 95% interval 32% to 65%). Signal D carries the same value in each of a firm's rows, so the interval is given for firms.

Rows without a dated post:

Table 12. Signal D: rows and firms without a dated post.

WHOLE STUDY	CONSISTENT (74 ROWS)	COMPARISON (45 FIRMS)
No dated post (no content published, content undated, or content behind a personal injury advertising gate)	15 (13 firms)	14

On the median, consistent firms had posted more recently. The spread is wide in both groups, so the median describes the middle of each group. Recency was measured at observation, after the runs, so a recent post may postdate a firm's appearances.

5.12 Signal E: commercial litigation consistent firms had a lower median review count

Signal E (Google review count) is an exploratory observation recorded for commercial litigation only. It is not recorded for personal injury, where the restrictions on personal injury advertising (Personal Injuries Proceedings Act 2002 (Qld), s 66) make review counts unreliable, or for family law, where client confidentiality is likely to depress review volumes (Section 9).

Table 13. Signal E: median Google review count, commercial litigation.

COMMERCIAL LITIGATION	MEDIAN REVIEWS
Consistent (24 rows, 24 firms)	34.5
Comparison (15 firms)	60

The consistent firms' median was lower, and the two ranges overlap almost fully. The direction is the same in two of the three commercial litigation combinations. This is one practice area with small groups.

5.13 Signal F: consistent firms were more often listed in the QLS directory

Signal F is an exploratory observation. It records presence in two directories: QLS Find a Solicitor (F1) and LawConnect (F2). It is scored 0 to 2, because FindLaw Australia, the third registered directory, ceased publication in December 2022 (Amendments 19, 27 and 39).

Table 14. Signal F: directory presence, QLS (F1) and LawConnect (F2).

PRACTICE AREA	QLS (F1): CONSISTENT	QLS (F1): COMPARISON	LAWCONNECT (F2): CONSISTENT	LAWCONNECT (F2): COMPARISON
Family law	20 of 25 (25 firms)	4 of 15 (15 firms)	13 of 25 (25 firms)	4 of 15 (15 firms)
Personal injury	15 of 25 (17 firms)	5 of 15 (15 firms)	12 of 25 (17 firms)	8 of 15 (15 firms)
Commercial litigation	21 of 24 (24 firms)	9 of 15 (15 firms)	12 of 24 (24 firms)	8 of 15 (15 firms)
Whole study	**56 of 74 (63 firms)**	**18 of 45 (45 firms)**	**37 of 74 (63 firms)**	**20 of 45 (45 firms)**

Three in four consistent rows were listed in the QLS directory (56 of 74, 76%; 95% interval 65% to 84%), against two in five comparison firms (18 of 45, 40%; 95% interval 27% to 55%); the direction is the same in all nine combinations. LawConnect showed little difference across the whole study (50%, 95% interval 39% to 61%, against 44%, 95% interval 31% to 59%), and its direction is the same in five of the nine combinations. Qualifications:

- QLS presence was recorded from the directory's Law Firm tab, a firm-level search. The registered rule records a firm as present if any of its solicitors appears for the location. This deviation may under-count presence (Amendment 37, Change 7).

5.14 What Part 2 shows

Against randomly drawn firms that the platforms did not name, the firms the platforms named consistently stood out on two signals. They had posted more recently on their own websites. They were more often listed in the QLS directory. In commercial litigation, their median Google review count was lower. About half of the consistent rows ranked in Google's organic top 10. Each comparison is a correlation between small, unevenly balanced groups. Part 3 adds structured data (Signal B) and Google Business Profile completeness (Signal C) after the blind re-check on 19 October 2026.

Appendix to Part 2: intervals for each share

Approximate 95% Wilson score intervals (Amendment 47), for every share published in sections 5.10, 5.11 and 5.13, in the unit each figure is published in. Consistent groups count rows unless stated; comparison groups count firms. The qualifications in section 5.9 apply. No interval is given for medians, for counts of combinations pointing the same way, or for any figure that is not reported.

Table 15. Intervals for each share published in Part 2.

SECTION	FIGURE	SHARE	95% INTERVAL
5.10	Signal A, family law, consistent rows	13 of 25 (52%)	33% to 70%
5.10	Signal A, personal injury, consistent rows	12 of 25 (48%)	30% to 67%
5.10	Signal A, commercial litigation, consistent rows	14 of 24 (58%)	39% to 76%
5.10	Signal A, whole study, consistent rows	39 of 74 (53%)	41% to 64%
5.10	Signal A, AI Overviews rows	24 of 30 (80%)	63% to 90%
5.10	Signal A, Perplexity rows	20 of 38 (53%)	37% to 68%
5.10	Signal A, ChatGPT rows	20 of 36 (56%)	40% to 70%

5.10	Signal A, Claude rows	18 of 33 (55%)	38% to 70%
5.10	Signal A, consistent only on Perplexity, ChatGPT or Claude	15 of 44 (34%)	22% to 49%
5.11	Signal D, posted within 90 days, consistent rows	46 of 59 (78%)	66% to 87%
5.11	Signal D, posted within 90 days, consistent firms	37 of 50 (74%)	60% to 84%
5.11	Signal D, posted within 90 days, comparison firms	15 of 31 (48%)	32% to 65%
5.11	Signal D, no dated post, consistent rows	15 of 74 (20%)	13% to 31%
5.11	Signal D, no dated post, consistent firms	13 of 63 (21%)	12% to 32%
5.11	Signal D, no dated post, comparison firms	14 of 45 (31%)	20% to 46%
5.13	QLS F1, family law, consistent rows	20 of 25 (80%)	61% to 91%
5.13	QLS F1, personal injury, consistent rows	15 of 25 (60%)	41% to 77%
5.13	QLS F1, commercial litigation, consistent rows	21 of 24 (88%)	69% to 96%
5.13	QLS F1, whole study, consistent rows	56 of 74 (76%)	65% to 84%
5.13	QLS F1, family law, comparison firms	4 of 15 (27%)	11% to 52%
5.13	QLS F1, personal injury, comparison firms	5 of 15 (33%)	15% to 58%
5.13	QLS F1, commercial litigation, comparison firms	9 of 15 (60%)	36% to 80%
5.13	QLS F1, whole study, comparison firms	18 of 45 (40%)	27% to 55%
5.13	LawConnect F2, family law, consistent rows	13 of 25 (52%)	33% to 70%
5.13	LawConnect F2, personal injury, consistent rows	12 of 25 (48%)	30% to 67%
5.13	LawConnect F2, commercial litigation, consistent rows	12 of 24 (50%)	31% to 69%
5.13	LawConnect F2, whole study, consistent rows	37 of 74 (50%)	39% to 61%
5.13	LawConnect F2, family law, comparison firms	4 of 15 (27%)	11% to 52%
5.13	LawConnect F2, personal injury, comparison firms	8 of 15 (53%)	30% to 75%
5.13	LawConnect F2, commercial litigation, comparison firms	8 of 15 (53%)	30% to 75%
5.13	LawConnect F2, whole study, comparison firms	20 of 45 (44%)	31% to 59%

Part 3 of 3: Signals B and C

Part 3 compares the two groups on website structured data (Signal B) and Google Business Profile completeness (Signal C). Both signals passed the blind re-check of 19 October 2026 as reliable (section 5.15), and each Signal B or C figure carries that result beside it.

The figures in sections 5.16 to 5.18 come from the Appendix C signal log and the consistent group in workbook v6. The agreement figures in section 5.15 come from the locked re-code and the dated export of the signal log. No firm or individual is named.

5.15 How the groups are compared, and how consistent the Signal B and C scoring was

The groups, the pooling and the minimum group size are those of Part 2 (section 5.9). Each consistent firm-and-query row (74 rows, 63 distinct firms) is compared with the comparison sample (45 rows, 45 distinct firms). Comparisons are pooled by practice area, with a whole-study figure as well, and each figure rests on at least 10 distinct firms with a recorded value for that figure (Amendment 41). Near-uniform figures are combined or withheld as in Part 2 (Amendment 43), and withheld figures read "not reported". The qualifiers in section 5.9 apply to every figure in this part: rows and firms, group balance, overlap, population frame, correlation only, and the registered comparison kept in the private analysis record.

Two further points apply to Signals B and C:

- **Timing.** Signals B and C were observed between 28 September and 5 October 2026, after both waves of runs.
- **Carry-over.** A firm consistent in more than one location of the same practice area carries the same Signal B value in each of those rows. A firm in two practice areas was observed separately for Signal B in each. Signal C was observed separately for each row's own listing (Amendment 22, Change 2).

The re-check. 13 of the 123 firm records, drawn at random, were re-scored on 19 October 2026 by the researcher, from the saved evidence only, without seeing the original scores and 14 days after the last observation (Amendment 30). At the researcher's request, the AI assistant gave procedural help during the re-score. For the first record, it said which file to open, which column to fill and which rule to read. When the researcher mentioned a schema type seen in that record, it referred her to two rules without saying whether the type counted; this is disclosed as arguably advice on a scoring judgement. It read out dates for two records on request. Before the lock, it read the sheet to

check for blank cells, note placement and names, and commented on no value. Before the lock it viewed no evidence file and gave no value for any record, and both records it helped with agreed with the original on every value. After the lock, it read the evidence files for the one record with disagreements and confirmed the researcher's reading. These steps depart from the registered method and are disclosed as deviations (Amendment 48, Changes 2 and 6). This is a consistency check by the same person, not an independent review, and the sample is small. Each signal receives one result from its Krippendorff's alpha for nominal data, applied to the unrounded point estimate: Reliable (0.800 or above), Tentative (0.667 to below 0.800), Unreliable (below 0.667), or Not banded (where alpha cannot be calculated, or one value makes up 90% or more of the values in either coding, the test being applied to each coding separately). The two signals are decided separately (Amendment 40).

Table 16. Signal B and C re-check: agreement results.

MEASURE	SIGNAL B (STRUCTURED DATA)	SIGNAL C (GOOGLE BUSINESS PROFILE)
Agreement	13 of 13 records (95% Wilson interval 77% to 100%)	62 of 65 element judgements (95% cluster bootstrap interval, resampling whole records, 86% to 100%)
Krippendorff's alpha	1.000	0.905
Excluding records whose value is fixed by the evidence	12 of 12, alpha 1.000	60 of 60, alpha 1.000
Result	Reliable	Reliable

For Signal B, the records excluded in the third row are those coded Not tested or no website in either coding. For Signal C, they are records with no listing in either coding. The bands are not applied to these figures (Amendment 30, point 10).

Signal C, agreement by element: C1 12 of 13, C2 12 of 13, C3 12 of 13, C4 13 of 13, C5 13 of 13. Exact match of the C total: 12 of 13 (95% Wilson interval 67% to 99%).

The figures above are the raw figures, calculated on the locked re-code and the dated export before any disagreement was examined. No original value was corrected as a clear error, and no disagreement was a judgement call. All three Signal C disagreements were in one record: the re-score had read the business panel of a different firm, and the original stands. This is counted as one re-code error (Amendment 48, Change 6). The agreement figures are not recalculated (Amendment 30, point 13, as changed by Amendment 40, Change 3(d)).

5.16 Signal B: structured data was detected more often for consistent rows

Signal B records whether a firm's website exposes structured data of three types: FAQPage, LocalBusiness or LegalService, and Person. It is tested with the Schema.org validator on three fixed pages: the homepage, the page for the row's practice area, and the first lawyer profile (Amendment 16). A row is "Detected valid" when at least one of these types is found without errors on at least one page, "Detected with errors" when the types are found only with errors, and "Not detected" when none is found. A firm with no website is recorded as not detected (no website), and counts as not detected in the comparison (Amendment 8). Rows where no page could be tested, because the whole site sits behind a personal injury advertising gate or every page exceeds the validator's size limit, are counted as not detected in every Signal B figure below (Amendment 43, Change 2(c)).

Re-check of this signal (section 5.15): 13 of 13 records agreed, alpha 1.000, Reliable.

Rows with structured data detected (valid or with errors), re-check: reliable, alpha 1.000:

Table 17. Signal B: structured data detected, consistent rows and comparison firms, by practice area.

PRACTICE AREA	CONSISTENT ROWS	COMPARISON FIRMS
Family law	17 of 25 (25 firms)	6 of 15
Personal injury	16 of 25 (17 firms)	5 of 15
Commercial litigation	13 of 24 (24 firms)	3 of 15
Whole study	**46 of 74 (63 firms)**	**14 of 45**

The direction is the same in eight of the nine combinations (re-check: reliable, alpha 1.000).

Whole study, by classification (re-check: reliable, alpha 1.000). The categories are combined under Amendment 43, Change 1(b):

Table 18. Signal B: whole study, by classification.

CLASSIFICATION	CONSISTENT (74 ROWS)	COMPARISON (45 FIRMS)
Detected	46	14
Not detected	28	31

The registered method reports Signal B separately for Google AI Overviews, where it describes structured data as "a confirmed Google

ecosystem signal", and for the other three platforms (Section 9). The two groups of rows below do not overlap. Each group has at least 10 distinct firms, and neither can be combined with the whole-study figure to reveal a group of fewer than 10 (re-check: reliable, alpha 1.000):

Table 19. Signal B: structured data detected, by platform group.

GROUP	ROWS WITH STRUCTURED DATA DETECTED
Rows consistent on Google AI Overviews (on that platform and possibly others)	17 of 30 (28 firms)
Rows consistent only on Perplexity, ChatGPT or Claude	29 of 44 (39 firms)
Comparison firms (not named for that query)	14 of 45

Google's Rich Results Test, recorded separately as a readout and not re-checked (Amendment 30, point 4): not reported.

Structured data was detected for 62% of consistent rows (95% interval 51% to 72%) against 31% of comparison firms (95% interval 20% to 46%) (re-check: reliable, alpha 1.000). The consistent share was higher in each practice area: family law 68% against 40%, personal injury 64% against 33%, and commercial litigation 54% against 20% (re-check: reliable, alpha 1.000). Both platform groups were above the comparison firms: 57% for rows consistent on Google AI Overviews and 66% for rows consistent only on the other three platforms, with intervals that overlap widely (re-check: reliable, alpha 1.000).

Qualifications:

- Three pages are a fixed sample of each website, not a full crawl. Structured data on other pages, such as a separate FAQ page, is not captured (Amendment 16).
- Rows that could not be tested, because of a personal injury advertising gate or the validator's size limit, are counted as not detected. This may affect the two groups unevenly.
- Structured data was observed after the runs and may have changed since a firm appeared.

5.17 Signal C: listings scoring 4 or 5 of 5 were more common among consistent rows

Signal C scores the firm's Google Business Profile listing from 0 to 5, one point for each element present at observation:

- C1: a business description
- C2: a primary category
- C3: a services section with at least one service
- C4: at least one photo uploaded by the firm showing its own staff, office or signage. Logo-only, stock and Google-generated images do not count (Amendment 21).
- C5: a post within the 90 days before observation

A firm without physical premises scores 0 on C4. The 0 counts in its total, and the firm is left out of photo comparisons (Section 9). A firm with no Google Business Profile listing scores 0 on every element (Amendment 26). Comparison firms were drawn from listings in the population.

Re-check of this signal (section 5.15): 62 of 65 element judgements agreed, alpha 0.905, Reliable.

Rows scoring 4 or 5, re-check: reliable, alpha 0.905:

Table 20. Signal C: scoring 4 or 5, consistent rows and comparison firms, by practice area.

PRACTICE AREA	CONSISTENT ROWS	COMPARISON FIRMS
Family law	14 of 25 (25 firms)	5 of 15
Personal injury	8 of 25 (17 firms)	2 of 15
Commercial litigation	10 of 24 (24 firms)	4 of 15
Whole study	**32 of 74 (63 firms)**	**11 of 45**

The direction is the same in seven of the nine combinations (re-check: reliable, alpha 0.905).

Whole study, score distribution (re-check: reliable, alpha 0.905). The scores are combined into the registered split under Amendment 43, Change 1(b):

Table 21. Signal C: whole study, score distribution.

SCORE	CONSISTENT (74 ROWS)	COMPARISON (45 FIRMS)
4 or 5	32	11
0 to 3	42	34

Whole study, by element (re-check: reliable, alpha 0.905):

Table 22. Signal C: whole study, by element.

ELEMENT	CONSISTENT	COMPARISON
C1 Business description	not reported	not reported
C2 Primary category	not reported	not reported
C3 Services listed	21 of 74 (63 firms)	9 of 45
C4 Own photos (firms without premises left out)	60 of 74 (63 firms)	26 of 45
C5 Post within 90 days	21 of 74 (63 firms)	8 of 45

By platform, rows scoring 4 or 5 were 14 of 30 (28 firms) for rows consistent on Google AI Overviews and 18 of 44 (39 firms) for rows consistent only on Perplexity, ChatGPT or Claude, against 11 of 45 comparison firms (re-check: reliable, alpha 0.905). The two platform groups gave similar Signal C figures: both are above the comparison firms, and their intervals overlap widely.

Listings scoring 4 or 5 were more common among consistent rows (43%, 95% interval 33% to 55%) than among comparison firms (24%, 95% interval 14% to 39%) (re-check: reliable, alpha 0.905). The consistent share was higher in each practice area: family law 56% against 33%, personal injury 32% against 13%, and commercial litigation 42% against 27% (re-check: reliable, alpha 0.905). Of the elements reported, the widest gap was in the firm's own photos: 81% of consistent rows against 58% of comparison firms (re-check: reliable, alpha 0.905).

Proposition 3 links consistent appearance with more complete listings and more recently published website content. Read with Signal D (section 5.11), the two signals point the same way: consistent firms had published website content more recently and had more complete listings. This combined statement takes the lower of the results involved, here Signal C's: reliable (alpha 0.905). Signal D was not re-checked.

Qualifications:

- C1 was read from the Google Search business panel, or from a second, more specific search where the location search showed only a multi-location panel (Amendment 19, Change 4(a), and Amendment 25, Change 4). C5 was read from the listing's posts on Maps and in the Search panel, with relative dates accepted where clearly inside or outside the window (Amendment 19, Change 4(c), and Amendment 20, Change 5).
- C4 requires a judgement on whether an image shows the firm's own staff, office or signage, read from the image alone (Amendment 21).
- Listings were observed after the runs and may have changed since a firm appeared.

5.18 Google Business Profile post recency, Google AI Overviews only: a recent post was more common among consistent rows

Post recency is an exploratory observation, registered for Google AI Overviews only and not framed as a proposition (Section 10). It uses element C5: a post on the firm's listing within the 90 days before observation. It takes Signal C's re-check result.

Table 23. Google Business Profile post within 90 days, Google AI Overviews only.

GROUP (RE-CHECK: RELIABLE, ALPHA 0.905)	POST WITHIN 90 DAYS
Rows consistent on Google AI Overviews	10 of 30 (28 firms)
Comparison firms (not named for that query)	8 of 45

A recent post was on the listing for 33% of rows consistent on Google AI Overviews (95% interval 19% to 51%), against 18% of comparison firms (95% interval 9% to 31%) (re-check: reliable, alpha 0.905). The two intervals overlap, and the groups are small.

5.19 What Part 3 shows

Against randomly drawn firms that the platforms did not name, the consistent rows more often had structured data detected on their websites (62% against 31%; re-check: reliable, alpha 1.000) and more often had a Google Business Profile listing scoring 4 or 5 of 5 (43% against 24%; re-check: reliable, alpha 0.905). Both patterns held in each practice area, and for rows consistent on Google AI Overviews as well as rows consistent only on the other three platforms. Rows consistent on Google AI Overviews more often had a recent Google Business Profile post (33% against 18%; re-check: reliable, alpha 0.905), an exploratory observation. Both signals passed the re-check as reliable (Signal B alpha 1.000; Signal C alpha 0.905). Each comparison is a correlation between small, unevenly balanced groups, observed after the runs.

Appendix to Part 3: intervals for each share

Approximate 95% Wilson score intervals (Amendment 47), for every share published in sections 5.16 to 5.18, in the unit each figure is published in. Consistent groups count rows; comparison groups count firms. The qualifications in sections 5.9 and 5.15 apply. No interval is given for counts of combinations pointing the same way, or for any figure that is not reported.

Table 24. Intervals for each share published in Part 3.

SECTION	FIGURE	SHARE	95% INTERVAL	RE-CHECK

5.16	Signal B detected, family law, consistent rows	17 of 25 (68%)	48% to 83%	reliable, alpha 1.000
5.16	Signal B detected, personal injury, consistent rows	16 of 25 (64%)	45% to 80%	reliable, alpha 1.000
5.16	Signal B detected, commercial litigation, consistent rows	13 of 24 (54%)	35% to 72%	reliable, alpha 1.000
5.16	Signal B detected, whole study, consistent rows	46 of 74 (62%)	51% to 72%	reliable, alpha 1.000
5.16	Signal B detected, family law, comparison firms	6 of 15 (40%)	20% to 64%	reliable, alpha 1.000
5.16	Signal B detected, personal injury, comparison firms	5 of 15 (33%)	15% to 58%	reliable, alpha 1.000
5.16	Signal B detected, commercial litigation, comparison firms	3 of 15 (20%)	7% to 45%	reliable, alpha 1.000
5.16	Signal B detected, whole study, comparison firms	14 of 45 (31%)	20% to 46%	reliable, alpha 1.000
5.16	Signal B detected, consistent on Google AI Overviews	17 of 30 (57%)	39% to 73%	reliable, alpha 1.000
5.16	Signal B detected, consistent only on Perplexity, ChatGPT or Claude	29 of 44 (66%)	51% to 78%	reliable, alpha 1.000
5.17	Signal C score 4 or 5, family law, consistent rows	14 of 25 (56%)	37% to 73%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, personal injury, consistent rows	8 of 25 (32%)	17% to 52%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, commercial litigation, consistent rows	10 of 24 (42%)	24% to 61%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, whole study, consistent rows	32 of 74 (43%)	33% to 55%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, family law, comparison firms	5 of 15 (33%)	15% to 58%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, personal injury, comparison firms	2 of 15 (13%)	4% to 38%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, commercial litigation, comparison firms	4 of 15 (27%)	11% to 52%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, whole study, comparison firms	11 of 45 (24%)	14% to 39%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, consistent on Google AI Overviews	14 of 30 (47%)	30% to 64%	reliable, alpha 0.905
5.17	Signal C score 4 or 5, consistent only on Perplexity, ChatGPT or Claude	18 of 44 (41%)	28% to 56%	reliable, alpha 0.905
5.17	C3 services listed, consistent rows	21 of 74 (28%)	19% to 40%	reliable, alpha 0.905
5.17	C3 services listed, comparison firms	9 of 45 (20%)	11% to 34%	reliable, alpha 0.905
5.17	C4 own photos, consistent rows	60 of 74 (81%)	71% to 88%	reliable, alpha 0.905
5.17	C4 own photos, comparison firms	26 of 45 (58%)	43% to 71%	reliable, alpha 0.905
5.17	C5 post within 90 days, consistent rows	21 of 74 (28%)	19% to 40%	reliable, alpha 0.905
5.17	C5 post within 90 days, comparison firms	8 of 45 (18%)	9% to 31%	reliable, alpha 0.905
5.18	Post within 90 days, consistent on Google AI Overviews	10 of 30 (33%)	19% to 51%	reliable, alpha 0.905
5.18	Post within 90 days, comparison firms	8 of 45 (18%)	9% to 31%	reliable, alpha 0.905

6. Limitations

Each limitation is stated once here, with the record that sets it out. Section 6.1 lists limits built into the registered design. Section 6.2 lists limits that arose in carrying the study out. Section 6.3 lists results the study registered and does not publish.

6.1 Limits of the registered design

Ten runs per query per platform is a small sample. It permits confidence intervals, and those intervals are wide. Larger studies used 60 to 100 runs per prompt (Fishkin and O'Donnell, SparkToro, January 2026). The findings are directional observations with quantified uncertainty (methodology v2.4, Section 12, limitation 1).

The confidence intervals are approximate. Wilson intervals assume each run is independent. Runs of one query on one platform may share platform state, so the stated 95% coverage is an approximation (methodology v2.4, Section 11).

Two waves give a short snapshot. Repeated-measurement studies show AI citation behaviour drifting over windows of a few weeks and longer (Schulte et al., arXiv 2604.07585; Sielinski, arXiv 2603.08924). The results describe August and September 2026 and may differ by the time of reading (limitation 2).

Every association is a correlation. Firms that AI platforms name consistently are likely to be established practices that score well on many signals at once. The study cannot attribute any difference to any single signal. That needs a controlled intervention study (limitation 5).

The comparison groups are small. The registered design compares the consistent firms in each practice area and location with five comparison firms. That supports a directional comparison only, and no statistical test is made (limitation 4). The two groups are also the two ends of the range: consistent firms were named in at least 7 of 10 runs on a platform, and comparison firms in none of that query's runs. Firms named in 1 to 6 runs on every platform are in neither group, and the differences may be wider than they would be across all firms.

The platforms' selection methods are not public. The study observes which firms the platforms named. The reasons for each selection are not observable (limitation 7).

ChatGPT and Claude mix training data with live retrieval. Their providers do not disclose the balance, so current website and listing signals may count for less in parametric-heavy responses (limitation 3).

Signals describe each firm as observed, and training data may be older. A platform's training data may predate a firm's current website and listing, so an association may reflect an earlier state of those signals (limitation 11).

The platforms may favour content from outside Australia. ChatGPT, Perplexity and Claude are developed in the United States and trained mainly on English-language content (limitation 8).

One query form was used. Each query asks for the "best" lawyers in a location. People looking for a personal injury lawyer often search with urgency and with proximity in mind, so the personal injury findings need extra caution (limitation 9; methodology v2.4, Section 5).

AI Overviews are part of Google Search. Their link to Google's organic rankings is partly built in, so the AI Overview arm of the Signal A comparison is coupled by construction. ChatGPT, Claude and Perplexity give the uncoupled test (methodology v2.4, Sections 4 and 10).

The population is defined by Google Business Profile listings. Firms without a listing, or with a listing that shows no street address, are outside it (limitation 10; Amendment 6). Practice-area and entity-type filters rest on the categories and names firms give themselves, and local government areas were assigned from listing addresses (Amendments 1 to 5 and 7). Government-funded and community legal services are kept in the population (Amendment 4). The population is a count of 805 listings, and a firm with more than one qualifying listing is counted for each (Amendments 7 and 37).

The study covers three locations. Ipswich was removed from the original design to fund more runs, so the regional contrast rests on the Sunshine Coast alone (limitation 12).

Every run came from one Sunshine Coast connection and device. Location services were off and each query names its location. Platforms may still infer location from the connection, which may pull results towards the Sunshine Coast (limitation 13).

The coast commercial litigation queries have no measurable Google search volume. The study kept "commercial litigation lawyers" in all three locations to hold the practice area constant (limitation 14). The methodology expected those two cells to be sparse. In the results they were not: both reached the threshold (section 5.5).

One researcher, with a commercial interest, collected and coded all the data. Pre-registration and the fixed run conditions reduce observer effects. They cannot remove them (limitation 6). Section 9 sets out the conflict of interest.

6.2 Limits that arose in carrying out the study

Google AI Overviews were recorded as first displayed. The registered method expanded each Overview before recording. AI Overview counts may therefore be lower than the registered method would have produced, and comparisons between AI Overviews and the other platforms carry this condition (Amendment 15, Change 2). Whether the practice was applied in every run cannot be confirmed from the captures (Amendment 37).

The AI Overview counting rule was set after collection. Some AI Overview runs had recorded firms from outside the Overview. A rule set on 11 October 2026 counts firms from inside the Overview in the main capture only, excluding faded text, and the researcher coded all 91 AI Overview runs again under it without opening the run log (Amendment 29). Limits of that audit:

- It was not blind. The researcher had seen the run log and the overall results.
- Two captures of the same search, minutes apart, sometimes showed different Overviews. Only the main capture counts.
- The AI assistant entered the resulting 57 corrections, and the researcher checked a random 6 of them (Amendment 32).
- Two steps departed from the audit rule: one run was treated as 1 firm as an exception, and named-individual counts changed in 34 of the 57 runs (Amendment 32, Change 1; Amendment 37, Change 4).
- The other three platforms were not audited in the same way. Their recording accuracy rests on the 37-run re-check (Amendment 29).

Some decisions were made after results were seen.

- The name-merging rule was set after the first result had been calculated, and moved it from 125 to 148 consistent combinations, before the AI Overview audit (Amendment 12).
- How a firm named in every run of a 9-run cell is counted was settled after the per-firm counts had been calculated and the headline figures seen (Amendment 34).
- The reporting rules in section 4.10 were set after the Part 1 and Part 2 figures had been calculated and seen (Amendments 43 and 45).
- The Rich Results Test rule for a record whose structured data could not be tested was decided with one such record's value known (Amendment 50, Change 3(b)).

- The fixed rule for what counts as a recorded source in the signal log was set after three checks had judged it differently, and kept after a fourth questioned two of its sources (Amendment 51, Change 5).
- The reading that the published medians do not point to a single firm was made with the medians known (Amendment 51, Change 9(e)).
- The photo element (C4) is published for both groups, departing from a reporting rule for this case, decided with the figures known (Amendment 52, Change 3).
- After the Signal B and C comparison, a third category, re-check error, was added for a disagreement where the original is shown correct, and two elements with no photo or post capture were not counted as unsupported (Amendment 48, Change 6(a) and (b)).
- Google's Rich Results Test was withheld for both groups after the Part 3 figures were calculated (Amendment 52, Change 2(c)).
- The second exception to the single-firm rule, for the adviser's firm's counts, was set on 22 October 2026 with all figures known (Amendment 52, Change 6).
- "Alone" was removed from the re-check wording fixed before the Signal B and C re-check, after both re-check results were known (Amendment 53).

Perplexity was run signed in for 87 of its 90 runs. The registered condition was signed out. A signed-in state may affect responses even with memory inactive (Amendment 15, Change 1; R9). At one run, earlier study threads were visible in the account's sidebar (R10; Amendment 37).

Two AI Overview cells depart from 10 runs. Query 02's replacement run was made 35 days after Wave 1 closed. The headline count is reported with and without it, and other figures without it are kept in the private record (Amendment 15, Change 4; Amendment 45). Query 03 has 9 AI Overview runs (Amendment 33).

An AI assistant transcribed the data. The registered methodology said the assistant had no role in data collection. In practice it drafted each log entry from the saved captures, and the researcher checked and entered every entry, apart from the corrections set out in section 4.11 (Amendment 15, Change 3). The assistant is made by the same provider as one of the platforms studied (Claude). Instances of the assistant also ran the record audit of 13 October and built the inventory of practitioners' names (Amendments 37 and 38). Its other roles, including the signal log corrections and the Part 3 calculations, are set out in section 4.11. The mention re-check was made by the researcher as a human check on the transcription, with the AI help set out in section 4.9 (Amendment 28; Amendment 52, Change 4).

Website and listing signals were observed once, after both waves. A firm's website or listing may have changed between the runs and the observation, and change between waves cannot be observed (Amendment 10).

Structured data was tested on three pages per website. Markup elsewhere on a site is not captured. Some sites could not be tested at all, because of a personal injury advertising gate or the validator's size limit, and are counted as not detected in the structured data figures (Amendments 16, 22 and 25; Amendment 52, Change 2(a)). For one record, structured data describes the homepage only, because its other pages sat behind an advertising gate (Amendment 50, Change 6(a)).

The directory signal is scored out of 2. FindLaw Australia ceased publication in December 2022, before registration (Amendments 19, 27 and 39). The QLS directory was searched by firm. The registered rule counts a firm as present if any of its solicitors appears for the location, so the search used may under-count presence (Amendment 37, Change 7).

Capture rules were set as cases arose. Each rule was applied to every later row and recorded within two days (Amendments 19 to 26). One directory value rests on a single search, because the repeat search the rules require was not captured (Amendment 51).

Signal log values were corrected after the re-check lock. Checks of the signal log against its own notes and saved captures found values that did not match. Five values were corrected on 19 October 2026 (Amendment 49), three more that day with one Rich Results Test value recoded (Amendment 50), and twelve on 21 October 2026 (Amendment 51). Part 2 figures changed with them, and each change is recorded in Amendments 49 to 51. Some published directory figures rested on values that did not match their notes or captures until 19 or 21 October (Amendments 50 and 51). Every analysed value now has a recorded source under a fixed rule (Amendment 51, Changes 5 and 6). Values supported only by the re-check, or by the photo re-check list of 30 September, were not checked again at source (Amendment 51). A value that matches its capture but was wrong on the day cannot be detected. None of the Signal B, Signal C or Rich Results Test values changed after the re-check lock (Amendments 50 and 51) is in the re-check sample, and the re-check agreement describes the sampled records only.

A change set was made to the signal log without a record. On the evening of 19 October 2026, 15 cells were changed with no record of the session. The changes were found on 21 October by comparing the live log with its dated copies, and each was checked at source and decided by the researcher. The changed notes cited an amendment that did not yet exist. Every changed value was kept, apart from one record whose note was wrong and whose value was corrected (Amendment 51, Changes 1 and 4(a)).

Signal D was observed again where the rule required it to be carried over. For three commercial litigation records on 4 October 2026, content recency was observed again, where the rule carried it over from the same firm's family law record. Two values were corrected to the carried-over date. The commercial litigation consistent median of 12 days depends on counting from the family law observation date; counted from the records' own date it would be 18 days (Amendment 49, Changes 4 and 7(a)).

Some published figures ran ahead of the signal log. For six days in October 2026, published directory figures counted values that the signal log did not yet hold. The figures themselves are correct (Amendment 49, Change 7).

Exclusion rules depend on how a platform framed its answer. Out-of-area firms and barristers' briefing suggestions are not counted where the platform framed them that way (Amendment 11). The barrister rule was not applied consistently: barristers or chambers were recorded

as firms in five runs (Amendment 37, Change 2(b)).

For multi-office firms, one listing was scored. Signal C for those firms describes one office (Amendment 14).

Comparisons are pooled by practice area. The registered comparison is within each practice area and location. Pooling, adopted on legal advice about identifiability, loses that location matching (Amendment 41). In the pooled groups:

- The comparison groups have 5 rows from each location and the consistent groups do not, so the two groups weight the locations differently.
- A firm consistent in several locations adds the same website values more than once (personal injury: 25 rows from 17 firms).
- Some firms sit on both sides: consistent in one location, drawn as a comparison firm in another. A comparison firm can also be linked to a consistent firm through shared premises or a principal. Both may narrow the differences (Amendments 13, 41 and 50).
- 11 of the 74 consistent rows are firms with no listing in the population for that practice area and location. Comparison firms were drawn only from population listings. All 74 stay in the consistent group, as the registered methodology provides, and a check without the 11 is kept privately (methodology v2.4, Sections 8 and 11; section 5.9; Amendment 51, Change 7).

Withheld figures hide some contrasts. Near-uniform figures are combined or not reported, and "not reported" can itself suggest that a group was near-uniform. The size and direction of some differences, including the comparison side of Signal A, cannot be read from the published figures (Amendment 43). The legal adviser confirmed on 14 October 2026 that these rules apply to the comparison group, and the comparison side of Signal A stays not reported (Amendment 45, Change 2(a)). In Part 3, structured data records that could not be tested are counted as not detected, and the structured data classifications are combined into detected against not detected. The Rich Results Test is not reported, and the GBP completeness score is given only as 4 or 5 against 0 to 3. The business description and primary category elements, the count of records with no listing, and distinct-firm counts within the platform groups are not reported (Amendment 52, Change 2).

The consistency re-checks are by the same researcher. They measure one person's consistency over time. The researcher had worked with the overall mention results for weeks before the mention re-check, which may raise agreement, and its sample of 37 runs gives wide intervals. Its two-week gap was measured from the end of AI response collection; measured from the end of all data collection, it started early (Amendment 28). The Signal B and C re-check was made on 19 October 2026, 14 days after the last observation, and its 13 records give wide intervals (Amendments 30 and 48). Google Drive showed two signal log files as viewed before its lock; the researcher's account is that she opened neither (Amendment 48, Change 3). The AI assistant helped in both re-checks, as set out in section 4.9. That departs from the method set for the re-checks (Amendment 48, Change 2; Amendment 52, Change 4; Amendment 53). On 11 October 2026 the researcher told the legal adviser that the mention re-check had no AI help. The advice of 12 October on the disclaimer wording rested on that statement, and the help was recorded 14 days after the re-check (Amendment 52, Change 4). The researcher sent the adviser a correction on 23 October 2026.

Practitioners' names were recorded in notes, against the registered rule. They have been removed from the working records, including 33 occurrences of names in 16 rows of the signal log, redacted on 19 October 2026. Older copies and the saved captures still show names, and named originals are held privately for the retention period. No count changed (Amendments 15, 38, 44 and 49).

The study's legal adviser's firm appears in the results. Section 9 sets out the disclosure. Its counts for that firm are published as a second exception to the single-firm rule. The firm is protected only while it cannot be identified, which depends on the redaction of the public notes being complete. The adviser declined to advise on the wording of the disclosure, independent advice was not obtained, and no lawyer reviewed it (Amendments 9, 35 and 36; Amendment 52, Changes 5 and 6).

6.3 Registered results that are not published

Results at the level of each query on each platform. The registered analysis reports frequencies for each firm, query and platform (methodology v2.4, Section 11). Under the reporting rules, those breakdowns are kept in the private analysis record, with the per-platform 10 of 10 counts, the highest count among firms below the threshold, and most sensitivity values (Amendment 45).

Two earlier commitments are changed. Figures without the judgement items (Amendment 12) and the 10 of 10 figure without the replacement run (Amendment 15) are kept in the private record (Amendment 45).

The registered comparison for each practice area and location. It is run and kept in the private record. The findings give the number of combinations that point the same way as each pooled result (Amendment 41, Change 4).

The comparison side of Signal A, and the Signal F score table (Amendment 43, Change 3).

Google's Rich Results Test. Registered as a separate readout beside Signal B, it is not reported for either group (Amendment 52, Change 2(c)).

The minimum, maximum and range of any group. Each gives one firm's value (Amendment 43).

A run-level figure for Proposition 1, and two exploratory tabs. No figure is reported for the share of runs that returned a firm from the organic top 10. The researcher reads Section 11's run-level statement as an example form, and Proposition 1 is addressed at firm level by Signal A. Read as a commitment, its omission is a deviation, disclosed in Amendment 46. Two unregistered tabs added on 22 September 2026 (AI vs Organic Overlap, Cross-Platform Agreement) were not re-run on the corrected firm list and are not reported (Amendment 46).

Figures that fail the reporting rules. Figures for groups below 10 firms, and near-uniform figures, are combined with others or not reported

(Amendments 41 and 43). Counts that would point to a single firm are not published, with two exceptions: one difference of one, disclosed beside the figures it concerns, and the counts for the adviser's firm in the conflict-of-interest paragraph in section 9 (Amendment 45, Change 4(e); Amendment 52, Change 6). Each withheld figure reads "not reported" in the findings.

7. Conclusions

7.1 Answer to the research question

The research question asked which observable digital signals are shared by South East Queensland law firms that appear consistently in AI-generated legal service recommendations. Against randomly drawn firms that the platforms did not name for the same query, the consistently named firms more often showed four signals:

- a listing in the QLS Find a Solicitor directory: 76% of consistent rows against 40% of comparison firms (section 5.13);
- structured data detected on the website: 62% against 31% (section 5.16);
- a Google Business Profile listing scoring 4 or 5 of 5: 43% against 24% (section 5.17);
- a more recent website post: a median of 15 days since the last dated post, against 101 days (section 5.11).

The QLS directory pattern was the most consistent. It held in all nine practice area and location combinations. Review count, recorded for commercial litigation only, ran the other way: the consistent firms' median was lower (section 5.12). LawConnect presence showed little difference (section 5.13).

These are correlations between small, unevenly balanced groups, observed after the runs. The two groups are the two ends of the range, and established firms are likely to score well on many signals at once (section 6.1). None of these patterns shows that a signal causes a firm to be named.

7.2 Result for each registered proposition

Proposition 1, organic ranking: not assessable. 39 of 74 consistent rows (53%) ranked in Google's organic top 10 for the same query. The comparison figure is withheld (sections 5.10 and 6.2), and without it the study cannot show whether organic ranking and consistent appearance are related. Rows consistent only on Perplexity, ChatGPT or Claude ranked in the top 10 less often (34%) than rows consistent on Google AI Overviews (80%), whose link to organic ranking is partly built in (section 5.10).

Proposition 2, structured data: the data point the same way as the proposition. The consistent share was higher in each practice area and in eight of the nine combinations. It was higher for rows consistent on Google AI Overviews and for rows consistent only on the other three platforms (section 5.16; re-check: reliable, alpha 1.000).

Proposition 3, profile completeness and content recency: the data point the same way as the proposition. Consistent rows more often had a listing scoring 4 or 5 of 5, in each practice area, and had posted on their websites more recently, in eight of the nine combinations (sections 5.11 and 5.17). This combined result takes Signal C's re-check result: reliable, alpha 0.905. Signal D was not re-checked.

Exploratory observations. QLS directory presence gave the clearest pattern in the study (section 5.13). A recent Google Business Profile post was more common among rows consistent on Google AI Overviews, 33% against 18%, with overlapping intervals (section 5.18). In commercial litigation, the consistent firms' median review count was lower (section 5.12).

7.3 Agreement between platforms

Each platform named many firms, and about two in five firm entries appeared in one run only (section 5.2). Just over half of the consistent rows (38 of 74) reached the threshold on one platform only, and 7 reached it on all four (section 5.4). In this market, consistent appearance on one platform was a weak guide to appearance on the others. This matches the instability reported in earlier studies (section 3.7.2).

7.4 Propositions for later confirmatory research

The registered methodology names propositions for later confirmatory research as one of the study's outputs (section 3.4). The findings suggest four:

1. Presence in the QLS Find a Solicitor directory is associated with consistent AI recommendation of Queensland law firms, after allowing for firm size and establishment.
2. Detectable structured data on a firm's website is associated with consistent AI recommendation, on platforms outside Google as well as on Google AI Overviews.
3. Google Business Profile completeness and website content recency, measured before the runs, are associated with consistent AI recommendation.
4. Consistent appearance on one AI platform is a weak predictor of consistent appearance on another.

A confirmatory test would need a second coder, signals recorded before the runs, comparison firms drawn from across the full range of appearance, and run counts large enough to test the differences between groups (sections 6.1 and 6.2).

7.5 Further research

The registered methodology names seven directions for later research (methodology v2.4, Section 16). Each one answers a limit of this study. One further direction, a composite-signal study, was considered when the study was designed and left out of the registered analysis. This report sets no timetable for any of them.

Directions registered in the methodology

1. **More runs.** A study with at least 30 runs per query per platform, to estimate how often firms are named with narrower intervals. At 10 runs, a firm named in 7 runs has an interval of about 40% to 89% (section 5.3).
2. **Twelve months of tracking.** The same nine queries run across 12 months, to show whether AI recommendations concentrate on fewer firms over time. This study covers two waves about a month apart (section 4.3).
3. **Problem-specific queries.** Queries that describe a legal problem, for personal injury and criminal defence, where people often search with urgency and want a lawyer nearby. This study used one query form, which asks for the "best" lawyers (section 6.1).
4. **Training data alone.** ChatGPT and Claude run with web search turned off, to separate the firms they name from training data from the firms they name from live retrieval. Their providers do not disclose the balance (section 6.1).
5. **A controlled intervention.** Naming frequency measured before and after a targeted change to one signal, such as adding structured data, to test whether an association in this study holds under controlled conditions. Every association in this study is a correlation, and established firms are likely to score well on many signals at once (section 6.1).
6. **Regional markets.** Ipswich, Toowoomba and Moreton Bay added. Ipswich was removed from this study's design to fund more runs, so its regional contrast rests on the Sunshine Coast alone (section 6.1).
7. **Google Gemini.** Gemini added, to test whether Google's standalone assistant behaves like AI Overviews or like the other three platforms. In this study, 80% of rows consistent on AI Overviews ranked in Google's organic top 10, against 34% of rows consistent only on the other three platforms (section 5.10).

A further direction, considered at design

A composite-signal study. A test of whether a combined score of the signals, with weights fixed in advance, predicts how often a firm is named. It was left out of this study's analysis in July 2026, before registration. Signal E is recorded in one practice area only, there was no defensible basis for setting weights, and the signals are correlated, so a combined score would mostly measure how established a firm is. A confirmatory version would need weights fixed before data collection, a new and larger sample, and a way to allow for firm size and establishment.

8. Changes from the registered method

The nine queries, the four platforms, the two waves of five runs and the 7 of 10 threshold were kept as registered. How the signals were captured, how the groups were compared and what is reported changed, as listed below. The study recorded 54 post-registration amendments, plus two restored entries (R9 and R10). Each records its trigger, the change, any limitation it leaves, and its timing. At publication, the full notes are on osf.io/c63dk, with firms replaced by study codes.

The changes that alter what the study can claim are listed here, grouped by when they were identified.

Before AI response collection (6 to 10 August 2026)

Table 25. Changes identified before AI response collection.

CHANGE	AMENDMENTS
Population filters added before capture: practice-area relevance, entity type, listings with no disclosed street address	1 to 6
Population verified before Wave 1, giving 807 listings (later corrected to 805, Amendment 37)	7
A signal log created before any signal was recorded	8

During AI response collection (10 August 2026)

Table 26. Changes identified during AI response collection.

CHANGE	AMENDMENTS
Perplexity run signed in to a study account for 87 of 90 runs; at one run, earlier study threads were visible in the account, and earlier sessions were deleted before each run after that	R9, R10, 15, 18, 37

After collection, before the website and listing signals were recorded (22 to 27 September 2026)

Table 27. Changes identified after collection, before the website and listing signals were recorded.

CHANGE	AMENDMENTS
Signals observed once, after both waves	10
Out-of-area firms and barristers' briefing suggestions not counted as firms	11, 37
Signal A pooled across platforms; firm names merged only on evidence; six comparison firms replaced; comparison firms that are consistent in another combination kept and disclosed	12, 13
Google AI Overviews recorded as first displayed, without the registered expansion step	15, 17, 37
An AI assistant transcribed the saved captures	15
One replacement run made outside its wave	15
The legal adviser's firm appears in the results: a conflict-of-interest disclosure added, later corrected for the platforms on which the firm is recorded (section 9)	9, 35, 36
Three fixed pages tested for structured data	16

During signal capture (28 September to 5 October 2026)

Table 28. Changes identified during signal capture.

CHANGE	AMENDMENTS
Capture rules set as cases arose, and applied to every later row	19 to 26
Directory signal scored out of 2: FindLaw Australia had ceased publication in December 2022	19, 27, 39

After signal capture, before the write-up (6 to 16 October 2026)

Table 29. Changes identified after signal capture, before the write-up.

CHANGE	AMENDMENTS
Mention re-check method fixed, with the gap measured per instrument	28
AI Overview counting rule set after collection, and all 91 AI Overview runs audited, not blind; 57 corrected, entered by the AI assistant, with a random 6 checked by the researcher; two departures from the audit rule	29, 32, 37
Query 03 AI Overviews counted out of 9 runs	33, 34
Practitioners' names removed from the working records, apart from 16 signal log rows redacted after the Signal B and C re-check (below); older copies and the saved captures still show names	38, 44
Signal B and C re-check method and report rules fixed before the re-check	30, 40
Minimum of 10 distinct firms for figures about firms' characteristics; comparisons pooled by practice area	41
Registered statement about the ACCC corrected	42
Near-uniform figures combined or not reported; no ranges; no figure that points to a single firm, apart from one disclosed difference of one	43, 45
Population corrected to 805 listings; QLS directory searched by firm, recorded as a deviation	37
Retention: coded six-year run log; no file naming a firm uploaded to OSF from 14 October 2026; the four named files on OSF to be deleted before the project is made public	44
No run-level figure for Proposition 1	46
Wilson intervals extended to the shares in the signal comparisons; methodology v2.4, Section 2 met in part	47

After the Signal B and C re-check, before publication (19 to 28 October 2026)

Table 30. Changes identified after the Signal B and C re-check, before publication.

CHANGE	AMENDMENTS
Signal B and C re-check results: both signals banded Reliable; the AI assistant's help during the re-check disclosed as a departure	48
Signal log corrected after the re-check lock; Signal D observed again on three records where the rule carried it over, two values corrected; practitioners' names redacted in 16 signal log rows	49
Full check of the signal log against its notes; three values corrected; a check of the data before each analysis or publication step made a standing rule	50
Second check of the signal log: a change set of 19 October 2026 with no record reviewed at source; every analysed value without a recorded source checked against its capture; twelve values corrected; 11 consistent rows outside the population list kept, as registered; published medians kept	51
Part 3 figures calculated twice and recounted; figures withheld or combined; the photo element published, departing from a reporting rule; the AI help in the mention re-check recorded; the conflict-of-interest disclosure published in full without independent legal advice, with the adviser's firm's counts as a second exception to the single-firm rule	52
"Alone" removed from the re-check wording fixed before the Signal B and C re-check, and the AI help described	53
The four named files on OSF deleted on 22 October 2026, each deletion recorded	44
The Amendment 43 rules file on OSF replaced by a coded copy and a fingerprint record; the original deleted on 28 October 2026, the deletion recorded	54

Each change's effect on the findings is stated in the findings and in section 6.

9. Conflict of interest

Felicity Jane Digital's interest. This research was conducted and is published by Felicity Jane Digital, a marketing agency that provides AI-visibility and search optimisation services to businesses, including law firms of the type examined in this study. Felicity Jane Digital is not an independent third party in relation to the market this research examines, and readers should take this relationship into account when considering the findings. The registered methodology states the same interest, and that the findings may inform the agency's services (methodology v2.4, front matter). Several of the signals measured, including structured data, Google Business Profile completeness and website content, are areas in which the agency sells services.

Clients. Felicity Jane Digital's current clients are outside the three study locations. Two former clients, whose engagements ended more than three years before the study, were named by the platforms. One is in the consistent group. The other was named in fewer than 7 of 10 runs on every platform, is in neither group, and was excluded from the population because its listing shows no street address (Amendment 6). The researcher checked the study records for both firms on 23 October 2026. Neither is named.

The study's legal adviser. A legal adviser was engaged to review the study's methodology and publication materials: defamation and identifiability risk, data retention, the adequacy of the study's disclaimer under the Australian Consumer Law, and claims about AI. The adviser's own firm appeared in the AI-generated results for one of the study's nine queries, in 4 of 10 runs on each of two platforms. That is below the threshold of 7 of 10 runs used to define consistent appearance. The firm was captured under the same protocol as every other firm, and the adviser took no part in collecting, coding or analysing the data. The adviser declined to advise on the wording of this disclosure, citing Rule 12 of the Australian Solicitors' Conduct Rules, and no lawyer has reviewed it. Neither the adviser nor the firm is named.

Independent advice on this disclosure was sought between 14 and 21 October 2026. Five firms declined. A sixth offered to advise by 26 October 2026, and the researcher declined that offer and published without independent advice, a departure from the legal advice of 12 October 2026. This disclosure replaces the one first recorded in Amendment 9 (Amendment 52, Change 5). The counts for the adviser's firm are published as a second exception to the single-firm rule, and the adviser is not named as legal reviewer (section 4.10; Amendment 52, Changes 5 and 6).

10. Acknowledgements

Dr Karen Sutherland reviewed Version 1 of the methodology in July 2026. Her review was of the methodology only. The study is an exploratory, observational study designed to generate hypotheses rather than establish causal or confirmatory conclusions.

The study's legal adviser was engaged to review the study's methodology and publication materials, as set out in section 9. The adviser is not named (Amendment 52, Change 5).

An AI assistant, Claude (Anthropic), was used throughout the study, in the roles set out in section 4.11. It made the corrections to the signal log through a spreadsheet connection, each approved by the researcher (Amendments 49 to 51). On the evening of 19 October 2026, 15 cells were changed with no record of who made the changes or of any approval; each was checked at source and decided by the researcher on 21 October (Amendment 51, Change 1). Separate instances checked the signal log before the Part 3 figures, under a standing rule set on 19 October for each later analysis or publication step (Amendments 50 and 51). Two separate instances calculated every Part 3 figure, each with its own code and without seeing the other's work, and a third recounted every figure (Amendment 52, Change 1). It also

gave the help in both re-checks set out in section 4.9. The researcher directed its work and made every decision. The checks the researcher made of its work are set out in section 4.11 and section 6.2.

11. Data availability and retention

What is public

- The registered methodology, query lock and run-log template: osf.io/wcvbk. The registration cannot be edited. It lists ten firms by website domain as examples of organic search results.
- The post-registration notes, with firms replaced by study codes: osf.io/c63dk. The files on it that named firms were deleted on 22 October 2026, before the project is made public.
- A SHA-256 fingerprint of each named original that has a published coded version, so the private record can be matched to it without identifying any firm.
- A copy of the registered materials on felicityjane.com.au, linked from this report (methodology v2.4, Section 3).
- The figures in this report. Only figures at the level of a group of firms are published. No per-firm rows, raw captures or AI responses are published.

What stays private

- The coded dataset and the coded run log.
- Every record that names a firm or an individual, and the key linking study codes to firms.

Retention

- Records that name a firm or an individual, and the code key, are kept for no more than three years from the date of first publication and then destroyed. Some are removed earlier, once their purpose has ended.
- The coded dataset, the coded run log and other de-identified working records are kept for six years from publication and then destroyed. Published versions stay as published.
- If a complaint, claim or regulator inquiry about the study is raised before a destruction date, the records it concerns are kept until it is resolved.
- Each destruction is recorded, naming no firm.

The study's data management plan sets out these periods and the disclosure-control rules summarised in section 4.10 (Amendments 38, 44 and 45).

12. References

Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Deshpande, A. (2024). GEO: Generative engine optimization. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)* (pp. 5-16). Association for Computing Machinery. https://doi.org/10.1145/3637528.3671900

Eimer, T., Schäpermeier, L., Biedenkapp, A., Tornede, A., Kotthoff, L., Leyman, P., Feurer, M., Eggensperger, K., Maile, K., Tornede, T., Kozak, A., Xue, K., Wever, M., Baratchi, M., Pulatov, D., Trautmann, H., Kashgarani, H., & Lindauer, M. (2025). *Best practices for empirical meta-algorithmic research: Guidelines from the COSEAL research network*. arXiv. https://arxiv.org/abs/2512.16491

Fishkin, R., & O'Donnell, P. (2026, January). *AIs are highly inconsistent when recommending brands or products; marketers should take care when tracking AI visibility* [Research report]. SparkToro. https://sparktoro.com/blog/new-research-ais-are-highly-inconsistent-when-recommending-brands-or-products-marketers-should-take-care-when-tracking-ai-visibility/

Gavoyannis, D. (2025, December 15). *Are AI Mode and AI Overviews just different versions of the same answer? (730K responses studied)*. Ahrefs. https://ahrefs.com/blog/ai-overviews-vs-ai-mode

Grossman, R., Liu, S., Chen, M. K., Smith, M., Borcea, C., & Chen, Y. (2026). How generative AI disrupts search: An empirical study of Google Search, Gemini, and AI Overviews. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26)* (pp. 448-459). Association for Computing Machinery. https://doi.org/10.1145/3805712.3809667

InterCore Technologies. (2026, July 20). *The state of AI search visibility for law firms (2026)* [Research report]. https://intercore.net/research/state-of-ai-search-visibility-for-law-firms-2026

Kirsten, E., Grosse Perdekamp, J., Wu, Q., Upadhyay, M., Gummadi, K. P., & Zafar, M. B. (2026). Characterizing web search in the age of generative AI. In *Findings of the Association for Computational Linguistics: ACL 2026* (pp. 10827-10848). Association for Computational Linguistics. https://doi.org/10.18653/v1/2026.findings-acl.526

Linehan, L. (2026, March 2). *Update: 38% of AI Overview citations pull from the top 10*. Ahrefs. https://ahrefs.com/blog/ai-overview-citations-top-10/

Martinez, O. (2026). *Optimizing visibility in generative engines: A critical survey of generative engine optimization (2023-2026)*. arXiv.

https://arxiv.org/abs/2607.14035

Mehta, M., Sharma, R., Kalluru, V., & Kotwal, A. (2026). Beyond the blue link: Empirical evaluation of generative engine optimization in stochastic retrieval systems. In *Proceedings of the 37th ACM Conference on Hypertext (HT '26)* (pp. 276-283). Association for Computing Machinery. https://doi.org/10.1145/3800935.3830864

Morrison, A. (2026, September 21). *Australian legal services AI visibility study: 100 search and AI attempts* [Research report]. Integral Media. https://www.integralmedia.com.au/hi/blog/legal-services-ai-visibility-study/

Schulte, J., Bleeker, M., & Kaufmann, P. (2026). *Don't measure once: Measuring visibility in AI search (GEO)*. arXiv. https://arxiv.org/abs/2604.07585

Schwab, S., & Held, L. (2020). Different worlds: Confirmatory versus exploratory research. *Significance, 17*(2), 8-9. https://doi.org/10.1111/1740-9713.01369

Sielinski, R. (2026). *Quantifying uncertainty in AI visibility: A statistical framework for generative search measurement*. arXiv. https://arxiv.org/abs/2603.08924

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tannenbaum, B. (2026). *Scoring with the engine: Retrieval exposure, cross-engine divergence, and the limits of engine-agnostic GEO scores*. arXiv. https://arxiv.org/abs/2609.22655

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), xiii-xxiii. https://doi.org/10.2307/4132319

Xu, H., Iqbal, U., & Montgomery, J. M. (2026). Measuring Google AI Overviews: Activation, source quality, claim fidelity, and publisher impact. In *Proceedings of the 2026 ACM Internet Measurement Conference (IMC '26)* (pp. 529-546). Association for Computing Machinery. https://doi.org/10.1145/3777912.3839818

Appendix A. The nine registered queries

Locked before registration (methodology v2.4, Appendix A). All 361 recorded runs used this wording exactly (Amendment 15).

Table 31. The nine registered queries.

#	PRACTICE AREA	LOCATION	QUERY
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. Signal definitions

As registered (methodology v2.4, Section 9), with the changes made during the study.

Table 32. Signal definitions, as registered and as carried out.

SIGNAL	REGISTERED DEFINITION	AS CARRIED OUT
A. Google organic ranking	Does the firm appear in Google's top-10 organic results for the same query, checked immediately after the AI response? Yes or No, with position	Pooled across every organic check recorded for the query, on all four platforms: Yes where the firm appears in more than half (Amendment 12)
B. Structured data	Detectable FAQPage, LocalBusiness or LegalService, or Person markup in the Schema.org validator: detected valid, detected with errors, or not detected; Google Rich Results Test recorded separately	Three fixed pages per website: the homepage, the matching practice-area page, the first lawyer profile (Amendment 16); capture rules in Amendments 19 to 26
C. GBP completeness	0 to 5: business description present; primary category set; services section populated with at least one service; at least one practice-uploaded photo showing people, premises or work context, excluding logo-only and Google-generated images (virtual practices scored 0 on this element and excluded from photo analysis); a GBP post visible within 90 days	Photo element narrowed to the practice's own photos of its staff, office or signage, with stock or generic imagery not counted (Amendment 21); other capture rules in Amendments 19 to 26; one listing scored for multi-office firms (Amendment 14)
D. Website content recency	Date of the most recent blog, news or journal post on the firm's own website	As registered, with capture rules in Amendments 19 to 26
E. Google review count	Total review count on the GBP listing; commercial litigation only	As registered
F. Legal directory presence	QLS Find a Solicitor, LawConnect and FindLaw Australia; scored 0 to 3	0 to 2: FindLaw Australia had ceased publication in December 2022 (Amendments 19, 27 and 39); QLS searched by firm (Amendments 19, 20 and 37)

All signals except A were observed once, after both waves (Amendment 10), between 28 September and 5 October 2026 (Amendments 19 to 26). Any later observation date is disclosed (Amendment 29, Change 4).

Appendix C. Run conditions

Table 33. Run conditions, as registered and as run.

CONDITION	REGISTERED	AS RUN
Session	New session per run, in its own Chrome incognito window	As registered, apart from one Perplexity run at which earlier study threads were visible in the account (R10)
Device and connection	One device; the researcher's Sunshine Coast connection; no VPN	No deviation recorded
Location services	Off	As registered
ChatGPT	Logged out, web search enabled	As registered
Claude	Dedicated study account, no prior history, memory and personalisation disabled where available, web search enabled	As registered
Perplexity	Logged out	Signed in to a dedicated study account, memory inactive, for 87 of 90 runs (Amendment 15)
Google AI Overviews	Incognito Google Search, no account; Overview expanded before recording	No account; the researcher decided to record the Overview as first displayed, and whether that was applied in every run cannot be confirmed (Amendments 15 and 37); firms counted from the main capture only (Amendment 29)
Organic check	Same window, straight after the AI response; for AI Overview runs, the same results page	Many AI Overview runs have a separate organic capture, and two captures minutes apart sometimes showed different Overviews; the main capture is the AI response record (Amendment 29)
Runs	5 per query per platform per wave, at least 24 hours apart; two waves at least three weeks apart; a run with no AI Overview replaced within the wave where feasible	Wave 1: 10 to 18 August 2026. Wave 2: 14 to 22 September 2026. One replacement run on 22 September 2026 counted in Wave 1 (Amendment 15). Query 03 has 9 AI Overview runs, with no replacement possible (Amendment 33)
Evidence	Full-page capture of every AI response and results page	Saved for every run; a few captures are partial (Amendment 29)

Appendix D. Glossary

- **Consistent appearance:** a firm named in at least 7 of the 10 runs of one query on one platform (7 of 9 for Query 03 AI Overviews) (section 4.5).
- **Firm, platform and query combination:** one firm, on one platform, for one query. 137 reached the threshold.
- **Practice area and location combination:** one of the nine queries (three practice areas in three locations).
- **Consistent row:** one firm for one query, where the firm reached the threshold on at least one platform. A firm consistent for two queries gives two rows. There are 74 consistent rows from 63 distinct firms.
- **Comparison firm:** a firm drawn at random from the population listings for a query that no platform named in that query's runs; 5 per query, 45 in all (section 4.8).
- **Population:** 805 Google Business Profile listings for law firms in the practice areas and locations studied (section 4.6).
- **Google Business Profile (GBP):** a business's listing in Google Search and Google Maps.
- **QLS Find a Solicitor:** the Queensland Law Society's public directory of solicitors and law firms. **LawConnect:** a public online directory of Australian lawyers and law firms.
- **Signals A to F:** A, Google organic ranking; B, structured data on the firm's website; C, Google Business Profile completeness, scored 0 to 5 on five elements, C1 to C5 (business description, primary category, services, the firm's own photos, a post within 90 days); D, website content recency; E, Google review count; F, directory presence (F1, QLS; F2, LawConnect). Section 4.7 defines each.
- **Personal injury advertising gate:** a page a visitor must click through before seeing a personal injury firm's content. Pages behind it could not be tested.
- **Re-check:** a second coding of a random sample by the same researcher, without seeing the first coding. There are two: the mention re-check and the Signal B and C re-check (section 4.9).
- **Positive specific agreement:** twice the number of firms recorded in both codings, divided by the total number of firm entries across the two codings.
- **Krippendorff's alpha:** a measure of agreement that corrects for chance; 1.000 is perfect agreement. The bands applied to it (for example "reliable") were fixed before the Signal B and C re-check (Amendment 40).
- **Wilson interval:** an approximate 95% confidence interval for a share, suited to small samples. **Cluster bootstrap:** an interval found by resampling whole runs or records, used where judgements within a record are related.
- **Near-uniform:** 90% or more, or 10% or fewer, of a group sharing one value. Such figures are combined with others or not reported (section 4.10).
- **Amendment:** a dated change recorded after registration. R9 and R10 are two entries of 10 August 2026 restored later (Amendment 18).