Concept Paper

Not peer-reviewed version

From Information Retrieval to Agentic Action: A Framework for Brand Visibility in AI-Mediated Markets

Marcos Guimaraes Figueira *

Posted Date: 18 May 2026

doi: 10.20944/preprints202605.1208.v1

Keywords: generative engine optimization; answer engine optimization; agentic AI; brand equity; consumer–AI delegation; large language models; retrieval-augmented generation

Concept Paper

From Information Retrieval to Agentic Action: A Framework for Brand Visibility in AI-Mediated Markets

Marcos Guimaraes Figueira [1,2]

[1] Fundação Getulio Vargas (FGV), Rio de Janeiro, Brazil; marcos.figueira@fgv.edu.br
[2] Wyse Brand Intelligence, Rua do Russell, 804/4º andar, Rio de Janeiro, 22210-010 RJ, Brazil

Abstract

The competitive ground of digital visibility has moved twice in eighteen months. The first move was from the ten blue links to the AI-generated answer; the second, still in progress, is from the answer to the autonomous action. This article develops a framework that integrates three optimization disciplines emerging in response to that shift—Answer Engine Optimization (AEO), Generative Engine Optimization (GEO), and Agentic Optimization (AgO)—under a single theoretical lens: delegated consumer–AI agency. Building on Puntoni et al.'s (2021) experiential perspective on consumer AI and Davenport et al.'s (2020) account of AI in marketing, we treat the AI assistant not as a channel but as a delegated decision-maker whose choices are governed by retrievability, attribution cost, and embedded brand associations. We argue that the central strategic risk of this transition is *brand erasure*—the systematic elimination of brand identity from synthesized answers and completed agentic actions—and we develop six propositions that link content and infrastructure choices to brand outcomes in this environment. We close with managerial implications for brands operating across the dual audience of human readers and machine intermediaries, and a research agenda focused on measurement, audit methods, and governance.

Keywords: generative engine optimization; answer engine optimization; agentic AI; brand equity; consumer–AI delegation; large language models; retrieval-augmented generation

---

1. Introduction: From Pages to Answers to Actions

In March 2025, Pew Research Center tracked 68,879 Google searches across 900 U.S. adults. When the search produced an AI Overview, users clicked a traditional result 8% of the time, against 15% for searches without one. They clicked a link *inside* the AI summary 1% of the time. The summaries themselves cited three or more sources in 88% of cases (Pew Research Center, 2025). That is not a softer click-through curve. It is a phase change in how attention reaches brands.

The phase change has a simple shape. Until recently, the search engine was an index—a directory that pointed users at pages where the work of persuasion happened. The page was the unit of competition. Now the answer is the unit of competition: the user receives a synthesized response, often with no need to leave the chat surface, and the brand either appears inside that response or it does not. A second move is already underway. Agents built on tool-augmented language models (Schick et al., 2023; Yao et al., 2023) and on computer-use APIs (Anthropic, 2024; OpenAI, 2025) no longer stop at the answer. They book the flight, write the SQL, request the refund. The unit of competition becomes the *executed action*. The brand either gets selected, called, transacted with—or it does not.

Practitioners have responded with three new acronyms. Answer Engine Optimization (AEO) targets the extraction of facts into featured snippets, voice assistants, and AI Overviews. Generative Engine Optimization (GEO), introduced as a research paradigm by Aggarwal et al. (2024), targets the probability of being *cited* in synthesized responses. Agentic Optimization (AgO)—a term still

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 May 2026 doi:10.20944/preprints202605.1208.v1

settling—targets the conditions under which an autonomous agent will *select your platform* to perform a task. These are usually treated as separate playbooks. They are not. They are three layers of the same problem.

This article makes three contributions. First, we propose a unified framework—AEO–GEO–AgO as the optimization stack—grounded in delegated consumer–AI agency rather than the loose "Agency Theory 2.0" framing that has circulated in trade literature. Second, we elevate *brand erasure*—the systematic suppression of brand identity inside synthesized answers and completed agentic actions—as the focal strategic risk of the transition, and offer detection and response mechanisms. Third, we translate the framework into propositions that can be tested with audit methods already in use elsewhere in marketing science, and we map the managerial KPI shift that follows from those propositions.

The argument runs as follows. Section 2 develops the theoretical foundation, drawing on consumer–AI delegation, brand-as-knowledge-network, and the embedding-bias literature. Section 3 lays out the optimization stack and develops six propositions. Section 4 treats brand erasure as a focal construct. Section 5 maps managerial implications onto the dual audience of human and machine. Section 6 addresses the regulatory environment, primarily the EU AI Act. Section 7 sets a research agenda.

2. Theoretical Foundations: Delegated Agency in AI-Mediated Markets

2.1. Why Classical Agency Theory Is the Wrong Starting Point

The instinct to reach for Jensen and Meckling (1976) when an intermediary appears between principal and outcome is understandable, but the fit is poor. Classical agency theory addresses information asymmetry, moral hazard, and incentive contracting between two human parties with separable interests. The AI assistant has none of these in the strict sense. It does not have interests; it has objective functions. It is not subject to monitoring costs in the human sense; it is subject to inspection at the level of weights, prompts, and tool calls. Its "shirking" is not laziness but hallucination, retrieval failure, and reranking bias. Borrowing the agency vocabulary without these adjustments produces a metaphor without theoretical traction.

A more productive starting point is the consumer–AI delegation literature. André et al. (2018) frame the autonomy trade-off when consumers offload choice to algorithmic systems. Puntoni et al. (2021) describe four experiential modes of AI in consumption—data capture, classification, delegation, and social—and emphasize that delegation reshapes the consumer's relationship to the brand by inserting a non-human filter. Hildebrand and Bergner (2021) show that conversational agents act as *surrogates of trust*: the user transfers part of the trust they would extend to a firm onto the agent that mediates it. Davenport, Guha, Grewal, and Bressgott (2020) place these effects within a broader account of AI's marketing implications. Across these works, the agent is not a channel; it is a *decision-making intermediary* whose criteria differ from the user's stated preferences in ways that matter for brand outcomes.

We will refer to this framing as **delegated consumer–AI agency**. It carries three commitments that classical agency theory does not. The intermediary is computational rather than economic; its choice procedure is governed by retrieval and reranking rather than incentive design; and the brand's leverage on the intermediary runs through the digital infrastructure the brand exposes, not through contracting.

2.2. The Principal–Machine–Brand Triad

The relationship to model is therefore not principal–agent but principal–machine–brand. A consumer (the principal) issues a request to an AI assistant (the machine). The assistant constructs a response by retrieving documents, reranking them, optionally calling tools, and generating language. Brands appear, get cited, get omitted, get transacted with—all inside the machine's procedure. The principal sees only the output.

 doi:10.20944/preprints202605.1208.v1

Three asymmetries follow. The first is *opacity*: the principal cannot inspect why one brand was selected over another, and the brand often cannot either. The second is *substitutability*: when the agent satisfies the underlying need, the specific brand identity becomes optional from the principal's standpoint. A user who asks for "a good travel jacket under $200" may not need to know whether the answer came from one retailer or five. The third is *infrastructure dependence*: brands that expose richer machine-readable signals—structured data, manifests, APIs—operate the choice procedure on more favorable terms than brands that do not. Visibility migrates upstream, into the architecture.

2.3. The Brand as an Entity in Latent Space

To reason about brand outcomes inside the machine's procedure, we need a representation of the brand the machine actually uses. Customer-Based Brand Equity (Keller, 1993) defined brand knowledge as a network of associations in consumer memory: a graph whose nodes carry awareness, valence, and uniqueness (Krishnan, 1996). The translation to language models is direct. The brand exists, for the model, as a vector—or more precisely, as a region in a high-dimensional embedding space—surrounded by attribute vectors. Cosine proximity between the brand vector and attribute vectors is not a metaphor for brand associations; it is the same construct, instantiated in a different substrate (Bolukbasi et al., 2016; Caliskan, Bryson, & Narayanan, 2017).

This matters because the embedding is the substrate the assistant *thinks in*. When a user asks the model for "a reliable mid-range running shoe," the relevant brands are surfaced via similarity to {reliable, mid-range, running, shoe} in the latent space. If your brand vector clusters with "discount" rather than "reliable," your visibility loss happens before retrieval, in a layer the SEO industry has historically ignored. The implication is twofold: brand-equity work and search-marketing work, long treated as adjacent crafts, become the same craft. The unit of analysis is the latent representation.

2.4. The Epistemic Effort Proposition

The retrieval and reranking literature makes a recurring claim: at decision time, the assistant minimizes verification cost subject to relevance. Liu, Zhang, and Liang (2023) show that production generative search engines cite sources but frequently misalign citations to claims, suggesting that what looks like fastidious sourcing is partly a heuristic for low-cost verification. RAG systems (Lewis et al., 2020; Gao et al., 2024) and late-interaction retrievers (Khattab & Zaharia, 2020; Santhanam et al., 2022) make this explicit: passages that are syntactically extractable, semantically grounded, and well-anchored to authoritative entities are cheaper to incorporate and easier to cite.

We propose, as a non-formal but testable claim:

> ***Proposition 1 (Epistemic Effort).*** *Holding semantic relevance and entity authority constant, an AI agent's probability of citing or selecting a source decreases as the verification cost of its claims increases.*

We resist the temptation to formalize this as an equation. The variables are not independently observable, and pseudo-equations import a precision the construct does not have. The proposition is sufficient to generate audit-style empirical work—and it organizes the optimization stack that follows.

3. The Optimization Stack: AEO, GEO, AgO

The three optimization disciplines are best understood as targeting three sequential stages of the assistant's procedure: extraction, synthesis, and execution. Treating them separately produces tactical confusion. Treating them as a stack clarifies what each is doing and what it is not.

3.1. AEO—The Extraction Layer

AEO targets *closed-domain* queries: questions with a single correct, factual answer that can be lifted from a passage. Featured snippets, voice-assistant readouts, and one-shot AI Overviews live here. The optimization tactic is propositional density: short, declarative sentences with explicit subjects, FAQ-style schema markup, and what we will call *atomic propositions*—single-fact units that can be

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 May 2026 doi:10.20944/preprints202605.1208.v1

extracted without surrounding context. The metric is fact-extraction accuracy: the share of the brand's factual claims that the assistant retrieves verbatim or near-verbatim.

AEO is the layer most continuous with classical SEO. Its failure mode is well-understood: when the assistant lifts the answer and the user's task ends, the brand provides the data and gets none of the attention. This is the smallest case of brand erasure—call it *factual disintermediation*—and the tactical response is to bind the fact to the brand's name in the same proposition ("Wyse's Q4 churn rate was...") so that extraction carries identity with it.

3.2. GEO—The Synthesis Layer

GEO targets *open-domain* queries where the assistant synthesizes a narrative from multiple sources. Aggarwal et al. (2024) introduced the term in their KDD paper and showed that visibility in synthesized responses is sensitive to specific content properties: inclusion of citations, statistics, quotations, fluency-improving edits, and easily extractable definitional language. Their benchmark (GEO-Bench) reports up to 40% visibility lift from these interventions.

The unit of competition here is *citation worthiness*. The metric is share-of-citation: among synthesized responses to a population of relevant queries, the share that include the brand as a cited or named source. Two underappreciated levers drive this share. The first is *information gain*—the marginal informational contribution of the source compared to others retrieved for the same query. Original statistics, proprietary data, and named definitions create dependency relationships the assistant has to honor. The second is *definitional anchoring*: when a brand defines a term that becomes load-bearing in the answer, the assistant cites the definer disproportionately, even when the term is widely repeated downstream.

A subtle implication: GEO rewards content that is *defensible* against paraphrase. Liu, Zhang, and Liang's (2023) audit found that 51.5% of generated sentences are fully supported by their citations and only 74.5% of citations support their associated sentence. The gap is not random. Paraphrase-friendly sources get reused without attribution; paraphrase-resistant sources—those with named methodologies, distinctive phrasing, or proprietary entities—survive the synthesis pass with their attribution intact.

Proposition 2 (Information Gain). *Sources whose informational contribution cannot be cheaply substituted from the rest of the retrieved corpus achieve higher share-of-citation in synthesized responses, controlling for entity authority.*

Proposition 3 (Definitional Anchoring). *Brands that introduce or operationalize terminology that becomes load-bearing in synthesized answers receive citations at rates disproportionate to their share of the underlying corpus.*

3.3. AgO—The Action Layer

AgO targets a different class of interaction altogether: the assistant is no longer producing a text but executing a task. Tool-using language models (Schick et al., 2023) and ReAct-style agents (Yao et al., 2023) issue API calls, fill forms, run queries, and complete transactions. Computer-use APIs (Anthropic, 2024) and GUI-controlling agents (OpenAI, 2025) extend the same logic to applications that lack APIs. Wang et al. (2024) survey the rapidly expanding literature on LLM-based autonomous agents.

The unit of competition is *interoperability*. The metric we propose is the agent execution rate: the share of tasks that an agent attempting work on the brand's surface successfully completes. The tactics are infrastructural—published manifests (the de facto standard now extends beyond the original `ai-plugin.json`), machine-readable schemas for product, pricing, and inventory, idempotent endpoints, low-latency authentication flows, clear error semantics, and explicit affordance documentation written for an agent reader rather than a human developer. The platforms that win at this layer are those whose machine surface is closer to a well-documented public API than to a website.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 May 2026 doi:10.20944/preprints202605.1208.v1

Proposition 4 (Interoperability Premium). *Among comparable platforms in a category, those that expose richer machine-readable affordances (manifests, structured product data, idempotent transactional endpoints) achieve higher agent execution rates and capture a disproportionate share of agentic transactions.*

The failure mode at this layer is the most consequential of the three. When an agent cannot complete a task on your platform—because the action surface is not legible—it does not retry. It substitutes. The brand is not delisted from a results page; it simply never enters the action.

3.4. The Stack, Summarized

The three layers are not substitutes. A brand can win at AEO and lose at AgO; a brand strong at AgO can still be erased from the answers that route consumers to its action surface in the first place. The strategic question is not which layer to invest in but which layer is currently rate-limiting for the brand's category. Categories dominated by transactional intent (booking, ordering, scheduling) are AgO-bound. Categories dominated by considered purchases (B2B software, financial services, education) remain GEO-bound. AEO is universal but rarely decisive on its own.

Table 1. The optimization stack. Each layer targets a distinct stage of the AI assistant's procedure—extraction, synthesis, execution—with a corresponding tactic, metric, and dominant failure mode.

Layer	Target	Optimization unit	Metric	Failure mode
AEO	Closed-domain extraction	Atomic propositions, schema markup	Fact extraction accuracy	Factual disintermediation
GEO	Open-domain synthesis	Information gain, definitional anchoring	Share-of-citation	Brand erasure
AgO	Task execution	Manifests, structured affordances	Agent execution rate	Action substitution

4. Brand Erasure as the Focal Strategic Risk

If the optimization stack is the *positive* program, brand erasure is the *negative* one—the failure mode the framework is built to prevent. We define it as follows.

> **Brand erasure** is the condition under which an AI assistant satisfies a user's underlying need without surfacing, citing, or transacting with the brand whose content, product, or service the response depends on.

Three observations make erasure the strategic question of the next decade.

First, it is *inherent to the technology*, not a side effect. Synthesis collapses sources into prose. Action collapses choices into a single call. A well-functioning assistant should be willing to omit attribution when attribution does not help the user—and most attribution does not help the user.

Second, it is *partially asymmetric*. Established, named brands with high-authority entity grounding (Wikipedia entries, Wikidata identifiers, well-anchored knowledge-graph relationships) are erased less often than challenger brands whose distinctiveness is carried in copywriting rather than in entity infrastructure. The legacy of brand investment shows up in the latent space, but only if the investment was indexed.

Third, it is *measurable*. Audit methods adapted from the algorithmic-bias literature can detect erasure at scale. The basic procedure is straightforward: define a set of category-relevant prompts, sample responses across multiple assistants and time, and compute brand-mention frequency relative to a baseline of corpus presence. The same audit can be run for actions: define a set of category-relevant tasks, attempt them with an agent, and measure successful completion on the brand's surface relative to substitutable alternatives.

4.1. Mechanism: When Erasure Occurs

Erasure operates through three mechanisms, each tied to a layer of the stack.

In AEO, erasure happens when a brand's factual content is extracted without identifying the source. The defense is binding facts to brand identity in the same proposition, and structuring the data so that schema markup forces attribution.

In GEO, erasure happens when a brand's content is paraphrased into a synthesis that omits the citation, often because the content was paraphrase-friendly in the first place. The defense is information gain: providing content that the assistant cannot synthesize without attribution because no comparable source exists.

In AgO, erasure happens when an agent fulfills a category-level task without selecting the brand's surface. The defense runs in two directions—*making the surface selectable* (interoperability) and *making the brand selection-influencing* upstream (the brand's presence in the model's category-level priors).

4.2. Detection: Measuring Erasure

Two measurement traditions can be combined to operationalize erasure detection.

The first is the *embedding probe*, adapted from word-embedding bias research (Bolukbasi et al., 2016; Caliskan et al., 2017). For a target brand, define a set of category-relevant attribute vectors (e.g., for an athletic brand: {performance, durability, innovation, premium, accessible}). Compute cosine similarities between the brand vector and each attribute vector across multiple foundation models. Repeat for direct competitors. Compare. Track over time. The probe yields a measurable construct—*semantic equity*—that can be benchmarked, monitored, and reported.

The second is the *response audit*. Define a representative set of prompts a real user might issue in the brand's category. Sample responses across assistants. Compute brand-mention frequency, share of citations, sentiment of the citing context, and—for AgO—task-completion rate. The audit can be automated, repeated, and used to compute month-over-month deltas that are far more strategically useful than the click-through metrics they replace.

Proposition 5 (Erasure Asymmetry). *Brands with higher entity-graph presence (canonical entries in public knowledge graphs, structured-data coverage, and stable named-entity recognition) experience lower erasure rates in both synthesized responses and agentic completions, controlling for category and corpus presence.*

4.3. Response: Gatekeeping and Anchoring

The strategic responses divide along an *open* and *closed* axis.

The *open* (offensive) strategy is semantic anchoring. The brand operates as if every public document is training data and every term it cares about is a candidate for definitional capture. It publishes its proprietary methodologies under named labels. It seeds its category vocabulary in academic, regulatory, and industry literature. It builds relationships in the entity graph through high-authority co-mentions. The point is to make the brand's terms inseparable from the category's terms, so that paraphrase cannot strip the brand from the synthesis without breaking the meaning.

The *closed* (defensive) strategy is agentic gatekeeping. Premium data, proprietary insights, and competitive intelligence are placed behind authentication. Agents that want access must identify themselves and accept attribution conditions. The risk of this strategy is also its mechanism: by forcing identification, the brand preserves its source identity at the cost of corpus presence. For categories where the asset is the proprietary insight (financial research, premium analytics, regulated data), this is a defensible trade. For commodity content categories, it is a way to disappear.

Most brands will need both, in proportions that vary by asset. Open content for the top of the funnel—where the goal is corpus presence and definitional anchoring. Closed content for the bottom—where the goal is forcing identification at the moment of value transfer.

Proposition 6 (Hybrid Gatekeeping). *In categories with mixed content economics, brands that combine open semantic anchoring with selective gatekeeping of bottom-of-funnel assets achieve higher long-term brand-attribution rates than brands pursuing a uniformly open or uniformly closed strategy.*

5. Managerial Implications

5.1. The Dual Audience: Machine Experience as a Layer, Not a Replacement

Trade commentary has begun to argue that *machine experience* (MX) replaces user experience (UX). It does not. The machine is an intermediary, not the principal. The user remains the ultimate consumer, and the agent's outputs are still consumed by humans—whether read, watched, or experienced as a completed transaction. Brynjolfsson's (2022) "Turing Trap" warning applies here: the move from augmentation to imitation is a real risk, and treating UX as obsolete is one form of taking the wrong side of the trap.

The correct framing is dual-audience design. Every public surface of the brand has two readers: the human who eventually consumes the result, and the machine that mediates access. The machine's reading happens first and shapes whether the human reading occurs at all. The two readings are not in tension when the machine layer is treated as additive: structured data does not replace prose, manifests do not replace user interfaces, and authoritative sourcing does not replace creative copy. They are the substrate on which the human layer becomes findable.

The organizational implication is that SEO, brand, content, and product-engineering functions need to operate against a shared technical substrate. The traditional handoff—brand specifies, content writes, SEO optimizes, engineering ships—produces machine-illegible artifacts at every step. The substrate has to be designed in.

5.2. The KPI Shift

The KPIs of the click economy—sessions, click-through rate, bounce rate, time on page—are not wrong, but they answer a question that is becoming peripheral. The relevant questions in the AI-mediated layer are different.

Table 2. The KPI shift from the click economy to the AI-mediated layer. None of the new KPIs is exotic; each can be measured with audit methods adapted from existing toolkits.

Click-economy KPI	AI-mediated KPI	What it measures
Click-through rate	Share-of-citation	How often the brand appears in synthesized answers in its category
Page rank position	Embedding proximity	Where the brand sits in the latent space relative to category attributes
Keyword density	Definitional anchoring	Whether the brand's terms become load-bearing in category answers
Page load speed	API latency	Whether the brand's machine surface clears agent timeouts
Conversion rate	Agent execution rate	Whether agents successfully complete tasks on the brand's surface
Brand search volume	Brand entity grounding	Whether the brand has stable, canonical representation in public knowledge graphs
Bounce rate	Hallucination rate	How often AI assistants misrepresent the brand's pricing, features, or claims

None of these new KPIs is exotic. Each can be measured with audit methods that already exist in the algorithmic-accountability literature (Diakopoulos, 2016) and adapted from the LLM evaluation toolkit. The reason they are not yet standard is organizational, not technical.

5.3. The Auditor Role

A consequence of the KPI shift is the emergence of a new functional role inside the marketing organization: the *AI visibility auditor*. The role's work is closer to that of an internal auditor than to that of a marketer. It runs prompt audits across foundation models, monitors embedding-probe drift, samples agent task completions, tracks hallucinations against published facts, and files takedowns and corrections through emerging mechanisms—including the AI Overrides registries that some assistants

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 May 2026 doi:10.20944/preprints202605.1208.v1

have begun to expose, and the structured feedback channels that platforms operate at varying levels of opacity.

This role is not a re-skinned SEO position. SEO operates against a known target (search engine algorithms) with a relatively stable measurement infrastructure. The AI visibility auditor operates against multiple non-stationary targets (foundation models that update, retrieval indexes that shift, agent platforms that change manifests) with measurement infrastructure they often have to build in-house. Glikson and Woolley's (2020) review of trust in AI is the closest organizational template: the auditor's job is to keep the brand's trust calibration accurate as the underlying system changes.

6. Ethical, Regulatory, and Governance Implications

6.1. The EU AI Act and the Question of Attribution

The Regulation (EU) 2024/1689, the Artificial Intelligence Act, entered the European legal corpus in July 2024 (European Parliament & Council, 2024). Among its provisions, the obligations on general-purpose AI systems include training-data summary requirements (Art. 53) and transparency obligations that, while not creating an enforceable "right to be cited," do begin to construct an attribution regime around the inputs to AI systems. Hacker, Engel, and Mauer (2023) and Veale and Zuiderveen Borgesius (2021) trace the regulatory architecture and its gaps.

Practitioners should be careful with one piece of trade-press framing: the so-called "right to be cited" is not, at present, an established legal doctrine. The closest analogues are attribution duties under the EU Digital Single Market Directive (Art. 17) for copyrighted material and the AI Act's transparency provisions for training data summaries. A brand seeking to ensure attribution should therefore work through the existing levers—copyright registration, structured `isBasedOn` and `citation` schema markup, and contractual terms with platforms that scrape—rather than wait for a doctrine that does not yet exist.

6.2. Hallucination, Takedown, and the Limits of robots.txt

Hallucinations are not symmetric across brands. Pricing hallucinations, feature hallucinations, and policy hallucinations create real damages—including liability exposure for misrepresentations a user takes to be authoritative. Glikson and Woolley (2020) and Lee and See (2004) provide the trust-calibration framework: when users transfer trust from a firm to an agent that mediates the firm, hallucinations become *firm-attributed* in the user's experience even when they are *agent-produced* in technical reality.

Robots.txt, the long-standing tool for crawl control, has been extended in practice with directives addressing AI training crawlers (e.g., GPTBot, ClaudeBot, CCBot). The directives are not legally binding in most jurisdictions; they are signals that well-behaved crawlers honor. Brands should set a deliberate policy across three crawler classes—search-index crawlers, training crawlers, and inference-time retrieval crawlers—recognizing that the choice for each has different consequences. Blocking all three optimizes for control at the cost of corpus presence; allowing all three optimizes for visibility at the cost of attribution leverage.

The emerging mechanism for correcting *specific* hallucinations is platform-mediated: AI Overrides, knowledge-graph correction submissions, and platform feedback channels. None of these provides the speed and scope of search-engine takedown notices. The practical implication is that *prevention*—getting the canonical facts into the training and retrieval corpora correctly the first time—has a higher payoff than *correction*.

6.3. Algorithm Aversion, Appreciation, and the Consumer's Role

Two consumer responses bound the brand's strategic options. The first is algorithm aversion: people erroneously avoid algorithms after seeing them err (Dietvorst, Simmons, & Massey, 2015), and aversion is stronger for subjective tasks (Castelo, Bos, & Lehmann, 2019). The second is algorithm appreciation: people often prefer algorithmic to human judgment when they have not seen the

 doi:10.20944/preprints202605.1208.v1

algorithm err (Logg, Minson, & Moore, 2019). Resistance is also domain-specific—Longoni, Bonezzi, and Morewedge (2019) show strong resistance to medical AI driven by perceived "uniqueness neglect."

The implication is that the brand's communication about AI mediation matters as much as the mediation itself. Brands that present the AI assistant as a transparent intermediary, with clear escalation paths to human service, capture the upside of algorithm appreciation while limiting the downside of aversion. Brands that hide the mediation—or that pretend the assistant is a brand spokesperson—are exposed to the trust collapse that follows the first visible error.

7. Research Agenda

The framework opens a set of empirical questions that the marketing-science toolkit is well-positioned to address.

The first is *measurement*. Embedding probes for semantic equity need standardized attribute sets, sampling protocols across foundation models, and convergent-validity work against existing brand-equity instruments (Aaker, 1996; Krishnan, 1996; John, Loken, Kim, & Monga, 2006). A brand-equity scale that works in latent space is not a substitute for survey-based instruments, but it is an addition that captures something the survey cannot—the brand's representation in the substrate the AI assistant uses.

The second is *audit infrastructure*. The Pew Research methodology (Pew Research Center, 2025) is a model for consented panel-based audit of AI search behavior, but it does not yet extend to agentic completion. The combined design—track both the AI summary and the agent action—is an open research opportunity. So is the cross-model audit: most published audits sample from a single assistant; the strategic question requires sampling across the population of assistants the user might use.

The third is *information gain operationalization*. Aggarwal et al. (2024) show that specific content properties lift visibility in synthesized responses, but the operational definition of "information gain" remains intuitive rather than formal. A retrieval-theoretic formalization—most plausibly grounded in late-interaction retrieval scoring (Khattab & Zaharia, 2020)—would let brands estimate the information gain of a candidate piece of content before publication.

The fourth is *erasure dynamics*. Erasure is a flow, not a stock. Brands lose attribution gradually as paraphrase chains compound. The dynamics—half-life, recoverability, intervention thresholds—are an open empirical question with direct strategic implications. Reisenbichler et al. (2022) on NLG-supported content marketing offers a methodological starting point.

The fifth is *governance and welfare*. The EU AI Act creates one regime, the U.S. patchwork another, the U.K. AI Safety Institute a third. Comparative work on how attribution outcomes vary across regulatory regimes—and on the welfare consequences of brand-erasure for content production incentives—is needed before the field can give policymakers actionable advice.

8. Conclusion: The Reader Who Is Both Person and Process

The conclusion that practitioners draw most often—*build for agents instead of for people*—is wrong in the same shape that "build for mobile instead of for desktop" was wrong a decade ago. The right conclusion is harder and more interesting. The audience is now plural by default. Every artifact you publish has at least two readers, and the machine reader sets the conditions under which the human reader gets to read at all.

The framework we have developed treats this plurality as the central design constraint. AEO, GEO, and AgO are three layers of the same problem; brand erasure is the failure mode the framework is built to detect and prevent; delegated consumer–AI agency is the theoretical vocabulary that holds the layers together. The propositions we have offered are testable with existing methods. The KPI shifts we have mapped are actionable inside any well-resourced marketing organization. The regulatory environment is real but partial, and brands that wait for it to settle will lose the period during which the substrate is still malleable.

 doi:10.20944/preprints202605.1208.v1

What the brand owns, in the end, is not its place in the SERP, the answer, or the action. What it owns is its representation in the substrate beneath all three. Build the substrate, and the surface forms become a problem of distribution rather than a problem of survival. Skip it, and the brand becomes the kind of thing the assistant can do without—which is the only outcome the framework was written to prevent.

References

Aaker, D. A. (1996). Measuring brand equity across products and markets. *California Management Review*, 38(3), 102–120.

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)*. ACM. https://doi.org/10.1145/3637528.3671900

André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., Huber, J., van Boven, L., Weber, B., & Yang, H. (2018). Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Needs and Solutions*, 5(1–2), 28–37.

Anthropic. (2024, October). *Introducing computer use, a new Claude 3.5 Sonnet, and Claude 3.5 Haiku* [Company announcement].

Bohnet, B., Tran, V. Q., Verga, P., Aharoni, R., Andor, D., Soares, L. B., Ciaramita, M., & Collins, M. (2022). *Attributed question answering: Evaluation and modeling for attributed large language models* [arXiv:2212.08037]. arXiv.

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29.

Brynjolfsson, E. (2022). The Turing Trap: The promise and peril of human-like artificial intelligence. *Daedalus*, 151(2), 272–287.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186.

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809–825.

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1), 24–42.

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56–62.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155–1170.

European Parliament and Council. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union, L 2024/1689, 12 July 2024.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., & Wang, H. (2024). *Retrieval-augmented generation for large language models: A survey* [arXiv:2312.10997]. arXiv.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.

Hacker, P., Engel, A., & Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*.

Hildebrand, C., & Bergner, A. (2021). Conversational robo advisors as surrogates of trust: Onboarding experience, firm perception, and consumer financial decision making. *Journal of the Academy of Marketing Science*, 49(4), 659–676.

Huang, M.-H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49(1), 30–50.

Izacard, G., & Grave, E. (2021). Leveraging passage retrieval with generative models for open-domain question answering. In *Proceedings of EACL 2021*.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360.

John, D. R., Loken, B., Kim, K., & Monga, A. B. (2006). Brand concept maps: A methodology for identifying brand association networks. *Journal of Marketing Research*, 43(4), 549–563.

Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., & Yih, W. (2020). Dense passage retrieval for open-domain question answering. In *Proceedings of EMNLP 2020*.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, 57(1), 1–22.

Khattab, O., & Zaharia, M. (2020). ColBERT: Efficient and effective passage search via contextualized late interaction over BERT. In *Proceedings of SIGIR 2020*.

Krishnan, H. S. (1996). Characteristics of memory associations: A consumer-based brand equity perspective. *International Journal of Research in Marketing*, 13(4), 389–405.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.

Liu, N. F., Zhang, T., & Liang, P. (2023). Evaluating verifiability in generative search engines. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 7001–7025).

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629–650.

Menick, J., Trebacz, M., Mikulik, V., Aslanides, J., Song, F., Chadwick, M., Glaese, M., Young, S., Campbell-Gillingham, L., Irving, G., & McAleese, N. (2022). *Teaching language models to support answers with verified quotes* [arXiv:2203.11147]. arXiv.

Mialon, G., Dessì, R., Lomeli, M., Nalmpantis, C., Pasunuru, R., Raileanu, R., Roziere, B., Schick, T., Dwivedi-Yu, J., Celikyilmaz, A., Grave, E., LeCun, Y., & Scialom, T. (2023). *Augmented language models: A survey* [arXiv:2302.07842]. arXiv.

OpenAI. (2025, January). *Introducing Operator* [Company announcement].

Pew Research Center. (2025, July 22). *Google users are less likely to click on links when an AI summary appears in the results*.

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131–151.

Reisenbichler, M., Reutterer, T., Schweidel, D. A., & Dan, D. (2022). Frontiers: Supporting content marketing with natural language generation. *Marketing Science*, 41(3), 441–452.

Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval*, 3(4), 333–389.

doi:10.20944/preprints202605.1208.v1

Santhanam, K., Khattab, O., Saad-Falcon, J., Potts, C., & Zaharia, M. (2022). ColBERTv2: Effective and efficient retrieval via lightweight late interaction. In *Proceedings of NAACL 2022*.

Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., Cancedda, N., & Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. *Advances in Neural Information Processing Systems*, 36.

Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the draft EU Artificial Intelligence Act. *Computer Law Review International*, 22(4), 97–112.

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J.-R. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6).

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the International Conference on Learning Representations (ICLR)*.