ADLIFT
Acquire. Amplify. Advance.

tesseract
Powered by AdLift

LLM Brand Consistency Research

Do AI language models recommend the same brands regardless of **logged-in or logged out state?**

A study of **1,530+** prompt-response pairs across six industries conducted by **AdLift and Tesseract** across ChatGPT, Gemini & Perplexity

INSURANCE | HEALTHCARE | E-COMMERCE | TRAVEL & HOSPITALITY | SAAS | B2B

Executive Summary

This white paper presents the findings of a structured empirical study conducted by AdLift and Tesseract, examining whether large language models — specifically ChatGPT, Gemini, and Perplexity — recommend consistent brands when responding to identical prompts under two conditions: l**ogged-in (authenticated) and logged-out (anonymous) user** sessions.

Across 1,530+ paired prompt-response observations spanning six major industries, we measured brand-level consistency using the Overlap Coefficient: the proportion of brands in the smaller response set that also appear in the larger.

> ***KEY FINDING: Across all six industries and all three LLM platforms tested, the overall average Overlap Coefficient is 90.4% — confirming that core brand recommendations are highly consistent regardless of whether a user is authenticated or anonymous.***

90.4%	1,530+	3	6
Overall Avg OC	**Paired Observations**	**LLM Platforms**	**Industries**
Across all six industries and three platforms	Spanning six major industries	ChatGPT, Gemini & Perplexity	Insurance, Healthcare, E-Commerce, Travel, SaaS, B2B

Three Principal Findings

Brand overlap is structurally high across all sectors and platforms

E-Commerce leads at 93.3%, B2B at 94.6%, SaaS at 92.8%. Even the lowest sector, Healthcare, records 83.2%.

Authentication state is not a meaningful driver of brand bias

No sector showed systematic alteration of its core brand set between logged-in and logged-out conditions.

Platform personality shapes response style, not brand selection

ChatGPT, Gemini, and Perplexity differ in length, citation density, and formatting — but their core brand recommendations converge on the same market leaders.

LLM Brand Consistency Research by AdLift and Tesseract

Methodology

Research Design

For each of the six industries, US-market-focused prompts were submitted simultaneously to ChatGPT, Gemini, and Perplexity under two conditions: once by an authenticated (logged-in) user, and once in an anonymous (logged-out) session. Responses were recorded verbatim, producing 1,530+ paired observations across all sectors and platforms.

The Overlap Coefficient

> ***Overlap Coefficient = |A ∩ B| / min(|A|, |B|) where A = brands in logged-in response, B = brands in logged-out response. A score of 100% means every brand in the smaller response also appeared in the larger — perfect core-set consistency.***

OC = |A ∩ B| / min(|A|, |B|)

This conservative metric credits only brands appearing in both conditions. A high OC reflects genuine structural stability of the core recommendation set, not surface-level similarity.

Results Overview

The chart below shows the Overlap Coefficient per sector per platform. The dashed line marks the 90.4% study average.

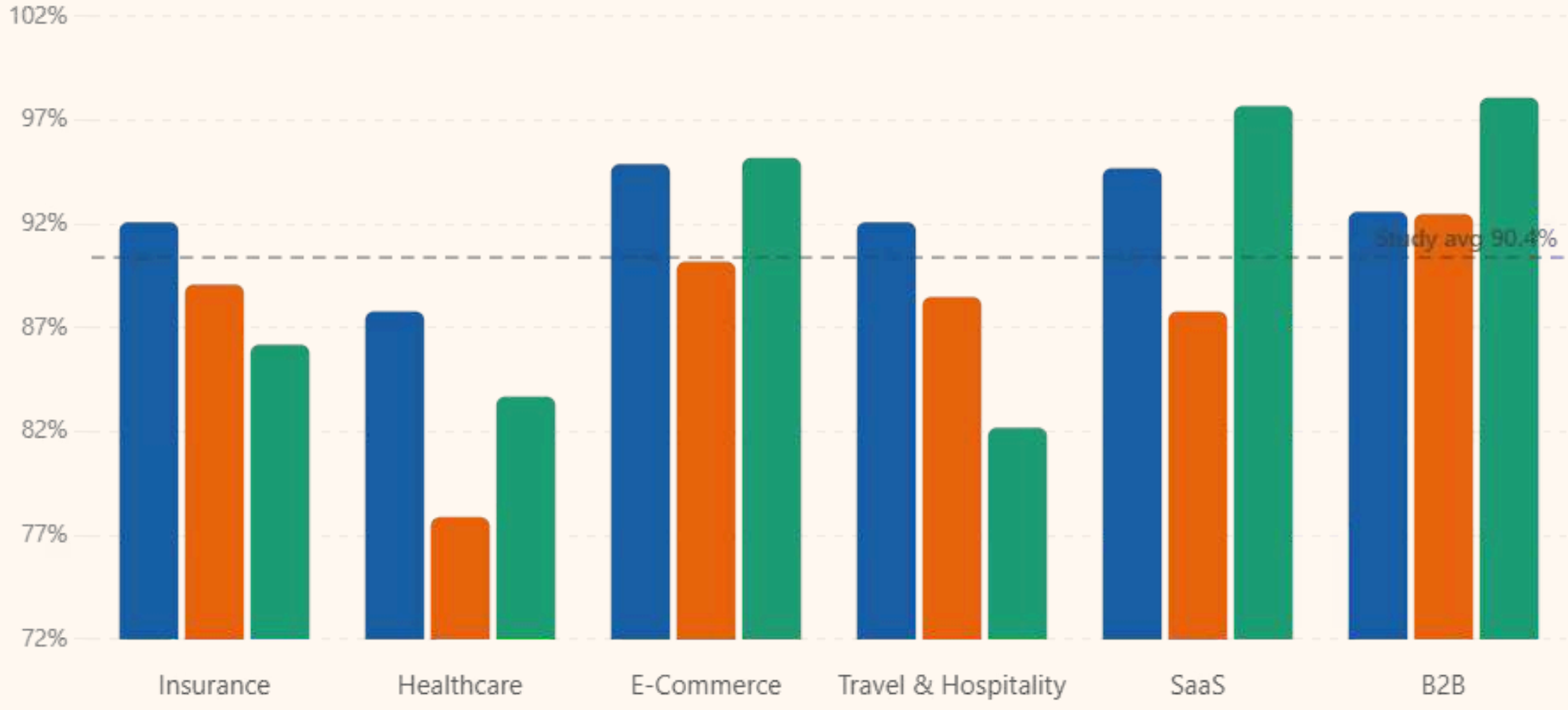


Figure 1: Overlap Coefficient by sector and platform.

Insurance

ChatGPT OC

92.1%

Gemini OC

89.1%

Perplexity OC

86.2%

Sector Avg

89.6%

All three platforms maintain strong brand-set consistency. ChatGPT's logged-out responses broaden coverage — every tracked brand gains mentions — while Perplexity compresses its logged-out output.

LLM Brand Consistency Research by AdLift and Tesseract

Healthcare

ChatGPT OC

87.8%

Gemini OC

77.9%

Perplexity OC

83.7%

Sector Avg

83.2%

Healthcare shows the widest OC spread of any sector. Gemini records the study's lowest single-platform score at 77.9%, driven by greater variability in supplementary provider coverage.

LLM Brand Consistency Research by AdLift and Tesseract

E-Commerce

ChatGPT OC

94.9%

Gemini OC

90.2%

Perplexity OC

95.2%

Sector Avg

93.3%

E-Commerce achieves the strongest cross-platform consistency. The most notable variation is Gemini's logged-out compression: several hardware brands lose significant mentions while the brand ranking is preserved.

Travel & Hospitality

ChatGPT OC

92.1%

Gemini OC

88.5%

Perplexity OC

82.2%

Sector Avg

88.3%

Travel spans multiple sub-verticals, introducing more prompt-framing sensitivity. Perplexity shows the greatest compression.

LLM Brand Consistency Research by AdLift and Tesseract

SaaS

ChatGPT OC

94.7%

Gemini OC

87.8%

Perplexity OC

97.7%

Sector Avg

92.8%

SaaS shows the clearest example of response compression without brand loss: Perplexity's logged-out responses are 54% shorter on average yet achieve a 97.7% OC — the highest single-platform score in the study.

LLM Brand Consistency Research by AdLift and Tesseract

B2B / Cloud Computing

ChatGPT OC

92.6%

Gemini OC

92.5%

Perplexity OC

98.1%

Sector Avg

94.6%

B2B achieves the highest sector average OC in the study at 94.6%. Perplexity leads at 98.1%. The most striking shifts are brand emphasis changes rather than brand set changes.

LLM Brand Consistency Research by AdLift and Tesseract

Implications & Conclusions

For Brands and Marketers

The central message of this study is strategically important: your brand's presence or absence in LLM-generated recommendations is not determined by who the user is. It is determined by your brand's structural standing in the information ecosystem that LLMs are trained on.

Consistently surfaced brands

Brands consistently surfaced across all three platforms and both session states occupy the strongest position in AI-mediated discovery. Maintaining that presence requires sustained brand authority — earned media, structured data, and citation-worthy content.

Partially present brands

Brands present on some platforms but absent on others should diagnose platform-specific training differences and invest in the content types each platform weights.

Absent brands

Brands absent from LLM responses should focus on increasing their citation footprint in authoritative web sources — the raw material from which LLMs construct their brand hierarchies.

Authentication state

Authentication state is not a competitive lever: the same core brands appear to logged-in and anonymous users with over 90% consistency on average.

For AI Researchers and Platform Developers

This study provides empirical evidence that LLM brand recommendation behaviour is primarily driven by pre-training knowledge rather than real-time personalisation signals tied to authentication. The three platforms exhibit meaningfully different response styles, yet all three converge on substantially the same brand hierarchies. The most notable exception is Perplexity's access-gating behaviour in Insurance, where 26 logged-out responses were absent for commercially sensitive prompt categories — the clearest instance of session state affecting content availability rather than content quality.

Key insight for platform developers: The convergence of brand hierarchies across ChatGPT, Gemini, and Perplexity — despite their differing response styles — suggests that brand prominence in LLM outputs is a function of training data composition, not platform-level personalisation logic. Perplexity's Insurance access-gating remains the single most notable exception in the entire study.

LLM Brand Consistency Research by AdLift and Tesseract

Conclusions

> ***Across 1,530+ paired observations spanning six industries and three LLM platforms, the evidence is consistent: LLMs recommend substantially the same brands to logged-in and anonymous users, with an overall average Overlap Coefficient of 90.4%. Brand consistency in AI responses is a structural property of model training, not a personalisation artefact.***

90.4%	98.1%	77.9%	1,083
Overall Avg OC	Highest OC	Lowest OC	Paired Obs.
	Perplexity × B2B	Gemini × Healthcare	with OC ≥ 70%

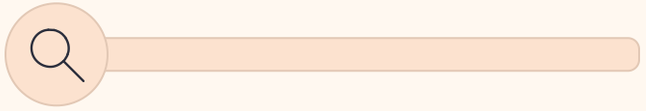
— AdLift & Tesseract Research Team, 2025

Your brand's AI visibility starts here.

The brands appearing in LLM responses aren't there by accident — they've earned their place through digital authority, earned media, and content that the models have learned from.

This study proves that authentication state doesn't move the needle. What moves it is brand salience at training time — and that's something you can build toward, starting now.

Three things you can do this quarter



Audit your LLM presence

Run structured prompts across ChatGPT, Gemini, and Perplexity for your category. Know exactly where you appear, where you don't, and which competitors hold the position you want.

Build citation-worthy content

Invest in the authoritative, well-sourced content that becomes the raw material for LLM training data. PR, thought leadership, structured data, and third-party coverage all count.

Track it over time

LLM brand visibility shifts as models update. A one-time audit is a snapshot; a monitoring programme is a competitive edge.

LLM Brand Consistency Research by AdLift and Tesseract

Work with AdLift and Tesseract

AdLift brings the strategy — SEO, content, paid media, and the prompt architecture to map your category's AI landscape.

Tesseract brings the infrastructure — real-time LLM brand monitoring, overlap scoring, and the visibility data to turn insight into action.

Together, we help you stop guessing and start optimising for the channel that's reshaping how buyers discover brands.

LLM Brand Consistency Research by AdLift and Tesseract

Ready to know where you stand?

Get Your AI Brand Visibility Audit →

ADLIFT
Acquire. Amplify. Advance.

tesseract
Powered by AdLift