International Journal of Web Applications

DLINE JOURNALS

Print ISSN: 0974-7710
Online ISSN: 0974-7729
IJWA 2026: 18 (4)
https://doi.org/10.6025/ijwa/2026/18/4/197-219

Algorithmic Gatekeepers in the Generative AI Era: Platform Asymmetry in AI-Mediated Business Visibility

Pit Pichappan
Digital Information Research Labs
Chennai. 600017. India
pichappan@dirf.org

ABSTRACT

The rapid proliferation of agentic artificial intelligence (AI) and large language models (LLMs) is fundamentally reshaping digital intermediation. Unlike traditional search engines that rank hyperlinks, generative AI agents synthesize and directly endorse commercial entities, acting as novel algorithmic gatekeepers. However, the structural dynamics of this AI-mediated commercial visibility remain largely unexamined. This study investigates platform asymmetry, industry specialization, and geographical consistency in AI business recommendations. Utilizing the AI Visibility Index, we analyze 79,077 business-level observations across 44 industries and 20 countries, auditing the zero-shot recommendation behaviors of ChatGPT, Claude, Gemini, and Grok. Our findings reveal profound structural asymmetry and severe cross-platform fragmentation. Gemini emerges as the dominant intermediary, capturing approximately 50% of the normalized visibility share and leading in 35 of 44 industries, particularly in local and consumer-facing services. Conversely, Claude demonstrates distinct specialization, dominating nine highly regulated, B2B, and financial sectors. Furthermore, we identify severe epistemic fragmentation, with only 261 business-location category combinations achieving visibility across all four platforms. These findings challenge foundational marketing paradigms, notably Keller's Customer-Based Brand Equity framework, by demonstrating that brand salience is now strictly contingent upon "algorithmic access." We discuss the imperative for Generative Engine Optimization (GEO), the strategic necessity of platform specific visibility management, and the urgent need for regulatory frameworks to address the opaque market shaping power of agentic AI gatekeepers.

Keywords: Generative AI, Algorithmic Gatekeeping, AI Visibility, Large Language Models (LLMs), Generative Engine Optimization (GEO), Digital Intermediation, Brand Salience

Received: 3 June 2026, Revised: 1 September 2026, Accepted: 12 September 2026

Copyright: DLINE

1. Introduction and Background

The rapid evolution of large language models (LLMs) has enabled the emergence of highly specialized systems capable of interacting with other models and digital services autonomously. These systems, commonly referred to as AI agents, represent a defining development in the broader shift toward agentic artificial intelligence. Agentic AI refers to software systems capable of executing actions involving third parties with minimal human

intervention beyond initial configuration, periodic supervision, and task-specific input requirements [1]. While existing scholarship increasingly examines AI agents from technical perspectives, including their architectures, applications, and limitations [2, 3, 4], social science research has primarily concentrated on specific societal risks, including labor displacement [5] and the spread of misinformation [6, 7, 8]. Similarly, the transformative implications of agentic technologies have attracted growing scholarly attention [9]. However, what remains largely absent is a systemic analysis that situates AI agents within the historical trajectory of digital intermediation and examines how these technologies are reshaping the structural relationship between users and online services.

Intermediation has long been a defining feature of digital environments. Across diverse domains such as media, finance, innovation, education, and sustainability, intermediaries function as actors or systems that connect, translate, and shape flows of information, value, and relationships. In its broadest sense, an intermediary is "a person or organization that acts as an agent between other people or things," while digital intermediaries refer to organizations that deliver content from third party providers to users through digital software, channels, and devices [10, 11]. As digital ecosystems have evolved, the nature and influence of intermediaries have become increasingly consequential for both information access and market participation.

Historically, gatekeeping was primarily exercised by traditional media institutions that controlled the production and distribution of information. The emergence of the internet and platform based infrastructures transformed this arrangement by shifting gatekeeping functions toward digital platforms and algorithmic systems responsible for content distribution and visibility. The rise of generative artificial intelligence represents a further stage in this evolution, significantly broadening the scope and influence of gatekeeping mechanisms [12].

This transformation has been conceptualized by Vukanoviæ, who views artificial intelligence as a dialectical force that simultaneously expands opportunities for cultural participation while recentralizing power through algorithmic reintermediation [13]. Contrary to early visions of the internet as a disintermediated space, contemporary digital platforms increasingly operate as computational governors that rank, recommend, and generate content through opaque and proprietary optimization processes [13]. In a related argument, Vukanoviæ contends that the growing concentration of artificial intelligence within the media sector reflects a broader structural recentralization of the post broadcast communication order. While digital disruption fragmented audiences, reduced distribution scarcity, and weakened traditional editorial and advertising authority, platform hyperscalers have leveraged AI as both infrastructure and interface to re-coordinate attention at scale [14].

The central role of algorithms in shaping access and visibility has established platforms as powerful structural gatekeepers. Appel et al. identify this growing gatekeeping function as a defining characteristic of contemporary digital markets [15].The consequences are particularly evident in digital marketing and business visibility. Organic reach on Facebook, for example, has declined from approximately 16 percent in 2012 to less than 2 percent in recent years [16, 17]. Similar reductions have been observed on other platforms, with TikTok business accounts reportedly experiencing significant declines in reach during 2025 [18]]. These developments suggest that brand visibility is increasingly determined by algorithmic systems rather than solely by consumer preferences or marketing effectiveness.

This shift carries important theoretical implications for marketing scholarship. Keller's Customer-Based Brand Equity (CBBE) framework emphasizes brand salience as a foundational element of brand equity [19]. However, the declining accessibility of organic visibility indicates that brand salience is increasingly contingent upon algorithmic access rather than direct consumer engagement alone. Consequently, marketing communications can no longer be understood exclusively through traditional models of consumer behavior, as market visibility itself is increasingly mediated by opaque algorithmic architectures that influence commercial outcomes.

1.1 Algorithmic gatekeeping

Beyond visibility, algorithmic gatekeeping has intensified structural dependencies throughout the digital economy. The concentration of power among dominant platform firms is illustrated by the fact that Alphabet,

Amazon, and Meta collectively account for approximately 55 percent of global advertising expenditure and 72 percent of advertising expenditure within the United States [20]. Within e-commerce ecosystems, algorithmic control is equally pronounced. The Amazon Buy Box algorithm, for instance, is estimated to influence more than 80 percent of sales occurring on the platform [21]. Academic research further demonstrates the market-shaping capacity of algorithms. Lee and Musolff show that algorithmic systems create asymmetric advantages among sellers [22] demonstrate that non-neutral recommendation algorithms can increase price dispersion and reduce competitive intensity [23]. Additional evidence suggests that algorithms actively influence pricing dynamics. Calvano et al. found that algorithmic systems can independently learn collusive pricing strategies [24], while Assad et al. provide empirical evidence that such practices can increase profit margins in real-world markets [25].Collectively, these findings indicate that algorithms not only allocate visibility but also coordinate market behaviors in ways that challenge classical assumptions regarding competition and market efficiency.

The evolution of algorithmic governance is further reinforced by advances in content evaluation and recommendation technologies. Major platforms increasingly employ sophisticated artificial intelligence systems capable of assessing content quality, relevance, and engagement at unprecedented levels of detail. YouTube, for example, has integrated large language models into its content evaluation processes [26, 27], while Instagram increasingly prioritizes content shared through direct messages as a signal for promoting original creators [28, 29]. These developments extend the foundational work of Chen et al. [30] on platforms as meta organizational structures. Contemporary algorithmic systems no longer merely facilitate interactions among market participants; they actively manage content quality, competitive dynamics, and revenue optimization simultaneously across global digital markets.

Taken together, these developments suggest that generative AI and agentic systems represent not simply technological innovations but a new phase in the evolution of digital intermediation. As algorithms become increasingly capable of generating, evaluating, ranking, and acting upon information autonomously, they assume an expanded gatekeeping role that shapes market visibility, competitive outcomes, and user access. Understanding these systems therefore requires moving beyond narrow technical analyses toward a broader examination of the structural power relations embedded within AI-mediated digital ecosystems.

2. Research Problem and Statement

2.1 Research Problem

The digital economy has historically been shaped by algorithmic gatekeepers platforms and search engines that mediate information flows, allocate consumer attention, and dictate market visibility. As digital ecosystems evolve, the emergence of agentic artificial intelligence (AI) and large language models (LLMs) represents a paradigm shift in this intermediation. Unlike traditional search engines that return a ranked list of hyperlinks, generative AI agents synthesize, evaluate, and directly recommend specific commercial entities in conversational, zero-shot environments. While existing scholarship has extensively documented the societal risks of agentic AI (e.g., labor displacement, misinformation) and the structural power of traditional platform algorithms (e.g., the erosion of organic reach, algorithmic pricing collusion), there is a critical empirical gap regarding how these new AI agents function as *commercial intermediaries*. Specifically, it remains largely unexamined whether generative AI platforms act as unified, consistent gateways to commercial information or as fragmented, asymmetrical silos that allocate visibility based on opaque, platform-specific architectures. This lack of transparency challenges foundational marketing frameworks, such as Keller's Customer-Based Brand Equity (CBBE) model, by suggesting that brand salience is no longer solely a function of consumer memory or traditional marketing spend, but is increasingly contingent upon algorithmic access and platform-specific recommendation heuristics.

2.2 Research Statement

This study investigates the structural asymmetry, industry specialization, and geographical consistency of AI-mediated business visibility across prominent generative AI platforms. By deploying a large scale, automated probing framework across 44 industries, 20 countries, and 108 geographic locations, we

systematically audit the recommendation behaviors of four leading AI agents (ChatGPT, Claude, Gemini, and Grok). The objective is to empirically map how agentic AI systems allocate commercial attention, quantify the fragmentation of cross-platform visibility, and establish a foundational understanding of generative AI as the next generation of digital market gatekeepers.

2.3 Research Design and Analytical Architecture

To systematically evaluate AI-mediated commercial intermediation, we designed a multi-layered analytical architecture that bridges empirical data extraction with theoretical synthesis. The research architecture is structured into four sequential layers:

• Layer 1: The Empirical Probing Environment (Data Generation) The foundation of the architecture is a stateless, zero-shot probing framework designed to isolate the parametric knowledge and inherent recommendation tendencies of the AI models. This layer controls for external variables (e.g., web search retrieval, geolocation metadata, conversational history) to ensure that observed visibility differences are attributable to the models' internal architectures rather than contextual artifacts.

• Layer 2: Entity Resolution and Normalization Pipeline Raw, stochastic LLM outputs are inherently noisy. The second layer employs a rigorous computational pipeline to harmonize data. This includes text normalization (stripping legal suffixes, accents, and articles) followed by LLM-assisted semantic clustering to merge variant spellings of the same business entity. A curated whitelist of manual overrides is applied to prioritize precision, ensuring distinct real-world businesses are not erroneously collapsed.

• Layer 3: Multi-Dimensional Analytical Matrix The core analytical engine evaluates the normalized data across four intersecting dimensions:

1. Platform-Level Asymmetry: Measuring mean visibility, non-zero coverage (extensive margin), and normalized relative composition.

2. Industry-Conditioned Dominance: Mapping platform leadership across 44 distinct business categories to identify sector-specific specialization (e.g., B2B/regulated vs. local services).

3. Geographical Consistency: Evaluating the stability of platform dominance across 20 countries and 4 macro-regions, utilising native-language prompting to test for cross-lingual robustness.

4. Cross-Platform Fragmentation: Quantifying the overlap of business visibility across platforms to test the assumption of a unified AI search ecosystem.

• Layer 4: Theoretical Integration and Inference The final layer maps the empirical findings back to theories of digital intermediation, algorithmic gatekeeping, and brand equity. Statistical inference (non-parametric repeated-measures testing) is utilized to validate the significance of observed asymmetries, while strict boundary conditions are applied to separate empirical measurement from platform market share or recommendation quality.

2.4 Experimental Setup

The experimental setup was designed to capture the stochastic nature of large language models while maintaining strict ecological validity across global markets.

2.41 Corpus and Scope

The experimental corpus encompasses 2,580 unique "pages," each representing a specific intersection of a business category and a geographic location. The scope includes 44 distinct business categories (stratified into 25 local trade services, 10 software/online services, and 9 consumer brands) and 108 distinct locations (88 cities and 20 national markets across North America, South America, Europe, and Oceania). To ensure ecological validity, all buyer-intent prompts were natively formulated in the primary language of the target location (spanning 11 languages), avoiding the semantic degradation of machine translation.

2.42 Prompting Protocol and Control Conditions

For each of the 2,580 pages, four distinct buyer-intent questions were posed to the AI assistants. To isolate the models' inherent parametric knowledge, the experimental environment was strictly controlled: all queries were executed in a stateless, zero-shot context. Web search capabilities, geolocation metadata, and prior chat history were explicitly disabled.

2.43 Model Sampling Strategy

Because LLM generation is inherently stochastic, a single query is insufficient to capture a model's true recommendation distribution. Therefore, we employed a multi-run sampling strategy:

- ChatGPT and Gemini: Queried three times each across all 2,580 pages.
- Claude: Queried three times on 825 pages, and once on the remaining 1,755 pages.
- Grok: Evaluated on a targeted subset of 104 pages.
- Retrieval-Augmented Generation (RAG) Baseline: To isolate the impact of live web search, a parallel single-run query with web search enabled was executed for ChatGPT (all pages) and Claude (680 pages).

In the October 2026 edition alone, this design generated 80,082 API calls without web search, yielding 80,059 valid responses. For this study, we concatenated the September and October 2026 editions, resulting in a final analytical sample of N = 79,077 business-level observations.

3. Dataset Description: The AI Visibility Index

The AI Visibility Index is a large-scale, longitudinal dataset designed to quantify the frequency and positioning of business recommendations generated by prominent large language models (LLMs) and conversational AI assistants. Published monthly by AIGROW and archived on Zenodo (DOI: 10.5281/zenodo.23188442 for the October 2026 edition), this dataset provides empirical, cross-sectional, and time-series measurements of "AI visibility." As AI-driven search and conversational agents increasingly mediate consumer discovery, this resource offers researchers a rigorous foundation for studying emergent phenomena such as Generative Engine Optimization (GEO), algorithmic bias, and the shifting dynamics of AI-mediated market competition. The dataset is released under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

The scope of the dataset is both geographically and categorically extensive, encompassing 108 distinct locations across 20 countries, including 88 specific cities and 20 national markets spanning North America, South America, Europe, and Oceania. Within these locations, the dataset tracks recommendations across 44 distinct business categories. These are stratified into 25 local trade services (such as dentists, plumbers, and lawyers) queried at the city level, 10 software and online service categories (such as CRM platforms and website builders) queried at the national level, and 9 consumer brand categories (including banks, mobile carriers, and supermarkets) also queried nationally. To ensure ecological validity, all queries were natively formulated in the primary language of the target location spanning English, German, French, Italian, Spanish, Portuguese, Dutch, Swedish, Danish, Norwegian, and Polish thereby avoiding the semantic artifacts often introduced by literal machine translation.

Data collection was executed through a highly structured, automated probing framework. The experimental design consists of 2,580 unique "pages," each representing a specific business category within a specific location. For each page, four distinct buyer-intent questions were posed to the AI assistants. To isolate the models' inherent parametric knowledge and recommendation tendencies, all queries were submitted in a stateless, zero-shot context, explicitly disabling web search, geolocation metadata, and prior chat history. To account for the inherent stochasticity of LLM generation, queries were sampled multiple times per assistant. ChatGPT and Gemini were queried three times each across all 2,580 pages. Claude was queried three times on 825 pages and once on the remaining 1,755 pages, while Grok was evaluated on a targeted subset of 104 pages.

In total, the October 2026 edition comprises 80,082 API calls without web search, yielding 80,059 valid responses. Additionally, a parallel baseline condition was established by running a single query with web search enabled for ChatGPT (across all pages) and Claude (across 680 pages), allowing researchers to isolate the impact of retrieval-augmented generation (RAG) on visibility metrics.

The primary data file is provided as a UTF-8 encoded, comma-separated values (CSV) document. To maintain data density and reduce noise, each row represents a unique business entity that was mentioned in two or more responses for a specific page; entities mentioned only once are aggregated into page-level totals but omitted from row-level detail. The dataset includes comprehensive identifiers, such as the edition month, category slug, place slug, and the ISO 3166-1 alpha-2 country code. Core visibility metrics include the absolute count of mentions, the total number of answered probes, and the overall mention share, accompanied by Wilson score 95% confidence intervals to denote statistical reliability. The data is further disaggregated to provide assistant-specific mention shares, the number of unique questions that elicited the business name, and the average ordinal rank of the business when it was mentioned. A dedicated column also captures the mention share under the web-search-enabled baseline condition.

To ensure accurate attribution, raw model outputs undergo a rigorous, multi-stage entity resolution pipeline. Initial text normalization strips accents, standardizes capitalization, and removes legal suffixes, leading articles, and bracketed annotations. Subsequently, a secondary language model clusters variant spellings or distinct branch locations that belong to the same underlying business entity. This automated clustering is governed by a curated whitelist of manual overrides, a design choice that intentionally prioritizes precision over recall to prevent the erroneous merging of two distinct, real-world businesses.

Researchers utilizing this dataset should note several important limitations. First, the index reflects model output frequency rather than an endorsement of business quality, operational status, or verified existence; LLM hallucinations may occasionally result in fictitious entities being named. Second, the stateless, location-agnostic querying protocol means that results may diverge from real-world user interactions, which often involve conversational history, explicit geolocation, or authenticated profiles. Third, due to the sampling nature of LLM generation, mention shares are probabilistic estimates, and entities with overlapping 95% confidence intervals should be interpreted as statistically tied in visibility. Finally, because model weights and training corpora are subject to continuous updates, visibility scores are inherently time-bound, and cross-edition comparisons should only be conducted when the querying protocol remains consistent.

Despite these constraints, the AI Visibility Index represents a foundational resource for scholars in computational marketing, human-computer interaction, algorithmic auditing, and the economics of digital search, offering unprecedented transparency into how modern AI systems shape commercial visibility.

4. Methodology

4.1 Data Aggregation and Variable Operationalization

The analysis utilizes the concatenated September and October 2026 editions of the AI Visibility Index, comprising 79,077 business-level observations. Because the dataset does not supply raw, platform-specific recommendation counts, visibility frequency is operationalized using the supplied platform-specific share variables (share_chatgpt, share_claude, share_gemini, share_grok). These variables represent the proportion of a platform's valid responses on a specific page that named a given business. All platform visibility columns are bounded between 0 and 1 and are converted to percentages for reporting.

To assess the relative concentration of visibility, we computed a normalized share metric. This is calculated by summing the mean visibility scores of the four platforms and expressing each platform's mean as a percentage of that total. *Crucially, this normalized metric reflects the relative composition of measured visibility within the experimental framework, not the consumer market share of the AI platforms, nor the absolute share of all AI-generated answers in the wild.*

4.2 Analytical Procedures and Statistical Modeling

The analytical strategy proceeds across platform, industry, and geographical dimensions. Platform-specific business rankings are derived by sorting the mean platform share for each category-place-country-business combination. We explicitly avoid using the dataset's supplied avg_rank variable for platform-specific inferences, as it represents an aggregate, non-platform-specific ordinal field.

To evaluate whether the four platform visibility scores differ systematically, we employed a repeated-measures non-parametric statistical framework. Given that the visibility distributions are boundedproportions and highly zero-inflated, parametric assumptions (e.g., normality) are violated. Therefore, the Friedman test was utilized as the omnibus test to evaluate overall distributional differences across the four platforms. Kendall's coefficient of concordance (W) was reported to assess the magnitude of the platform effect.

For post-hoc pairwise comparisons, we applied the Wilcoxon signed-rank test with a Holm correction for multiple comparisons. Because the large sample size (N=79,077) inherently inflates statistical power—rendering even trivial differences statistically significant ($p < .001$)our interpretation prioritizes practical effect sizes (mean visibility differences in percentage points) over p values alone.

4.3 Boundary Conditions and Analytical Caveats

Several methodological boundary conditions govern the interpretation of this study:

1. Structural Zeros and Sparse Coverage: The Grok platform was queried on a limited subset of pages. Consequently, its zero values largely reflect a lack of measurement rather than a substantive absence of recommendation propensity. Grok's visibility metrics must be interpreted as features of the observed dataset coverage, not as definitive evidence of its underlying commercial behavior.

2. Cross-Platform Consistency Definition: Nonzero visibility across multiple platforms indicates that a business was mentioned, but it does not imply identical recommendation positions, sentiment, or contextual framing.

3. Temporal Limitations: The dataset covers only two consecutive monthly editions. While this permits robust short-term matched comparison, it does not support claims regarding long term trends, seasonality, or model evolution. The findings are strictly framed as a cross-sectional analysis of AI visibility in late 2026.

4. Nested Observations: Businesses are nested within category place pages, meaning observations are not strictly independent. While the Friedman and Wilcoxon tests provide robust within sample evidence of distributional differences, the resulting values should be interpreted as descriptive indicators of structured asymmetry rather than definitive population level inferences. Future research should employ multilevel modeling or clustered bootstrapping to account for this nesting.

5. Analysis

5.1 Comparative Business Visibility Across Four Generative AI Platforms

The analysis of the AI Visibility Index provides a systematic assessment of how commercial entities are surfaced by four prominent generative AI systems: ChatGPT, Claude, Gemini, and Grok. By combining the September and October 2026 editions, the analysis draws on 79,077 business-level observations spanning 44 industries, 20 countries, and 108 geographic locations. The central empirical object is not consumer demand, business quality, or actual market share, but rather AI-mediated visibility: the extent to which a business appears in model generated responses to standardized buyer-intent prompts. This distinction is important because the dataset captures recommendation behavior under controlled prompting conditions rather than the full complexity of real world user interactions.

The results indicate a pronounced asymmetry in platform specific business visibility. Gemini exhibits the highest average visibility across the pooled dataset, followed by ChatGPT and Claude, while Grok has very limited observed representation. These differences are not confined to a small number of sectors or countries.

Rather, Gemini's advantage appears consistently across industry level and country level comparisons, suggesting that the observed pattern is a broad structural feature of the dataset rather than a localized anomaly. However, the interpretation of this dominance must remain measurement-specific: Gemini's higher visibility means that, within the sampled business page observations and under the dataset's probing protocol, businesses receive higher average mention shares from Gemini than from the other platforms. It does not imply that Gemini has a larger user base, greater market share, or superior recommendation quality.

5.2 Initial platform visibility findings

Preliminary dataset results

The average platform specific visibility scores across all business observations indicate a clear difference in platform representation.

Average business visibility by AI platform

Mean of platform-specific visibility scores across 79,077 observations; expressed as percentages.

Platform	Mean visibility	Observations with nonzero visibility
Gemini	28.12%	72.15%
ChatGPT	14.91%	47.32%
Claude	12.79%	37.53%
Grok	0.30%	0.80%

Table 1. Platforms and visibility

Table 1 reports the mean platform-specific visibility scores and the proportion of observations in which each platform records nonzero visibility. The results show that Gemini has a mean visibility score of 28.12%, compared with 14.91% for ChatGPT, 12.79% for Claude, and 0.30% for Grok. In practical terms, Gemini's average visibility is approximately 1.89 times that of ChatGPT and 2.20 times that of Claude. This magnitude indicates that the difference is not merely statistical but substantively meaningful within the measurement framework.

The nonzero visibility rates provide an additional and important dimension of interpretation. Gemini records nonzero visibility in 72.15% of observations, whereas ChatGPT and Claude do so in 47.32% and 37.53% of observations, respectively. This suggests that Gemini's higher mean visibility is not driven solely by a small number of extremely high scores. Rather, Gemini appears more frequently across the dataset and therefore has broader extensive-margin coverage. In contrast, the median visibility scores reported for ChatGPT, Claude, and Grok are zero, while Gemini has a median visibility of 16.7%. This distributional difference is analytically important: for the majority of ChatGPT, Claude, and Grok observations, a given business is not visible at all on that platform, whereas Gemini more often registers positive visibility.

Figure 1 visually reinforces this asymmetry. The figure should be interpreted as showing a two-level structure of platform visibility. Gemini occupies a clearly separate upper tier; ChatGPT and Claude form a second, comparatively similar tier; and Grok remains a marginal category in this dataset. The visual separation between Gemini and the other platforms is useful for communicating the central descriptive finding, but the figure should be accompanied by a clear note that these values are averages of platform specific visibility indicators and not shares of total AI usage or consumer-facing recommendation traffic.

The distinction between mean visibility and nonzero coverage also helps clarify the nature of AI-mediated commercial exposure. A platform may recommend a relatively narrow set of businesses very frequently, or it may distribute mentions across a wider set of entities. Here, Gemini's advantage appears on both dimensions: it has both the highest average visibility and the highest proportion of nonzero observations. This makes Gemini's dominance more robust descriptively than would be the case if its lead were observed only in mean values. Nevertheless, because the dataset lists only businesses mentioned in at least two answers on a page, the analysis should be understood as examining recurring visibility rather than the full universe of all possible business mentions, including rare or singleton recommendations.

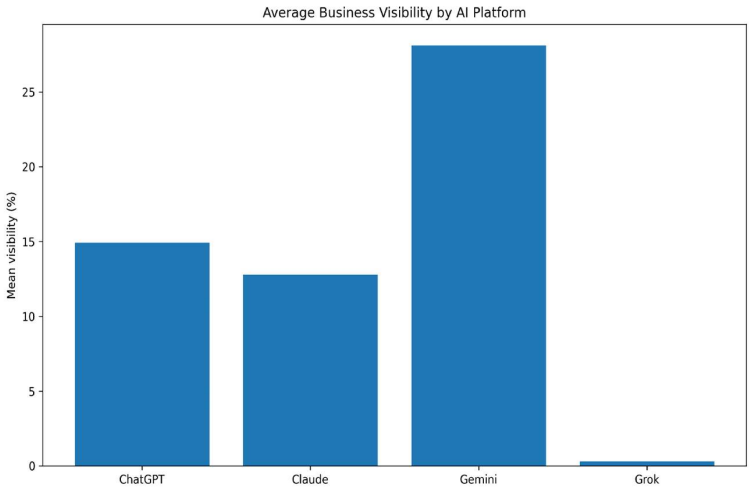

Figure 1. Platform Mean Visibility

5.3 Platform-wise recommendation share

To compare the relative contribution of each platform, we normalized the four average visibility scores to 100%.

5.31 Relative platform visibility composition

Platform	Mean_visibility_pct	Median_visibility_pct	Nonzero_observations	Nonzero_observation_pct	Total_visibility_score_sum	Normalized_share_of_four_platform_mean_pct
ChatGPT	14.90585379	0	37415	47.31464269	11787.102	26.56231877
Claude	12.78525361	0	29673	37.52418529	10110.195	22.78339684
Gemini	28.1226956	16.7	57056	72.152459	22238.584	50.1148083
Grok	0.302735309	0	634	0.801750193	239.394	0.539476093

Table 2. Relative Platform visibility

5.4 Relative Platform Visibility Composition

Table 2 extends the descriptive comparison by normalizing the four mean platform visibility scores to sum to 100%. Under this transformation, Gemini accounts for 50.11% of the combined four-platform mean visibility

score, while ChatGPT contributes 26.56%, Claude 22.78%, and Grok 0.54%. This normalized composition is useful because it expresses the relative distribution of measured visibility across the four platforms in an intuitive form. Figure 2 provides a visual representation of this concentration and shows that approximately half of the measured platform visibility is associated with Gemini.

However, this normalized measure requires particularly careful interpretation. It is not a market-share statistic, nor does it represent the percentage of all AI-generated recommendations received by each platform. Instead, it is a relative composition of the four average visibility scores included in the analysis. The denominator is the sum of the mean visibility scores, not the total number of prompts, users, businesses, or recommendations in the wider AI ecosystem. Therefore, the appropriate interpretation is that Gemini contributes approximately half of the measured four-platform visibility signal in this dataset.

The substantive implication is that visibility is highly concentrated across platforms. Businesses that appear in Gemini-generated recommendations are, on average, more strongly represented in the measured visibility landscape than businesses surfaced by the other systems. For research on generative engine optimization and algorithmic intermediation, this suggests that platforms may differ substantially in their propensity to surface commercial entities, even when responding to comparable buyer-intent prompts. The finding also indicates that a business's AI visibility strategy cannot be treated as platform-neutral. Visibility on one assistant does not necessarily translate into equivalent visibility on another.

At the same time, the normalized composition should not be overinterpreted as evidence of competitive dominance in the AI assistant market. Platform usage, user trust, interface design, subscription models, and integration into search environments are outside the scope of the dataset. The measure captures recommendation visibility under controlled experimental conditions. Its value lies in identifying differences in model output behavior, not in estimating the commercial reach of the platforms themselves.

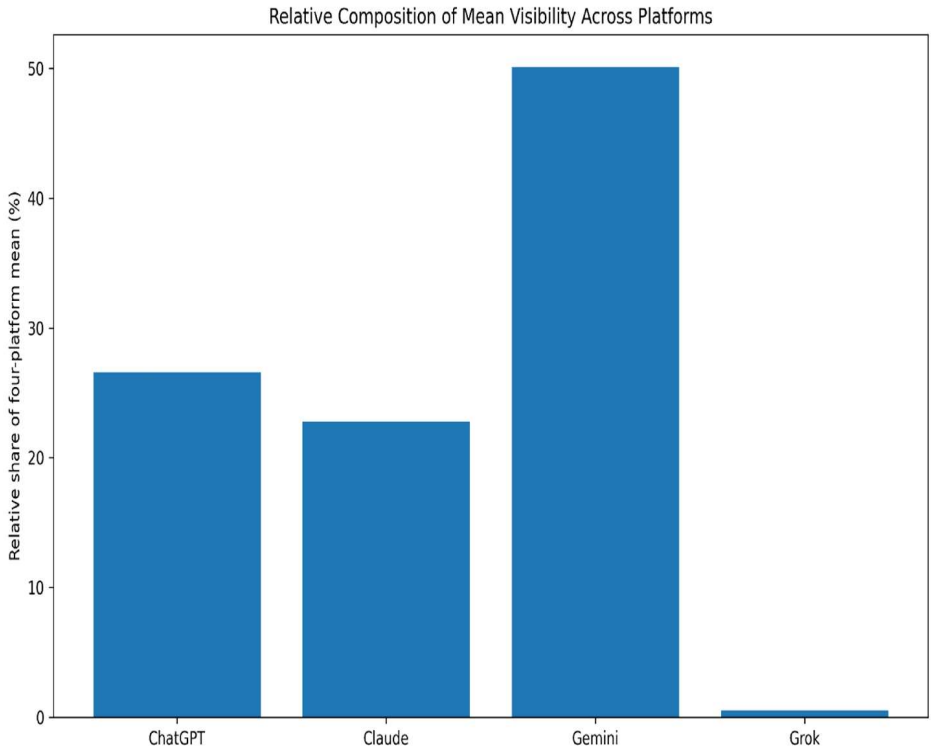

Figure 2. Platform Relative Share

Important distinction: The dataset does not contain separate raw recommendation counts for each AI platform. The calculations therefore use share_chatgpt, share_claude, share_gemini and share_grok, rather than dividing the overall mentions variable into platform-specific counts.

5.5 AI platform dominance by industry

The analysis identifies the platform with the highest mean visibility in each of the 44 industries.
Number of industries dominated by each platform

Industries dominated by Claude:

1. Internet providers.

2. Car insurance.

3. Project management.

4. Online brokers.

5. Mobile carriers.

6. Accounting software.

7. Banks.

8. Email marketing.

9. Website builders.

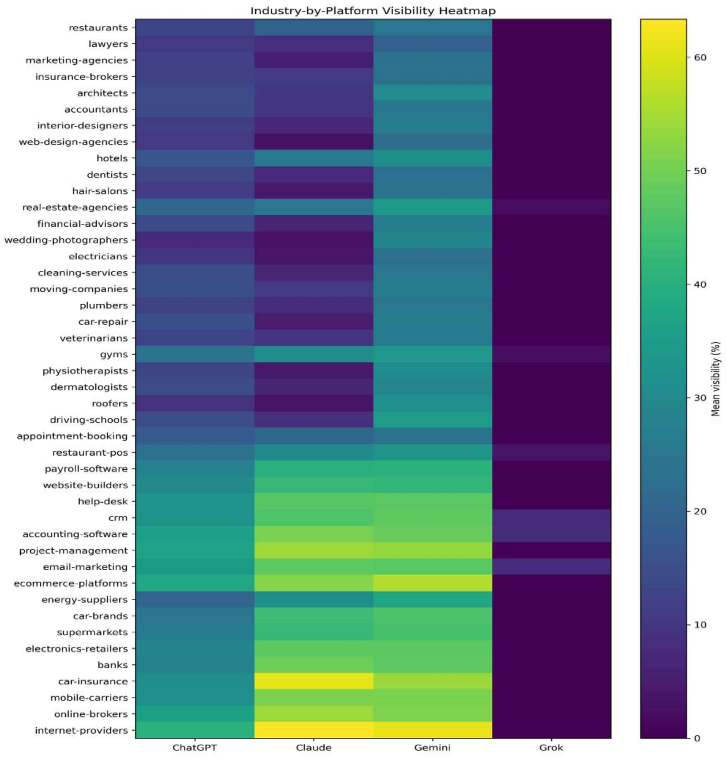

Figure 3. Industry Platform Heatmap

5.6 Industry-Level Platform Dominance

The industry-level analysis provides a more nuanced picture of platform visibility. Although Gemini leads overall, its dominance is not uniform across all commercial sectors. Figure 3 shows that Gemini has the highest mean visibility in 35 of the 44 industries, corresponding to 79.55% of the categories examined. Claude leads in the remaining nine industries, while ChatGPT and Grok do not lead any industry under the mean-visibility criterion.

This pattern suggests that Gemini has broad cross-sector coverage, particularly in local service and consumer-facing categories. The dataset indicates especially strong Gemini visibility in sectors such as restaurants, hotels, dentists, lawyers, gyms, real estate agencies, and other professional services. These categories typically involve location sensitive consumer choice, routine service discovery, and high frequency local search behavior. Gemini's strong performance in such sectors may indicate that its generated responses more readily surface named commercial entities when users express purchase or service seeking intent. However, the dataset alone cannot determine whether this pattern arises from training data composition, model alignment policies, response formatting tendencies, retrieval behavior, or other architectural and operational factors.

Claude's leadership in nine industries is equally important because it qualifies the overall Gemini dominance narrative. The industries in which Claude leads include internet providers, car insurance, project management software, online brokers, mobile carriers, accounting software, banks, email marketing, and website builders. These sectors are not random. Many are information dense, comparison oriented, regulated, financial, telecommunications-related, or business to business service categories. This suggests that Claude may have comparatively stronger visibility patterns in domains where recommendations involve structured comparison, institutional familiarity, or specialized product knowledge.

The industry level findings therefore support a interpretation of *category-conditioned platform visibility*. The platforms do not appear to behave as interchangeable recommendation engines. Instead, their visibility profiles vary by sector, indicating that generative AI systems may organize commercial knowledge differently across domains. For businesses, this implies that platform specific optimization may be especially important in sectors where the overall leader is not necessarily the dominant platform. For researchers, the finding raises important questions about how model training corpora, safety policies, commercial knowledge representations, and answer-generation strategies shape sector specific recommendation patterns.

Figure 3 should therefore be interpreted not simply as a dominance chart but as evidence of heterogeneity in AI-mediated market visibility. Gemini's broad lead indicates general prominence, but Claude's sector-specific strength demonstrates that platform leadership is conditional. A high level journal interpretation should emphasize this interaction between platform and industry rather than treating overall mean visibility as a complete summary of platform behavior.

5.7 Platform preference by geographical region

The country level analysis shows that Gemini has the highest average visibility in all 20 countries.

The broad geographical comparison is:

Region	ChatGPT	Claude	Gemini	Grok
Europe	14.69%	12.28%	28.36%	0.30%
North America	15.27%	13.97%	26.94%	0.43%
Oceania	15.90%	14.76%	27.28%	0.00%
Latin America	15.54%	12.55%	32.29%	0.00%

Table 3. Regional Variation

The geographical analysis examines whether platform visibility differences are consistent across regions and countries. Table 3 reports mean platform visibility across four broad regional groupings: Europe, North America, Oceania, and Latin America. Gemini records the highest mean visibility in all four groups. Its regional mean visibility is 28.36% in Europe, 26.94% in North America, 27.28% in Oceania, and 32.29% in the Latin American grouping, which contains Brazil only in this dataset. ChatGPT and Claude follow a similar ordering across regions, while Grok remains minimally represented.

The regional results indicate that Gemini's lead is not confined to a single geographic context. Instead, it persists across diverse linguistic and national markets. This is an important finding because the dataset's prompts were formulated in the language of each place rather than translated mechanically. The consistency of Gemini's lead across multilingual settings suggests that its visibility advantage is not simply an artifact of English-language prompting. However, the results should not be interpreted as evidence that geography is irrelevant. Rather, they show that the rank ordering of platforms is geographically stable even though the magnitude of visibility varies by region.

Table 4 provides a more granular country level view and confirms that Gemini is the dominant platform in all 20 countries examined. The highest Gemini mean visibility is observed in Brazil, followed by Norway, Portugal, Mexico, and Ireland. The United States records the lowest Gemini mean visibility among the listed countries, although Gemini still leads all other platforms there. This pattern suggests that the magnitude of Gemini's advantage varies across national contexts, even when its leading position remains unchanged.

The ChatGPT–Claude relationship shows more geographical nuance. In many countries, ChatGPT exceeds Claude, but the gap is often modest. In Sweden, Denmark, Ireland, and Norway, Claude records visibility levels that are very close to or slightly above ChatGPT. For example, in Norway, Claude's mean visibility is 18.98%, compared with ChatGPT's 18.17%. These cases indicate that the second tier ranking is less stable than the first-tier Gemini lead. They also suggest that local language, market structure, or category composition may influence the relative visibility of ChatGPT and Claude.

Figure 4 should be interpreted as a visualization of cross national consistency with localized variation. The dominant visual pattern is Gemini's repeated lead across countries, but the figure also allows readers to see differences in the size of the gap between platforms. A country such as Brazil shows especially high Gemini visibility, whereas the United States shows a more compressed distribution among the leading platforms. Such differences may be substantively relevant for international businesses seeking to understand AI-mediated discovery across markets.

Two cautions are necessary. First, the regional groupings are analytical constructions derived from country codes and do not necessarily reflect standard geopolitical classifications. In particular, the North America grouping includes the United States, Canada, and Mexico, while the Latin America grouping contains only Brazil. Therefore, the Latin America estimate should be described specifically as Brazil-based rather than generalized to the broader Latin American region. Second, country-level averages may reflect differences in the composition of industries and locations represented in the dataset. Without additional modeling, it is not possible to separate pure country effects from category composition effects.

country	share_chatgpt	share_claude	share_gemini	share_grok	n_observations	dominant_platform
US	14.88640292	13.74425773	25.52064529	0.696821586	10414	Gemini
GB	15.69538759	14.10136884	27.05342955	0.815890492	6721	Gemini
DE	13.04014145	9.355028463	26.77954114	0.619544592	5797	Gemini

FR	14.52298236	10.05018707	27.81345092	0	5613	Gemini
IT	12.96355157	9.633924653	28.63200439	0.594257498	5468	Gemini
CH	13.57681307	11.14008172	27.98439224	0	4895	Gemini
ES	14.99390722	10.84855758	29.36473575	1.005815832	4333	Gemini
AU	15.89934778	14.75457722	27.28171442	0	4293	Gemini
CA	16.31969027	15.36752704	27.96733038	0	4068	Gemini
NL	13.9819952	10.46140304	26.21539077	0	3749	Gemini
BE	13.78016578	11.46823564	28.27255773	0	3378	Gemini
PL	13.92377173	11.50661376	28.98805745	0	2646	Gemini
SE	16.56948126	17.16228701	29.55221507	0	2641	Gemini
BR	15.53906433	12.54873294	32.29138402	0	2565	Gemini
AT	14.60123309	13.09427208	28.71682578	0	2514	Gemini
MX	15.15831631	12.62525541	31.27977115	0	2447	Gemini
IE	17.69348791	17.71006413	30.51583621	0	2027	Gemini

Table 4. Geographic variation

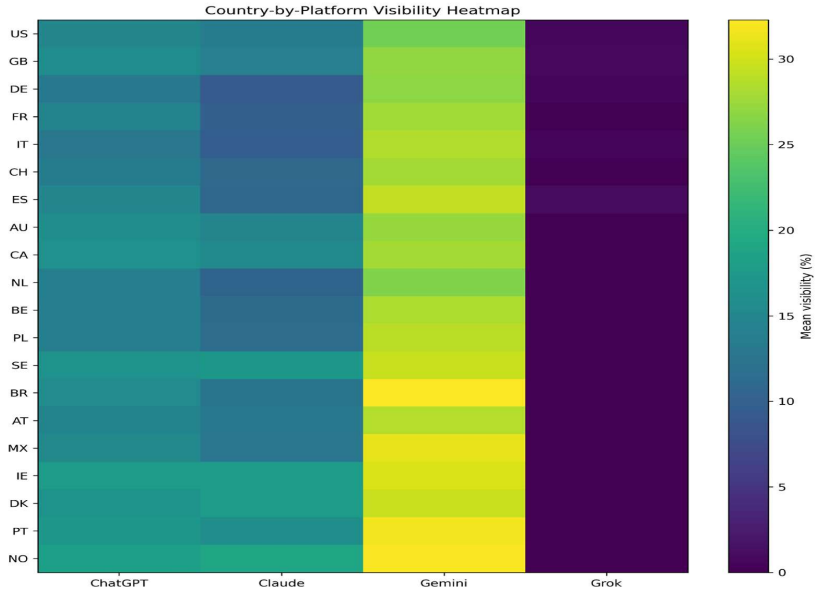

Figure 4. Country Platform

The geographical findings indicate:

- Gemini maintains its lead across all four regional groupings.
- ChatGPT and Claude have relatively similar visibility in most regions.
- Grok has very limited representation.
- Brazil has the highest Gemini visibility among the broad regional groups examined.

The consistency of Gemini's lead across countries is an important finding, although it should not be interpreted as evidence that geography has no effect on visibility.

5.8 Statistical differences between AI platforms

A repeated-measures statistical analysis was conducted because all four platform scores are recorded for the same business observations.

Statistical measure	**Results**
Observations	79,077
Friedman X^2	74,400.80
Degrees of freedom	3
p-value	p less than 0.001
Kendall's W	0.314

Table 5. Statistical results

The repeated-measures statistical analysis evaluates whether the four platform visibility scores differ systematically when measured on the same business observations. The Friedman test is appropriate as a non-parametric omnibus test because the visibility distributions are bounded, highly zero-inflated, and unlikely to satisfy normality assumptions. The test yields a Friedman chi-square statistic of 74,400.80 with three degrees of freedom and a p-value below .001. This indicates that the null hypothesis of equivalent visibility distributions across the four platforms is rejected.

Kendall's coefficient of concordance is 0.314. This suggests a meaningful, moderate overall platform effect. The result is important because it indicates that the observed differences are not solely a consequence of the large sample size. Nevertheless, the very large number of observations means that even small differences can achieve statistical significance. Therefore, the substantive interpretation should focus on the magnitude and direction of differences rather than on p-values alone.

The pairwise Wilcoxon signed-rank tests with Holm correction show statistically significant differences for all six platform pairs. The largest mean difference is between Gemini and Grok, at 27.82 percentage points. The next largest differences are between Claude and Gemini, at 15.34 percentage points, and between ChatGPT and Gemini, at 13.22 percentage points. These are substantial gaps within the measurement framework and reinforce the descriptive conclusion that Gemini occupies a distinct visibility tier.

By contrast, the difference between ChatGPT and Claude is much smaller, at 2.12 percentage points. Although this difference is statistically significant after correction, its practical magnitude is more modest. This distinction

is central for journal-level interpretation. The analysis does not simply show that "all platforms differ." Rather, it shows that the platform ecosystem is structured by a large separation between Gemini and the other major platforms, a smaller but measurable separation between ChatGPT and Claude, and a very large gap between Grok and the rest of the field as represented in this dataset.

The Grok comparisons require particular caution. Grok was queried on only a small subset of pages, and its zero values on other pages reflect lack of measurement rather than necessarily absence of recommendation propensity. Consequently, the low Grok mean and the large Gemini Grok difference should be interpreted as features of the observed dataset coverage, not as definitive evidence of Grok's underlying commercial recommendation behavior. A more conservative interpretation would state that Grok has extremely limited observed visibility in the available sample.

For a high-level journal paper, the inferential results should therefore be presented as supportive rather than definitive. The statistical tests confirm that platform visibility distributions differ, but the main contribution lies in the structured descriptive pattern: Gemini's broad lead, Claude's sector-specific strength, the moderate ChatGPT–Claude difference, and Grok's sparse coverage. The p-values establish that the differences are unlikely to be random under the test assumptions; the effect magnitudes explain why the differences matter.

5.9 Industry-specific AI recommendation patterns

The industry-level results reveal three important patterns.

Pattern 1: Gemini has broad industry coverage.

Gemini leads in 35 industries, suggesting that its business visibility is relatively broad across the dataset.

Pattern 2: Claude has stronger visibility in selected specialized industries.

Claude leads in nine industries, particularly in areas such as financial services, insurance, telecommunications and business software.

Pattern 3: Platform dominance does not mean universal business dominance.

A business with strong Gemini visibility may still have stronger or comparable visibility on Claude or ChatGPT. This is why platform-level and business-level analysis should be interpreted separately.

The industry comparison spreadsheet includes the mean visibility of all four platforms for every industry, the number of observations and the dominant platform.

5.10 Cross-Platform Consistency and Fragmentation of AI Visibility

One of the most consequential findings is the limited cross-platform consistency observed in the dataset. Only 261 business-location-category combinations exhibit nonzero visibility across all four platforms. Relative to the full dataset, this represents a very small degree of overlap. This suggests that AI-mediated business visibility is fragmented: a business visible on one platform is not necessarily visible on others.

This finding has important theoretical and practical implications. Traditional search engine optimization has often been framed around improving visibility within a relatively unified search environment, even though search engines differ in ranking algorithms. Generative AI recommendation environments appear more fragmented because each assistant may produce different named entities, different recommendation lists, and different degrees of commercial specificity in response to similar buyer intent prompts. As a result, AI visibility may be better understood as a platform-specific asset than as a general property of a business.

However, the cross-platform overlap statistic must be interpreted cautiously because Grok's limited coverage mechanically constrains the possibility of all-four-platform visibility. Since Grok was not queried across the

full set of pages, the number of businesses visible on all four platforms is partly a function of research design. Therefore, the 261 overlapping cases should not be treated as a general estimate of cross-platform consistency across the entire AI assistant ecosystem. A stronger future analysis would compute pairwise overlap among platforms with comparable coverage and separately examine overlap within the subset of pages where Grok was actually queried.

Despite this limitation, the observed fragmentation remains analytically important. It suggests that businesses, policymakers, and researchers should avoid assuming that AI systems converge on the same commercial recommendations. Different platforms may create different visibility opportunity structures, and a business that benefits from one model's recommendation pattern may remain absent from another's outputs. This has implications for digital competition, consumer information exposure, and the emerging practice of generative engine optimization.

6. Research interpretation and conclusions

The findings contribute to an emerging understanding of generative AI systems as commercial intermediaries. Search engines and digital platforms have long influenced which businesses become visible to consumers. Generative AI assistants extend this intermediary role into conversational environments, where users may receive a short list of recommended businesses rather than a ranked page of search results. In such settings, being named by the model may constitute a consequential form of algorithmic attention.

The observed Gemini advantage suggests that AI visibility is unevenly distributed across platforms. This unevenness may shape how businesses benefit from conversational discovery. If one platform is more likely to surface named businesses across many categories and countries, then visibility on that platform may become strategically important for firms seeking exposure in AI-mediated markets. However, the results also show that platform advantage is not universal. Claude's stronger position in selected industries indicates that sectoral context moderates platform visibility. This supports a more differentiated theory of AI-mediated recommendation in which platform, industry, and geography interact to shape commercial discoverability.

The findings also raise questions about the mechanisms underlying platform-specific visibility. The dataset does not reveal whether differences arise from training data, reinforcement learning from human feedback, safety policies, model architecture, answer length, retrieval integration, or post-processing rules. Therefore, causal explanations should be avoided. Nevertheless, the patterns are consistent with the possibility that generative AI platforms encode and reproduce different structures of commercial knowledge. Some platforms may be more inclined to provide concrete business names, while others may produce more generic advice, cautionary language, or category-level responses. These differences matter because they influence whether businesses are named at all and how commercial attention is allocated.

From a marketing perspective, the results suggest that generative engine optimization should not be conceptualized as a single uniform practice. Instead, businesses may need platform-specific visibility strategies. A restaurant, dental clinic, or law firm may face a different AI visibility environment from a bank, insurance provider, or software company. Similarly, a business operating across multiple countries may encounter different degrees of platform concentration even when the leading platform remains the same. These findings point toward a research agenda on how firms can understand, monitor, and ethically improve their visibility in AI-generated recommendations.

From a public-interest perspective, the results also raise concerns about transparency and accountability. If conversational AI systems increasingly mediate consumer choice, then the criteria by which businesses are recommended become economically and socially significant. The dataset records what models say, not whether the recommendations are accurate, current, unbiased, or commercially appropriate. Because some recommended entities may be outdated, poorly matched to user needs, or hallucinated, AI visibility should not be equated with quality or trustworthiness. This distinction is essential for avoiding the implicit endorsement of businesses merely because they appear frequently in model outputs.

7. Methodological Interpretation and Analytical Caveats

First, the dataset covers only two monthly editions. Although this permits short-term matched comparison, it does not support claims about long-term trends, seasonality, or model evolution. The findings should therefore be framed as a cross-sectional and short-horizon analysis of AI visibility in September and October 2026.

Second, the analysis uses platform-specific share variables rather than raw platform-level recommendation counts. The dataset does not provide separate raw counts of recommendations for each assistant. Therefore, the analysis appropriately relies on share_chatgpt, share_claude, share_gemini, and share_grok. These variables measure visibility within each platform's answers and should not be reconstructed by dividing the overall mentions variable across platforms.

Third, the supplied avg_rank variable is not platform specific. It represents an overall rank field and cannot be used to infer business ranking within individual assistants. The platform specific rankings used in this analysis are therefore derived from mean platform visibility scores. This is a valid proxy for relative visibility but should not be confused with actual ordinal placement inside each model's response.

Fourth, zero values require careful interpretation. For ChatGPT, Claude, and Gemini, a zero generally indicates that the business was not observed in that platform's answers within the relevant page-level sampling. For Grok, however, many zeros arise because the platform was not queried on most pages. These are structurally different forms of zero. Treating all Grok zeros as substantive absence would overstate conclusions about Grok's recommendation behavior. The current analysis can describe Grok's observed representation in the dataset, but stronger claims require restricting comparisons to pages where Grok was actually evaluated.

Fifth, the observations are not necessarily independent in a strict statistical sense. Businesses are nested within category-place pages, and the same or similar entities may appear across editions. Therefore, although the Friedman and Wilcoxon tests provide useful within-sample evidence of distributional differences, their p-values should be interpreted cautiously. In a journal submission, the statistical results would be strengthened by robustness checks using clustered inference, bootstrap procedures at the page level, or multilevel models that account for the nesting of businesses within categories, places, and editions.

Finally, the visibility measures are bounded proportions with many zeros. Future work could model visibility using a two-part framework: first estimating the probability of nonzero visibility, and then modeling positive visibility shares conditional on being mentioned. Such an approach would distinguish between a platform's tendency to mention a business at all and the intensity of visibility once the business appears. This distinction is especially relevant here because Gemini differs from the other platforms both in nonzero coverage and in average visibility.

8. Overall Interpretation

Comprehensively, the analysis demonstrates that AI-mediated business visibility is highly uneven across platforms, moderately differentiated across industries, and broadly consistent across countries. Gemini emerges as the most visible platform in the measured dataset, leading in overall mean visibility, normalized visibility composition, most industries, and all countries examined. ChatGPT and Claude occupy comparatively similar intermediate positions, although Claude shows stronger relative visibility in selected specialized sectors. Grok has very limited observed representation, but this finding is strongly conditioned by its restricted coverage in the dataset.

The principal contribution of the analysis is to show that generative AI platforms do not function as interchangeable gateways to commercial information. They produce distinct visibility regimes, and these regimes vary by sector and market. For businesses, this implies that AI visibility must be monitored platform by platform. For researchers, it suggests that generative AI systems are becoming important infrastructures

of commercial discovery whose recommendation patterns warrant systematic empirical scrutiny. For policymakers and platform designers, the findings highlight the need for transparency around how AI systems generate, rank, and frame business recommendations.

The results should therefore be interpreted as evidence of structured platform asymmetry in AI-generated business recommendations, not as a definitive ranking of platform quality or market power. Within that boundary, the AI Visibility Index provides a valuable empirical foundation for studying how generative AI systems allocate commercial attention in increasingly consequential conversational search environments.

9. Discussion

The empirical analysis of the AI Visibility Index reveals a fundamental shift in the architecture of digital intermediation. Generative AI platforms do not function as monolithic, interchangeable gateways to commercial information; rather, they operate as distinct, asymmetrical gatekeepers that allocate visibility based on opaque, platform-specific optimization processes. By mapping the recommendation behaviors of ChatGPT, Claude, Gemini, and Grok across 44 industries and 20 countries, this study provides the first large-scale evidence of how agentic AI systems are restructuring commercial discoverability.

9.1 Theoretical Contributions: Redefining Algorithmic Gatekeeping and Brand Equity

Our findings extend the literature on digital intermediation and algorithmic gatekeeping (Appel et al., 2026; Vukanoviæ, 2025). Historically, platform algorithms were shown to suppress organic reach and create structural dependencies (e.g., the Amazon Buy Box). However, generative AI represents a qualitative leap: it transitions the platform's role from *distributing* ranked links to *synthesizing* and *endorsing* specific commercial entities This shift poses a direct challenge to foundational marketing theories, notably Keller's (1993) Customer-Based Brand Equity (CBBE) framework. Keller posits that brand salience the foundational layer of the equity pyramid—relies on the frequency and ease with which a brand comes to mind during consumer purchase situations. Our findings demonstrate that in the agentic AI era, brand salience is no longer solely a function of consumer memory or traditional marketing spend; it is strictly contingent upon algorithmic access. If an AI agent fails to surface a brand in a zero-shot, conversational query, that brand effectively lacks salience in that specific decision-making context. Consequently, marketing theory must evolve to account for "algorithmic salience," where the power of a brand resides as much in the parametric knowledge and retrieval heuristics of LLMs as it does in the minds of consumers.

Furthermore, our discovery of severe cross platform fragmentation (with only 261 business-location-category combinations visible across all four platforms) contradicts the assumption of a unified AI search ecosystem. Unlike traditional search engines, which share a broadly similar underlying logic of indexing and ranking web pages, LLMs appear to encode fundamentally different structures of commercial knowledge. This epistemic fragmentation suggests that AI platforms are not merely reflecting the open web, but are actively curating distinct, proprietary "commercial realities."

9.2 Category-Conditioned Intermediation and Platform Specialization

A critical contribution of this study is the identification of *category-conditioned platform visibility*. While Gemini exhibits broad, cross-sector dominance (leading in 35 of 44 industries, particularly in local services and consumer facing categories), Claude demonstrates distinct specialization in nine highly regulated, information dense, or B2B sectors (e.g., banking, insurance, accounting software).

This divergence indicates that AI platforms are not neutral arbiters of commercial quality. Instead, their visibility profiles are heavily moderated by their underlying training corpora, alignment policies, and safety guardrails. Platforms like Claude, which are often fine-tuned for high-stakes, analytical, or enterprise tasks, appear to exhibit greater "commercial caution," favoring institutional familiarity and structured comparison over local, consumer facing entities. Conversely, Gemini's architecture or retrieval mechanisms appear more permissive in surfacing concrete, localized business names. This finding necessitates a more nuanced theory

of AI-mediated recommendation, wherein platform advantage is deeply conditional on the sectoral context of the query.

9.3 Managerial Implications:
The Imperative of Generative Engine Optimization (GEO)

For marketing practitioners, the erosion of organic reach on traditional social platforms and the rise of fragmented AI gatekeepers necessitate a paradigm shift from Search Engine Optimization (SEO) to Generative Engine Optimization (GEO). The assumption that a strong SEO strategy will automatically yield high AI visibility is empirically unsupported by our findings.

Because AI visibility is highly platform and industry specific, businesses must adopt bespoke, platform-aware visibility strategies. A dental clinic or restaurant must optimize for the broad, consumer facing recommendation heuristics of platforms like Gemini, whereas a B2B software provider or financial institution must tailor its digital foot print to satisfy the specialized, risk averse evaluation criteria of platforms like Claude. Furthermore, managers must recognize that AI visibility is a probabilistic asset. The high variance in LLM outputs means that brand monitoring in the AI era requires continuous, automated auditing rather than static ranking checks.

Crucially, managers must also decouple AI visibility from brand quality. The dataset captures model output frequency, not operational status or consumer satisfaction. Because LLMs are prone to hallucinations and may recommend outdated or fictitious entities, appearing in an AI response does not constitute an endorsement. Brands must therefore develop active reputation management protocols to correct AI hallucinations and ensure their digital footprints provide accurate, structured data that LLMs can reliably parse.

9.4 Public Policy and Societal Implications

The concentration of AI-mediated commercial discovery raises profound public interest concerns. As conversational AI agents increasingly replace traditional search interfaces, the criteria by which businesses are recommended become economically and socially consequential. The opacity of these recommendation engines creates a new digital divide: businesses that understand how to signal their relevance to LLMs will thrive, while those that do not may face algorithmic obscurity.

Policymakers and antitrust regulators must recognize that algorithmic gatekeeping in the generative AI era extends beyond pricing coordination (as noted by Calvano et al. and Assad et al.) to include the unilateral control of market access. If a dominant AI platform systematically favors its own affiliated services or disproportionately elevates specific market incumbents through its parametric weights, it could stifle competition and reduce consumer choice. The findings of this study underscore the urgent need for regulatory frameworks that mandate transparency in how AI systems generate, rank, and frame commercial recommendations, ensuring that the new gatekeepers of the digital economy remain accountable to market fairness and consumer welfare.

10. Conclusion

The transition from the search engine era to the agentic AI era represents a defining moment in the evolution of digital markets. This study utilized the AI Visibility Index to conduct a large scale, cross national audit of how four prominent generative AI platforms, ChatGPT, Claude, Gemini, and Grok mediate business visibility. By analyzing 79,077 observations across 44 industries and 20 countries, we established that AI-mediated commercial discovery is characterized by profound platform asymmetry, industry specialization, and severe cross platform fragmentation.

Gemini emerged as the dominant commercial intermediary in the measured dataset, leading in overall visibility, industry coverage, and geographical reach. However, the nuanced dominance of Claude in specialized, regulated sectors, combined with the extreme fragmentation of cross-platform recommendations, proves

that generative AI systems do not operate as a unified ecosystem. Instead, they function as distinct algorithmic gatekeepers that encode and reproduce divergent structures of commercial knowledge.

While this study provides a foundational empirical mapping of AI-mediated business visibility, it also opens critical avenues for future research. The current analysis is constrained by its short term, cross-sectional design. Future scholarship must employ longitudinal designs to track how model updates, reinforcement learning from human feedback (RLHF), and shifts in retrieval-augmented generation (RAG) alter commercial visibility over time. Methodologically, future work should advance beyond descriptive asymmetry testing by employing two part econometric models to disentangle a platform's probability of mentioning a business from the intensity of that mention, alongside multilevel modeling to account for the nested nature of businesses within geographic and categorical contexts.

Ultimately, this research demonstrates that generative AI is not merely a new channel for marketing communications; it is a new infrastructure of commercial intermediation. As AI agents increasingly act as the primary interface between consumer intent and market supply, understanding the structural power relations embedded within these algorithmic gatekeepers is no longer optional, it is essential for the future of competitive strategy, digital marketing, and market regulation.

References

[1] Rondina, M., Vetrò, A., De Martin, J. C., Bargis, M., Elia, G. (2026). AI agents and the progressive intermediation between users and online services: A critical analysis. *Future Internet*, *18*(8), 419. https://doi.org/10.3390/fi18080419.

[2] Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., et al. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, *18*, 186345.

[3] Abou Ali, M., Dornaika, F., Charafeddine, J. (2025). Agentic AI: A comprehensive survey of architectures, applications, and future directions. *Artificial Intelligence Review*, *59*, 11.

[4] Bandi, A., Kongari, B., Naguru, R., Pasnoor, S., Vilipala, S. V. (2025). The rise of agentic AI: A review of definitions, frameworks, architectures, applications, evaluation metrics, and challenges. *Future Internet*, *17*, 404.

[5] Acemoglu, D., Restrepo, P. (2018). *Artificial intelligence, automation and work* (Working Paper). National Bureau of Economic Research.

[6] Lin, S., Hilton, J., Evans, O. (2022). TruthfulQA: Measuring how models mimic human falsehoods. In S. Muresan, P. Nakov, A. Villavicencio (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (p. 3214–3252). Association for Computational Linguistics.

[7] Capraro, V., Lentsch, A., Acemoglu, D., Akgun, S., Akhmedova, A., Bilancini, E., Bonnefon, J. F., Brañas-Garza, P., Butera, L., Douglas, K. M., et al. (2024). The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS Nexus*, *3*, 191.

[8] Bontridder, N., Poullet, Y. (2021). The role of artificial intelligence in disinformation. *Data & Policy*, *3*, e32.

[9] Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., et al. (2025). The rise and potential of large language model based agents: A survey. *Science China Information Sciences*, *68*, 121101.

[10] Foster, R. (2012). *News plurality in a digital world* (Technical Report). Reuters Institute for the Study of Journalism.

[11] González-Tosat, C., Sádaba-Chalezquer, C. (2021). Digital intermediaries: More than new actors on a crowded media stage. *Journal of Media*, *2*, 77–99.

[12] Zwingmann, N., Hoffmeister, E. I., Engert, S. P., Hess, T. (2025). *Shifting the gate–Generative AI's impact on media gatekeeping*.

[13] Vukanoviæ, Z. (2026). AI recentralization and reintermediation in media platforms, cloud infrastructure, and intermediaries. In *AI and the recentralization of the global media economy*. Palgrave Macmillan. https://doi.org/10.1007/978-3-032-28082-4_26.

[14] Vukanoviæ, Z. (2026). The transformational impact of artificial intelligence on media market concentration. In *AI and the recentralization of the global media economy*. Palgrave Macmillan. https://doi.org/10.1007/978-3-032-28082-4_25.

[15] Appel, G., Grewal, L., Hadi, R., Stephen, A. T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science*, *48*(1), 79–95.

[16] Marketing Scoop. (2024). *Facebook organic reach statistics* https://marketingscoop.com.

[17] Social Cat. (2025). *Nano-influencer marketing statistics* https://thesocialcat.com.

[18] Napolify. (2025). *TikTok organic reach decline report* https://napolify.com.

[19] Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, *57*(1), 1–22. https://doi.org/10.1177/002224299305700101.

[20] WARC. (2025). *Alphabet, Meta and Amazon to take half of $1 trillion ad market in 2025*. https://www.warc.com.

[31] Amazon Sellers Lawyer. (2025). *Amazon Buy Box statistics and strategies 2025*. https://www.amazonsellerslawyer.com

[22] Lee, K. H., Musolff, L. (2024). *Entry into two-sided markets shaped by platform-guided search* [Working Paper]. Yale/Princeton.

[23] Libai, B., Babic Rosario, A., Beichert, M., Donkers, B., Haenlein, M., Hofstetter, R., Zhang, L. (2025). Influencer marketing unlocked: Understanding the value chains driving the creator economy. *Journal of the Academy of Marketing Science*, *53*(1), 4–28.

[24] Calvano, E., Calzolari, G., Denicolò, V., Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, *110*(10), 3267–3297.

[25] Assad, S., et al. (2024). Algorithmic pricing and competition: Evidence from the German retail gasoline market. *Journal of Political Economy*, *132*(3), 723–771.

[26] Buffer. (2025). *How the YouTube algorithm works in 2025*. https://buffer.com/resources/youtube-algorithm/.

[27] Sprout Social. (2025). *YouTube algorithm guide* https://sproutsocial.com.

[28] Dataslayer. (2025). *How the Instagram algorithm works in* https://dataslayer.ai.

[29] NetInfluencer. (2025). *Instagram algorithm updates* https://www.netinfluencer.com.

[30] Chen, L., Tong, T. W., Tang, S., Han, N. (2022). Governance and design of digital platforms: A review and future research directions on a meta-organization. *Journal of Management*, *48*(1), 147–184.