

WHITEPAPER | VERSION 1.1 | APRIL 2026

The AI Representation Risk Framework™

*A 7-Layer Model for Diagnosing Business Misrepresentation in
AI Search*

Renee Gersteling

Founder, Iconic Marketing

[iconic.marketing](https://www.iconic.marketing)

Iconic Marketing | Sydney, Newcastle, Penrith NSW | April 2026

Executive Brief

AI Representation Risk is the commercial risk that occurs when AI systems inaccurately surface, interpret, compare, or describe a business during buyer research, causing potential customers to form incorrect, incomplete, or commercially damaging impressions before direct enquiry.

The issue is not simply whether a business appears in AI search. A business can appear in an AI-generated response and still be described vaguely, misclassified, compared with the wrong competitors, connected to outdated sources, or presented with uncertain language that weakens buyer trust.

The AI Representation Risk Framework™ diagnoses this issue across seven layers: Discovery Suppression, Interpretation Drift, Factual Hallucination, Competitor Drift, Authority Gap, Entity Confusion, and Source Decay.

The framework is not an SEO ranking model, a governance model, or a traffic report. It is a commercial diagnostic tool for understanding how a business is discovered, interpreted, compared, and trusted inside AI-mediated buyer research.

Executive Summary

AI-assisted search tools are changing how buyers research businesses. Tools including Google AI Overviews, ChatGPT search, Perplexity, Gemini, and Microsoft Copilot can synthesise answers from multiple public sources instead of simply returning lists of links. In some buyer research sessions, a shortlist impression may be formed before a website is visited.

This creates a commercial risk that existing marketing frameworks have not adequately addressed. A business can perform well in traditional search, maintain a well-structured website, and still be described inaccurately, positioned weakly, or excluded entirely from AI-generated responses during buyer research.

This whitepaper introduces the AI Representation Risk Framework™, a seven-layer diagnostic model developed by Iconic Marketing. The framework was developed through structured observation of AI outputs across audits of more than 25 Australian businesses conducted between mid-2025 and early 2026, using ChatGPT and Google Gemini as the primary platforms. The model defines AI Representation Risk as a distinct commercial problem: a business can be visible in AI responses and still carry significant risk if the representation is inaccurate, incomplete, or commercially weak.

The seven layers are Discovery Suppression, Interpretation Drift, Factual Hallucination, Competitor Drift, Authority Gap, Entity Confusion, and Source Decay. Each describes a different failure mode observed in AI-generated business descriptions, and each carries distinct commercial consequences.

The framework is a practical diagnostic instrument, not a governance model or search ranking tool. It applies to any business for which buyer research is a meaningful part of the sales or enquiry process.

The Shift from Search Visibility to AI Representation

For two decades, digital marketing has treated search visibility as a ranking problem. Businesses invested in appearing in search engine results pages for relevant queries. The model assumed that buyers would receive a list of results, visit multiple websites, and form their own views through direct comparison.

That model is no longer complete.

AI-assisted search and answer tools can deliver synthesised responses instead of traditional lists of results, depending on the platform, query, and search context. When a user queries an AI tool, they may receive a curated output that includes business recommendations, descriptions, service comparisons, and credibility assessments assembled from multiple public sources. The user may not click through to a website at that stage.

Pew Research Center published analysis in 2025 showing that when a Google search page included an AI summary, users clicked a traditional search result in 8% of visits, compared with 15% of visits without an AI summary. The same analysis found that users were more likely to end their browsing session after a page with an AI summary than after a page without one. (Pew Research Center, 2025)

OpenAI usage research published in 2025 reported that seeking information grew from 14% to 24% of ChatGPT usage over the period studied, with practical guidance and information-seeking together representing a major share of use. (OpenAI, 2025) That shift indicates AI tools are becoming a meaningful layer in how people research products, services, and providers.

The commercial implication is that the role of a business website is no longer limited to serving direct visitors. It is also a source signal that AI systems may assess when constructing summaries. When that signal is weak, inconsistent, or outdated, AI-generated descriptions may not reflect the business accurately, regardless of how well the website is designed for human readers.

Search visibility determines whether a business appears in a list. **AI representation** determines what a buyer understands about that business before they make contact. These are different problems and require different diagnostic approaches.

Australian Business Context

AI adoption is no longer confined to large enterprises. The Australian Government's AI Adoption Tracker monitors how small and medium businesses perceive and adopt artificial intelligence, using monthly survey data from 400 businesses. (Department of Industry, Science and Resources, 2025a; 2025b) This matters because Australian SMEs are increasingly exposed to AI-enabled discovery and decision support, while many lack internal AI, data, brand intelligence, or digital governance teams.

The COSBOA and Square State of Small Business Report found that digital transformation is already well advanced among Australian small businesses, with 85% using at least one digital tool and 39% adopting artificial intelligence for tasks such as marketing content generation, customer analytics, and social media automation. (COSBOA and Square, 2025) That adoption creates opportunity, but it also increases the need for accurate public business signals.

For regional businesses, professional services firms, education providers, retail businesses, consultants, and other high-trust categories, AI representation risk is commercially relevant because buyers often research fit, credibility, location, service boundaries, and comparisons before making contact. If AI systems rely on weak or outdated information, the business may be misrepresented before it has any opportunity to correct the impression.

Why AI Representation Risk Matters

Traditional digital marketing risk is measurable. Ranking drops, traffic declines, and conversion rate problems leave data trails. The cause can usually be investigated through search, analytics, advertising, or conversion data.

AI Representation Risk is structurally different. It is largely invisible without active testing. A business receives no notification that an AI system has described it incorrectly. There is no analytics event recording that a buyer formed a wrong impression and chose a competitor. A lost shortlist position in an AI-mediated research session leaves no trace in standard reporting.

The commercial consequences fall across several categories.

Lost shortlist inclusion	When an AI system does not surface a business in response to a relevant buyer query, that business is excluded from consideration before a single enquiry is made. In competitive categories, this represents a structural disadvantage that standard marketing metrics are not designed to detect.
Incorrect buyer assumptions	When an AI system describes a business inaccurately, buyers arrive at first contact carrying wrong expectations. That creates friction at the

	earliest point of the sales relationship and may result in enquiries being abandoned once the discrepancy is identified.
Weakened credibility	Buyers researching professional service providers use AI tools to assess credibility and fit, not just availability. A business described with hedged language or missing authority signals starts at a lower credibility position than competitors described with confidence and specificity.
Misclassification	An AI system that describes a specialist as a generalist, or positions a premium provider alongside low-cost alternatives, creates misaligned buyer expectations. This affects both enquiry quality and conversion, and may damage positioning over time.
Competitor drift	AI systems may acknowledge a business in passing while directing substantive attention to competitors with stronger public signal density. The business loses shortlist prominence not because it was compared and found lacking, but because the AI had more structured material to draw on for competitors.
Stale information	For businesses that have rebranded, expanded services, changed location, or evolved their offering, AI systems may construct descriptions based on outdated public information. The buyer forms a view of a previous version of the business.
Website displaced by weaker sources	A business cannot assume its website will always be the dominant source for AI-generated descriptions. AI systems may draw on directory listings, review platforms, cached content, and third-party references, particularly when website signals are sparse or poorly structured.

Definition: AI Representation Risk

AI Representation Risk is the commercial risk that occurs when AI systems inaccurately surface, interpret, compare, or describe a business during buyer research, causing potential customers to form incorrect, incomplete, or commercially damaging impressions before direct enquiry.

Several aspects of this definition are worth stating clearly.

The risk is commercial, not technical. It does not sit primarily within AI governance, model safety, or software compliance. It is a business consequence: lost enquiries, weakened shortlist position, incorrect buyer expectations, and misaligned first contact.

The risk occurs before direct contact. By the time a business has an opportunity to correct a misconception, the shortlist decision may already have been made.

Visibility and representation are distinct. A business can appear in an AI response and still carry significant representation risk if the description is vague, outdated, factually wrong, or commercially weak.

The risk is partly a function of the business's own public signals. AI systems synthesise from available information. When that information is thin, inconsistent, or outdated, the risk of inaccurate representation increases. Businesses have more influence over this than is commonly understood.

How the Framework Was Developed

The AI Representation Risk Framework™ was developed through structured observation of AI outputs across audits of more than 25 Australian businesses conducted between mid-2025 and early 2026. The audits tested ChatGPT and Google Gemini as the primary platforms, with web search enabled where available.

Each audit followed a consistent method. An unbranded discovery query was generated from the business's website signals and run independently three times on each AI platform. AI-generated outputs were then compared against verified website content and public business information across service description, pricing, geographic scope, entity identity, and authority signals. Findings were assessed for commercial impact rather than technical accuracy alone.

Across those audits, seven distinct failure modes emerged consistently. Some were present in most audits. Others appeared in specific categories or for businesses with particular signal gaps. These failure modes form the seven layers of the framework.

The framework has not been subjected to peer review and does not claim to be exhaustive. It is a practical diagnostic model grounded in observed AI behaviour across a specific sample of Australian businesses, tested over a defined period. AI systems update frequently, and the patterns described here may shift as models evolve. The framework will be updated accordingly.

The AI Representation Risk Framework™

The framework identifies seven distinct ways in which AI-generated business descriptions can diverge from commercial reality. The layers are not mutually exclusive. Multiple layers may be active for the same business at the same time, and the combination determines the overall risk profile.

No.	Layer	Core Risk	Primary Commercial Impact
1	Discovery Suppression	AI systems fail to surface the business in response to buyer-style prompts.	The business is absent from shortlist formation before a buyer visits a website or makes contact. Standard marketing metrics will not capture this loss.
2	Interpretation Drift	AI systems partially understand the business but describe it in vague, incomplete, or commercially weak terms.	The business appears but is described in ways that fail to differentiate it. Buyers cannot assess fit accurately before contact.
3	Factual Hallucination	AI systems invent, distort, or misstate facts about services, pricing, locations, credentials, or business details.	Enquiries may arrive with wrong expectations, wasting resources and eroding trust at first contact. In credentialled or regulated industries, the risk can extend beyond commercial loss.
4	Competitor Drift	AI systems shift attention towards competitors or substitutes in a way that weakens the business's shortlist position.	The business loses shortlist position to competitors whose public content is more structured and more frequently referenced, regardless of actual capability.
5	Authority Gap	AI systems describe the business with uncertainty or qualification because authority signals are weak, inconsistent, or absent.	Buyers interpret uncertain language as lower credibility. In professional services and other high-trust categories, trust signals are central to shortlist inclusion.
6	Entity Confusion	AI systems confuse the business with another entity, brand, location, person, or organisation.	Buyers receive incorrect information about the business identity, location, or offering. Enquiries may be directed to the wrong provider.
7	Source Decay	AI systems rely on outdated or stale public information that no longer reflects the business accurately.	Buyers form impressions of who the business was, not who it is. The gap is widest after significant changes to services, leadership, pricing, or positioning.

Layer 01 | Discovery Suppression

AI systems fail to surface the business in response to buyer-style prompts.

Observed in practice

In an audit of an Australian specialist education provider, one AI platform surfaced the organisation consistently across three independent runs, while another excluded it entirely and returned competing providers instead. From that platform's response pattern, the organisation was absent from the buyer's shortlist context despite matching the query intent.

Commercial consequence

The business is absent from shortlist formation before a buyer visits a website or makes contact. Standard marketing metrics will not capture this loss.

Common causes

Sparse digital presence, insufficient third-party references or directory listings, inconsistent business name across platforms, and weak entity recognition by AI systems.

Correction priorities

Strengthen public business signals across directories and third-party sources. Ensure service and location information is explicit on the website and in structured data. Improve entity name consistency across all platforms.

Layer 02 | Interpretation Drift

AI systems partially understand the business but describe it in vague, incomplete, or commercially weak terms.

Observed in practice

In an audit of a specialist service provider, AI outputs described the organisation in broadly accurate but commercially flattened terms. The response omitted important service boundaries, support model details, audience fit, and contextual differentiators visible on the website. The summary was directionally correct, but it made the provider sound interchangeable with generalist alternatives.

Commercial consequence

The business appears but is described in ways that fail to differentiate it. Buyers cannot assess fit accurately before contact.

Common causes

Ambiguous website copy, absent service definitions, inconsistent positioning language across platforms, and mission-led language that reads well for humans but is too imprecise for machine interpretation.

Correction priorities

Define services, target audiences, and geographic scope explicitly. Use consistent language across the website and all public channels. Avoid positioning language that requires prior context to interpret correctly.

Layer 03 | Factual Hallucination

AI systems invent, distort, or misstate facts about services, pricing, locations, credentials, or business details.

Observed in practice

In one audit, an AI platform presented a specific pricing claim as current fact even though the business did not publish that pricing on its main website pages. The statement appeared to have been synthesised from older or secondary material and presented without adequate qualification.

Commercial consequence

Enquiries may arrive with wrong expectations, wasting resources and eroding trust at first contact. In credentialled or regulated industries, the risk can extend beyond commercial loss.

Common causes

Generative AI synthesising from incomplete or contradictory source material, stale cached content, model inference, and factual gaps filled with plausible but incorrect detail.

Correction priorities

Ensure the website explicitly states what the business does and does not offer. Publish factual, specific content. Correct or remove outdated public content that could be drawn into AI synthesis.

Layer 04 | Competitor Drift

AI systems shift attention towards competitors or substitutes in a way that weakens the business's shortlist position.

Observed in practice

In an audit of a local retail business, the business appeared inconsistently for a direct local product query, while several competitors were named consistently across runs. When the business did appear, it was described briefly. Competitors were described with stronger product, category, and location signals, making their relevance more obvious.

Commercial consequence

The business loses shortlist position to competitors whose public content is more structured and more frequently referenced, regardless of actual capability.

Common causes

Competitors holding higher volumes of authoritative public content, more structured data, or stronger citation frequency, while the business has relatively sparse external references.

Correction priorities

Build authoritative third-party references. Create structured content that reinforces positioning and points of difference. Monitor how competitors are described in AI responses and identify signal gaps.

Layer 05 | Authority Gap

AI systems describe the business with uncertainty or qualification because authority signals are weak, inconsistent, or absent.

Observed in practice

An AI system may qualify its description with phrases such as 'it appears', 'it may offer', or 'I could not confirm' rather than providing a confident factual summary. Credentials and expertise signals may be absent from the response even when they are visible on the website, particularly if they are not reinforced through structured data or third-party references.

Commercial consequence

Buyers interpret uncertain language as lower credibility. In professional services and other high-trust categories, trust signals are central to shortlist inclusion.

Common causes

Limited third-party endorsement, absent or inconsistent credentials in public profiles, insufficient external citations, and lack of structured data reinforcing expertise.

Correction priorities

Publish structured evidence of expertise including qualifications, awards, professional memberships, and media mentions. Ensure credentials appear consistently across the website and all public profiles.

Layer 06 | Entity Confusion

AI systems confuse the business with another entity, brand, location, person, or organisation.

Observed in practice

In a retail audit, a business operating under a name similar to another entity had overlapping information in external listings. AI and search results treated some details as closely related or interchangeable. Buyers researching the specific store could encounter information that blended details from different entities, creating uncertainty about which listing was correct.

Commercial consequence

Buyers receive incorrect information about the business identity, location, or offering. Enquiries may be directed to the wrong provider.

Common causes

Non-unique business name, inconsistent entity information across sources, and absent entity identifiers such as ABN, registered location, and structured data.

Correction priorities

Ensure entity information is explicit, consistent, and structured across the website and all public business profiles. Use schema markup to define the business entity. Claim and complete all relevant directory listings.

Layer 07 | Source Decay

AI systems rely on outdated or stale public information that no longer reflects the business accurately.

Observed in practice

In several audits, AI descriptions drew on older directory listings, superseded documents, or outdated public pages that predated changes to the business's service scope, pricing, location, or positioning. In one case, AI referenced information from an older public document and presented it as current.

Commercial consequence

Buyers form impressions of who the business was, not who it is. The gap is widest after significant changes to services, leadership, pricing, or positioning.

Common causes

Outdated website content, unclaimed or unupdated directory profiles, stale third-party articles, and insufficient recent authoritative content to displace older sources.

Correction priorities

Audit and update all public-facing content. Claim and maintain directory profiles. Publish recent, date-stamped content. Update schema markup to reflect current business reality.

Why the Website Should Be the Anchor Truth Source

AI systems do not treat all public sources equally, and source weighting varies by platform, query type, and model version. For most businesses, the website should be the single most important source of authoritative self-description in the public domain. If website signals are weak, AI systems may draw from whatever else is available, which may be lower-quality, outdated, or partially wrong.

Vendor research on AI citation patterns suggests that answer engines tend to favour sources combining credibility with machine-readable structure, though the precise weighting varies by platform, query type, and model version. (SurferSEO, 2025) This source should be treated as directional vendor research, not peer-reviewed evidence. The practical implication remains: a website needs to function not only as a human-facing communication tool, but also as a structured information source that AI systems can parse reliably.

For a website to serve that function, it needs to address the following clearly and explicitly.

- Who the business serves, including specific industries, business types, or buyer profiles.
- What it offers, described with specificity rather than marketing generality.

- Where it operates, including geographic coverage and any service or location boundaries.
- What it does not offer, to reduce the risk of misclassification or false attribution.
- Credentials and authority markers, including qualifications, memberships, awards, and recognitions.
- Evidence of expertise, including case summaries, publications, media references, and client outcomes where these can be disclosed.
- Entity information, including the registered business name, operating location, ABN, and contact details consistent with directory listings.
- Structured data where relevant, using schema markup to define the business entity, services, location, and people in a format AI systems can parse reliably.

Absence, ambiguity, or inconsistency in any of these elements increases the likelihood that one or more framework layers will be active.

The website is not the only source that matters. Third-party directories, Google Business Profile, Bing Places, review platforms, professional association listings, and media coverage all contribute to the public signal set that AI systems assess. The website should be the anchor, but the consistency of signals across all external sources determines how confidently AI systems can construct an accurate description.

From Diagnosis to Correction

Identifying which layers are active is the starting point. Correction requires prioritised action on the specific signals that are weak, inconsistent, or absent. The following priorities are broadly applicable, though the order of urgency depends on diagnostic findings.

Clarify website content	Audit website content against the elements listed above. Identify where service definitions, geographic scope, credentials, and entity information are absent or ambiguous. Revise to be explicit and specific. Marketing language that persuades human readers is often too vague for reliable machine interpretation.
Strengthen public profiles	Claim, complete, and maintain all relevant directory and business profile listings, including Google Business Profile, Bing Places, industry-specific directories, and professional association listings. Ensure the business name, address, phone number, and service categories are consistent across all platforms.
Improve entity consistency	The business name, structure, and identifying details should appear in the same format across the website, directories, social profiles, and all public references. Inconsistency across sources is a primary contributor to Entity Confusion and Authority Gap.
Remove or update outdated content	Identify public content that no longer accurately reflects the business and update or remove it. This includes old website pages, superseded directory listings, and stale third-party content. Outdated material

	does not disappear from AI synthesis simply because it has been replaced on the website.
Add proof signals	Publish structured evidence of expertise and authority. This includes professional qualifications, awards, memberships, media mentions, and case summaries that can be disclosed without breaching confidentiality. These signals give AI systems material to construct a confident, credible description.
Implement structured data	Use schema markup appropriate to the business type, including LocalBusiness, Service, Person, and Organisation schemas as applicable. Structured data reduces the reliance of AI systems on inference by providing machine-readable definitions of the business.
Monitor AI outputs	AI representations change as models are updated, public signals evolve, and competitors adjust their content. Businesses should test AI outputs periodically using buyer-style prompts across multiple platforms, compare those outputs against current website content, and identify any representation drift.

| Application: The AI Brand Signal Audit

The AI Brand Signal Audit is the first commercial application of the AI Representation Risk Framework™. It applies the seven-layer model to a specific business, testing how AI systems describe, compare, and surface that business, then translating the findings into evidence, severity ratings, and prioritised corrective actions.

The audit generates unbranded discovery and interpretation queries from the business's website signals and runs those queries independently across tested AI platforms, with web search enabled where available. It compares AI-generated outputs against verified website content and public business information. It identifies which of the seven framework layers are active and assesses the severity of each. It provides specific evidence of representation gaps and recommends corrective actions in order of commercial priority.

The audit is not a search ranking tool, a traffic report, or a brand sentiment measure. It is designed to answer a specific question: when AI systems are asked about this business, what do they say, and how does that compare to what is actually true?

Limitations and Responsible Use

AI outputs are not stable	AI systems do not produce identical outputs for identical queries. Responses vary across time, platform, prompt phrasing, user context, and model version. A single test is a snapshot. Findings should be treated as diagnostic signals that require pattern confirmation across multiple tests and platforms.
No corrective action guarantees accurate AI representation	Businesses cannot instruct AI systems to describe them in a particular way. The goal of corrective action is to provide better inputs so that AI systems have more accurate and authoritative material to draw on. This reduces representation risk but does not eliminate it.
Findings require contextual interpretation	An inaccurate AI output observed during a test does not mean the system always produces that output. An accurate output does not confirm consistent accuracy. Findings should be weighted by frequency across multiple tests and interpreted in the context of the platforms and prompt types used.
Signal improvement requires ongoing maintenance	Strengthening public signals is not a one-time project. AI models are updated, competitors adjust their content, and business information changes. Signal quality needs to be maintained and monitored over time.
Ethical responsibilities	The framework is a diagnostic and correction tool for improving accuracy of representation. Businesses should not use any understanding of AI systems to publish false or misleading information, inflate authority signals artificially, or attempt to manipulate AI outputs in ways that misrepresent the business to buyers. (NIST, 2023)

Conclusion

AI answer tools are not simply replacing traditional search. They are adding a research layer that operates on different mechanics and creates different commercial risks. That layer is already present across a range of buyer research contexts and is likely to grow in scope.

The businesses most exposed to AI Representation Risk tend to share common characteristics: sparse or inconsistent directory coverage, website content that is persuasive but not specific, absent structured data, inconsistent entity information across platforms, and limited third-party authority references. These are also the businesses for whom existing SEO and digital marketing tools offer the least useful diagnostic information, because those tools measure traffic, rankings, and clicks rather than representation accuracy.

Managing AI Representation Risk is not simply a new form of search optimisation. It is a discipline of maintaining accurate, consistent, and structured public information so that AI

systems have reliable inputs when buyers ask about the business. The commercial stakes are the same as any other buyer research channel: shortlist inclusion, accurate first impressions, and a credible starting position for the sales conversation.

The AI Representation Risk Framework™ provides a structured basis for diagnosing where representation is falling short and what corrective action is warranted. As AI-mediated buyer research becomes more common, this type of diagnostic capability is likely to become an important part of responsible marketing practice.

A Note on Framework Positioning

The AI Representation Risk Framework™ builds on established work across AI risk management, bias detection, search visibility, entity resolution, and digital trust. Adjacent frameworks from NIST, the Cloud Security Alliance, and KPMG address AI risk at a governance and model management level. AI search optimisation literature addresses citation frequency, visibility, and traffic. Appendix A provides a comparison of these adjacent frameworks and their relationship to the problem this framework addresses.

To our knowledge, the AI Representation Risk Framework™ is a new diagnostic model for the specific problem of commercial representation accuracy in AI-generated responses. We have not identified an established framework that combines these seven risk dimensions into a single commercial diagnostic instrument. We have not conducted a systematic academic literature review and do not claim this is the first framework of its kind. It is positioned as a practical response to a problem that current frameworks and tools do not adequately address.

References

Sources used in this whitepaper are listed below. Where a source is vendor-published or not independently peer-reviewed, this is noted. Claims based on interpretation rather than direct citation are presented as such in the body text.

Pew Research Center (2025). Do people click on links in Google AI summaries?

<https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/>

OpenAI (2025). How people are using ChatGPT. <https://openai.com/index/how-people-are-using-chatgpt/>

Chatterji, A., Deming, D., et al. (2025). How People Use ChatGPT. OpenAI economic research working paper. <https://cdn.openai.com/pdf/a253471f-8260-40c6-a2cc-aa93fe9f142e/economic-research-chatgpt-usage-paper.pdf>

Department of Industry, Science and Resources (2025a). AI Adoption Tracker. <https://www.industry.gov.au/publications/ai-adoption-tracker>

- Department of Industry, Science and Resources (2025b). AI adoption in Australian businesses for 2025 Q1. <https://www.industry.gov.au/news/ai-adoption-australian-businesses-2025-q1>
- COSBOA and Square (2025). State of Small Business Report, August 2025. <https://cosboa.org.au/wp-content/uploads/2025/10/State-of-Small-Business-Report-August-2025.pdf>
- NIST (2023). Artificial Intelligence Risk Management Framework. <https://www.nist.gov/itl/ai-risk-management-framework>
- Cloud Security Alliance (2024). AI Model Risk Management Framework. <https://cloudsecurityalliance.org/artifacts/ai-model-risk-management-framework>
- KPMG (2024). Effective Model Risk Management Framework for AI/ML-based Models. <https://assets.kpmg.com/content/dam/kpmgsites/in/pdf/2024/10/effective-model-risk-management-framework-for-ai-ml-based-models.pdf>
- RACGP / Australian Journal of General Practice (2023). Bias in artificial intelligence and data-driven diagnostic tools. <https://www1.racgp.org.au/ajgp/2023/july/making-decisions>
- Griffith University (2025). Griffith research explores the psychology of AI-induced consumer vulnerability. <https://news.griffith.edu.au/2025/10/13/griffith-research-explores-the-psychology-of-ai-induced-consumer-vulnerability/>
- SurferSEO (2025). AI Citation Report 2025: Which Sources AI Overviews Trust Most. <https://surferseo.com/blog/ai-citation-report/> [Vendor research; directional, not peer-reviewed.]
- DemandScience (2026). AI Search Brand Misrepresentation. <https://demandscience.com/resources/blog/ai-search-brand-misrepresentation/> [Vendor content; directional framing only.]

Appendix A: Adjacent Framework Landscape

The following table summarises the frameworks most relevant to AI Representation Risk and explains how each differs in scope and focus. The comparison is based on a scoped review of publicly available framework documentation. It does not constitute a systematic academic literature review.

Framework	Primary focus	Scope	Gap relative to AI Representation Risk
NIST AI Risk Management Framework (2023)	Governance, trustworthiness, and safety of AI systems	AI developers, deployers, and operators. Covers bias, explainability, robustness, and privacy.	Addresses AI system risk broadly, not the commercial accuracy of how specific businesses are described in AI outputs.
Cloud Security Alliance AI Model Risk Management Framework (2024)	Model bias, hallucination, misuse, and reputational harm in AI deployments	Enterprises deploying AI. Covers validation, monitoring, and governance of AI models.	Covers hallucination as a model risk, not as a commercial representation problem experienced by the business being described.
KPMG AI/ML Model Risk Management Framework (2024)	Model validation, bias, explainability, and fit-for-purpose assessment	Financial services and regulated industries. Covers model lifecycle and risk governance.	Addresses internal AI model risk management. Does not address how external AI systems describe or misrepresent third-party businesses.
SEO and Generative Engine Optimisation (GEO) literature	Citation frequency, traffic, rankings, and AI search visibility	Marketers, SEO practitioners, and publishers. Covers being cited in AI answers and maintaining organic traffic.	Optimises for visibility and citation volume. Does not diagnose the accuracy, completeness, or commercial quality of AI-generated business descriptions.
AI Representation Risk Framework™ (Iconic Marketing, 2025)	Accuracy and commercial quality of AI-generated business descriptions during buyer research	Business owners, marketers, consultants, and advisers. Covers discovery, interpretation, hallucination, competitor positioning, authority,	Designed specifically for this problem. Diagnoses representation risk across seven dimensions rather than measuring traffic, citation frequency, or model governance.

Framework	Primary focus	Scope	Gap relative to AI Representation Risk
		entity identity, and source quality.	

Sources: NIST (2023); Cloud Security Alliance (2024); KPMG (2024); SurferSEO (2025); DemandScience (2026).

Appendix B: Methodology Summary

This appendix summarises the observed-audit method used to develop the AI Representation Risk Framework™. It is included to make the framework's evidence base and limitations clear.

Method area	Approach used
Sample base	More than 25 Australian businesses audited between mid-2025 and early 2026.
Business types	Service-led, product-led, and hybrid organisations across local, regional, and professional categories.
Primary platforms tested	ChatGPT (OpenAI) and Google Gemini. Web search enabled where available.
Test frequency	Each core discovery query was run independently three times per platform to identify consistency, variance, and recurring failure patterns.
Prompt types	Buyer-style prompts across discovery, comparison, service interpretation, credibility, and location-relevant scenarios.
Truth-source comparison	AI-generated outputs compared against verified website content, public business profiles, directory listings, public documents, and other accessible business signals.
Assessment focus	Findings assessed for commercial impact, not technical accuracy alone. The question was whether the representation could affect discovery, trust, fit, comparison, or buyer expectation.
Framework development	Recurring failure patterns grouped into seven diagnostic layers: Discovery Suppression, Interpretation Drift, Factual Hallucination, Competitor Drift, Authority Gap, Entity Confusion, and Source Decay.
Limitations	The sample was practical, not academic. It was not peer-reviewed and does not claim statistical representation of all Australian businesses. AI outputs change across time, platform, user context, and model version. Findings should be treated as diagnostic evidence rather than fixed behaviour.

Responsible interpretation

The framework is intended to identify commercial risk patterns and guide corrective action. It should not be interpreted as a guarantee of AI behaviour, a ranking system, or a claim that any platform will produce identical results for every user. The purpose is to help businesses strengthen the public signals that AI systems may use when buyers research, compare, and evaluate them.

Iconic Marketing

iconic.marketing

Sydney | Newcastle | Penrith, NSW, Australia