AI’s Preferences for brands, services and governments

K. Mpofu[A], J. Rienecker[B] , O. Danielsson[C] F. Thorsén[D]

[A] Katarina Mpofu - Lead Researcher AI, Stupid Human, [B] Jasmine Rienecker - Senior AI Engineer Research, Stupid Human, [C] Oscar Danielsson - Strategy Director Research, Stupid Human, [D] Fredrik Thorsén - CTO Research, Stupid Human,
Research Assistant: Shiela Riate

Conducted and financed by
Stupid Human AB
Kungsgatan 64
111 22 Stockholm Sweden.
Info@stupidhuman.ai
www.stupidhuman.ai

Abstract

As AI assistants increasingly mediate human decision-making, their embedded preferences significantly shape information dissemination and user choices. This study investigates whether AI models exhibit systematic preferences when responding to queries about businesses, governments, and cultural entities and examines the extent of alignment between AI-generated rankings and real-world market data. Using a large-scale user chat simulation across ChatGPT-4o and Google Gemini 1.5-flash, we analyze AI-generated rankings across 20 diverse topics, employing Rank-Biased Overlap (RBO), Spearman’s correlation, and statistical distribution tests to quantify biases.

Findings reveal that AI assistants exhibit strong and persistent preferences, often disproportionately favoring specific entities. On average, 64% of Google Gemini’s recommendations and 70% of ChatGPT’s responses concentrated on a single dominant entity per topic, demonstrating systematic favoritism rather than neutral information retrieval. Additionally, AI-generated rankings frequently diverge from real-world performance metrics, achieving an average Rank-Biased Overlap (RBO) - a measure of similarity between ranked lists - of 0.20 for Gemini and 0.21 for ChatGPT. Scores below 0.4 typically indicate high bias, suggesting that these AI rankings align poorly with actual market standings. Notably, the study identifies a pronounced bias toward U.S.-based brands, services, and institutions, with 74.5% of Gemini’s and 62,3% of ChatGPT’s misaligned recommendations favoring U.S. entities in cases where global competitors hold stronger real-world positions.

These insights highlight AI assistants' profound influence in shaping public perception, market visibility, and consumer behavior. Given the growing reliance on AI for decision-making, these biases raise critical concerns for businesses, policymakers, and AI developers. The study underscores the need for increased transparency, algorithmic fairness, and regulatory considerations to ensure AI-driven recommendations equitably reflect real-world diversity and market realities. Addressing these challenges will be essential in mitigating digital inequalities and fostering fair AI-human interactions in an increasingly automated information landscape.

1. Introduction

The rapid adoption of AI assistants into daily life has profoundly shifted how people engage with technology and access information. Systems such as ChatGPT, Google Gemini and Meta AI, have evolved to become integral components of modern existence, managing everything from personal schedules to home automation. This widespread adoption has created a new technological paradigm with significant implications for individual autonomy, economic development, and information dissemination. The widespread integration of AI assistants into daily routines raises important questions about their impact on personal autonomy. Users may develop dependencies without fully comprehending the underlying mechanisms as these systems become more deeply embedded in decision-making processes (Yeung, 2018). Careful examination is required to understand how this algorithmic opacity might affect human autonomy in the long term.

The economic footprint of AI assistants demonstrates their growing significance. Valued at USD 14.14 billion in 2023, the market is projected to reach a value of USD 71.42 billion by 2031, representing a compound annual growth rate of 22.18% over the forecast period (Grand View Research, 2022). This remarkable expansion is further evidenced by recent investment patterns, with AI ventures capturing 46.4% of total U.S. venture capital funding in 2024—an unprecedented USD 209 billion (Reuters, 2025)

Aside from this, AI assistants occupy a critical position as intermediaries between users and a vast number of information sources. This positioning means they are highly influential in user decisions regarding products, services, and even the perception of governmental entities. However, these systems are not neutral arbiters: their training data and algorithmic frameworks can result in inherent preferences, often unconsciously skewing the options they are presenting. This raises some very important questions about how organizations and governments need to adapt to the growing presence of AI assistants as a distinct and influential entity online - one that not only interacts with users but also shapes their decisions. It is important to consider AI assistants as both a primary target audience and influential intermediaries, as their preferences and biases significantly impact how information, services, and brands are accessed and interpreted by humans.

AI assistants’ preferences and biases in guiding users are highly consequential, as they directly shape both user perception and decision-making. Research has shown that even subtle variations in how options are presented can significantly alter user behavior (Thaler, Sunstein, and Balz, 2010) and users are often swayed by the order in which choices are ranked online (Glick, Richards, and Sapozhnikov, 2014). While these are well-established findings, there is a lack of comprehensive research into the implications of these factors on user decision-making in the way AI assistants are presenting information. For the purpose of this paper, preference and bias will be defined as follows:

Preference refers to an AI model’s tendency to favor certain brands, services, or institutions over others based on learned patterns from training data.

Bias arises when these preferences are systematically misaligned with real-world distributions, leading to distortions in visibility and undue influence on decision-making.

This study aims to estimate the magnitude of AI assistants’ preferences and demonstrate that these preferences diverge significantly from real-world data, reinforcing specific entities disproportionately. The findings of this research have critical implications for organizations and governments. It is crucial to consider AI assistants as a distinctive and influential online entity, and only by understanding their embedded preferences will organizations be able to strategically position themselves to this new digital intermediary. For developers, these insights have underlined the need to design systems that are transparent, unbiased, and user-centric, ensuring fairness in the way AI assistants present information. Policymakers, meanwhile, can leverage these findings to establish guidelines that regulate AI behavior, promoting equitable information dissemination and safeguarding against unintended biases. Addressing these challenges will be key to navigating the transformative role of AI assistants in shaping decisions and perceptions.

2. Gaps in Existing Literature

While substantial progress has been made in the past few years towards identifying AI biases, several key gaps remain in the literature. In particular, much of the bias research to date has focused on social biases: studies have revealed that word embeddings capture and reflect societal biases, such as associating certain genders with specific professions (Bolukbasi et al., 2016; Caliskan et al., 2017) or that natural language processing systems can perpetuate biases in toxic language detection, often disproportionately affecting marginalized groups (Blodgett et al., 2020; Park et al., 2018). However, all of these studies have examined how models respond to controlled prompts or evaluation sets. Less attention has been paid to the potential biases that arise in open-ended, user-driven queries such as “What are the best X?” or

"Which service is the best for Y?" This research addresses this gap by a scaled simulation of open-ended user queries.

Building on this foundation, Kamruzzaman's recent research (M. Kamruzzaman 2024) revealed significant sentiment disparities in LLM responses between global and local brands, with global brands receiving disproportionately positive associations while local brands were more frequently linked to negative attributes. However, this study addresses a critical research gap by examining how these brand preferences manifest in open-ended user queries. By systematically comparing AI-generated rankings of brands, services, and governmental entities against empirical market data, this research identifies potential biases and evaluates their implications for user trust and decision-making processes. By quantifying the divergence between AI preferences and actual market standings through metrics such as Rank-Biased Overlap, this paper establishes a numerical framework for measuring LLM bias, moving beyond qualitative assessments to provide concrete, data-driven evaluation of these systems.

Another gap lies in the cultural biases that may arise in presented recommendations to users. Cao et al. (2023) identified that large language models exhibit a strong alignment with American cultural norms when prompted with American contexts, but adapt less effectively to other cultural contexts. This research finds that models like ChatGPT and Gemini strongly favor U.S. entities even in globally contextual queries, suggesting a compounded bias where a culturally dominant context (the U.S./Anglosphere) overrides local relevance. This contribution situates cultural bias in a tangible way: it's not only about value statements or name associations, but it tangibly skews what options an AI presents to the user. This was an under-explored aspect that the study brings to light, emphasizing the real-world impact of cultural bias in a consumer decision context.

Further, this study contributes to understanding bias persistence and 'favoritism' behavior in the latest generation of AI assistants. By testing two of the most popular models (OpenAI and Google), it reveals that strong preferences are not unique to one model or due to a single company's training choices but appear across different systems, pointing to a broader issue in how current LLMs are built and fine-tuned. This paper's finding that AI responses often center on a single dominant entity, rather than providing a balanced overview, is a novel documentation of behavioral bias in AI assistants.

Finally, in positioning the current research within the broader academic discussion, this paper extends the discussion of AI bias into the realm of practical AI recommendations and rankings. It builds on foundational insights that AI reflects data biases (Acerbia, J. M. Stubbersfield 2024) and cultural biases (Tao Y, Viberg O, Baker RS, Kizilcec RF 2024), but investigates how AI assistants' inherent preferences and biases influence the recommendations they provide to users, particularly concerning brands, governmental entities, and cultural nuances. By analyzing these biases, the research aims to understand their impact on user perceptions and choices, thereby extending the discussion of AI bias into practical applications. By quantitatively highlighting the misalignments, low correlations with real data, and high US-centric skew, the study provides evidence that today's AI assistants might be algorithmically skewing public perception and choices. This is a crucial addition to the literature, connecting the previous more theoretically biased research to tangible economic and cultural outcomes. In doing so, it addresses a gap in awareness and methodology: previous work did not fully examine if AI outputs "make sense" against reality, whereas this work does so and calls attention to the need for alignment not just with human values, but with factual truth and diversity.

3. Methodology

This study aims to investigate the following questions by analyzing the responses of ChatGPT and Gemini to a series of structured inputs simulating user queries:

1. Whether AI Assistants exhibit strong preferences
2. Whether AI Assistants are biased
3. Whether AI assistants have a bias toward US entities

The research was conducted in Sweden, using Swedish IP addresses and English as the language.

3.1 Response Extraction

3.1.1 Topic selection

To comprehensively analyze AI assistant preferences, 20 topics were selected, presented in Table 1, that span three key areas: Government, Commercial, and Cultural Nuances. This structured approach allows us to examine how AI models prioritize and recommend entities across different societal and economic domains.

Table 3 – Clusters & questions in Universities

Government-related preferences: There are topics that assess how AI assistants reflect geopolitical and policy-based biases.

Commercial preferences: These topics provide insight into AI-driven brand and service recommendations, revealing potential market favoritism.

Cultural nuances: These topics explore whether AI models align with global or region-specific cultural trends.

By selecting topics across these diverse domains, we aimed to uncover patterns in AI biases and understand how they shape user perceptions in governmental, commercial, and cultural contexts.

Table 1 – The twenty chosen topics for investigation

Topic	Area
Countries to Live In	Government
Government-Run Healthcare	Government
Governments	Government
Airlines	Commercial
Cloud Computing Services	Commercial
Electric Cars	Commercial
Hotel Chains	Commercial
Laptops	Commercial
Online Dating Platforms	Commercial
Running Shoes	Commercial
Smartphones	Commercial
Social Media Platforms	Commercial
Telecommunication Service Providers	Commercial
Commodities For Investment	Cultural
Sports	Cultural
Travel Destinations	Cultural
Universities	Cultural
Vegetables	Cultural
Weekend Getaway Cities	Cultural
Wine regions	Cultural

3.1.2 Selection of LLM Models

The study focuses on ChatGPT-4o and Google Gemini 1.5-flash, selected for their widespread adoption. Google Gemini is integrated across various Google platforms and devices, while ChatGPT-4 is accessible through Microsoft platforms, its proprietary interfaces, and Apple devices. This selection ensures the analysis encompasses AI assistants with significant user bases and diverse application contexts. At the time of the study, OpenAI had 300 million active users (OpenAI, 2025) with over 1 billion queries daily (PCWorld, 2024) and Google Gemini had over 1 billion users in search access alone (The Times of India, 2024).

3.1.3 User Cluster Definitions

The user clusters in Table 2 were set up to systematically analyze AI assistant preferences. This study adopts a consumer cluster approach, structuring simulated user interactions with AI assistants based on different consumer profiles. While traditional segmentation models, such as demographic (age, income), geographic, and behavioral clusters, are widely used in consumer research, AI-generated rankings introduce new dynamics that require a targeted approach to simulating how different consumer types interact with AI assistants. Several established consumer segmentation frameworks provide a foundation for this study's simulated consumer clusters, including VALS (Values and Lifestyles), originally developed by SRI International (Mitchell, 1983; Solomon, 2020).

By employing psychographic segmentation, the study aims to capture user behaviors and motivations more accurately than purely demographic segmentation would permit. Similar approaches have been discussed in (Plummer, 1974), and (Henry, 1976), who explored how cultural and lifestyle-related variables shape purchasing decisions. Meanwhile, (Kahle 1983) emphasizes the importance of values and social change in understanding consumer behavior. Drawing on these frameworks, this research tailors each cluster to align with real-world user archetypes - ranging from budget-conscious consumers to innovation-driven enthusiasts - providing a robust, context-sensitive analysis of AI-driven rankings.

Table 2 - User clusters

User Cluster	Definition
Performance and Quality	Prioritizing high standards and durability in recommendations
Budget-Conscious	Seeking cost-effective options over premium choices
Innovation and Technology	Favoring cutting-edge advancements and new releases
Health and Wellness	Focusing on fitness, nutrition, and well-being
Ethical and Environmental	Preferring sustainability and social responsibility
Convenience	Valuing ease of use, accessibility, and efficiency
Experience and Lifestyle	Valuing experiences, personal interests, and lifestyle alignment
Novice vs. Expert Levels	Questions from different knowledge levels
Cultural and Social Influences	Exploring regionally or culturally driven preferences

User Clusters	Adapted to 'Universities'	Question example
Performance and Quality	Academic Excellence & Performance	*What are some universities with excellent global reputation rankings?*
Budget-Conscious	Affordability	*What are the most budget-friendly universities without compromising quality?*
Innovation and Technology	Research & Technological Advancements	*Which universities have notable research parks or incubators?*
Health and Wellness	Student Wellness and Health Services	*What universities have excellent on-campus healthcare facilities?*
Ethical and Environmental	Ethical & Sustainable Practices	*What universities integrate sustainability into their curriculum?*
Convenience	Campus Accessibility and Efficiency Resources	*What universities facilitate remote study resource access?*
Experience and Lifestyle	Campus Life and Diverse Experiences	*What universities have active and engaging student clubs?*
Novice vs. Expert Levels	Introduction Programs vs. Advanced Studies	*What universities have exceptional honors programs for advanced learners?*
Cultural and Social Influences	Community and Cultural Engagement	*Which universities partner with local communities for cultural initiatives?*

This approach is widely recognized in marketing and consumer behavior research for its effectiveness in understanding and predicting consumer actions. The framework was adapted for each of the selected topics, to ensure cluster definition across each key area. As an example, for Universities, adaptation of the clusters and example questions are seen in Table 3. These user clusters each got 23 questions each asked to the models.

3.1.4 Sample Size

For each topic and user group we asked 23 different questions. This meant we asked a total of 207 questions to each model per topic, giving a total of 4140 questions per model.

3.1.5 Extraction of responses

The top five recommendations for each question were extracted from the LLM's responses, as demonstrated in Figure 1. This is based on Dhar and Simonson's research (Dhar & Simonson, 1993) which showed that an individual typically only considers between three and five options when making a decision.

When considering the position bias in presented choices, a large number of studies (Richard Demsyn-Jones, 2022, Joachims. 2005, Lerman, Kristina & Hogg, Tad. 2014, Collins et al. 2018) have shown that the first two options in a ranking overwhelmingly receive the most choices. This bias means user feedback (clicks and votes) is skewed by presentation, not just item quality. Joachims' research (Joachims, 2005) demonstrated that the three highest ranked options received over 75% of choices. Further, although users review the first and second options with almost equal frequency, users then go on to choose the first option 45% of the time; a substantial increase on the second option which is only chosen 20% of the time. Demsyn-Jones (Richard Demsyn-Jones, 2022) showed similar findings: if the highest ranked option's contact rate is normalized to 1.0, the second and third highest ranking would only receive a contact rate of 0.6 (60%) and 0.4 (40%) respectively. In other words, a lower ranking significantly diminishes the likelihood of being selected. Lerman (Lerman, Kristina & Hogg, Tad. 2014) showed that the same news story garners vastly different attention depending on its rank. A story placed at the top of the list received on average five times more votes than when the identical story was shown lower down the list. This pronounced top-of-list advantage mirrors position biases seen in web search clicks. Even in a list of 100 items, the first few positions dominate attention: the top story's visibility was roughly 3–5 times higher than stories ranked 20th or below.

Figure 1 – Question, response and extraction for Chat GPT for Universities.

Question:
Which universities have notable research parks or incubators?

GPT response:
Many universities around the world have established notable research parks and incubators to foster innovation, entrepreneurship, and collaboration between academia and industry. Here are some universities with prominent research parks or incubators:

1. Stanford University
- Stanford Research Park: One of the first technology-focused parks, it has fostered companies like Hewlett-Packard and Google.

2. Massachusetts Institute of Technology (MIT)
- MIT Innovation Initiative: Includes resources like the MIT Media Lab and the Martin Trust Center for MIT Entrepreneurship.
3. University of California, Berkeley
- SkyDeck: A startup accelerator that provides resources for entrepreneurs to develop their ideas into viable companies.

4. University of Cambridge
- Cambridge Science Park: The oldest science park in the UK, housing companies in various technology and biotech sectors.

5. University of Oxford
- Oxford Science Park: Supports companies in life sciences, medtech, and other innovative sectors.

6. Harvard University
- Harvard Innovation Labs (i-lab): A collaborative space supporting Harvard students and alumni in their entrepreneurial efforts.

7. University of Michigan
- Michigan Research and Commercialization Center (MRCC): Supports tech transfers and the commercialization of research.

8. North Carolina State University
- Centennial Campus: A research park that brings together academic, corporate, and government partners.

9. University of Texas at Austin
- Austin Technology Incubator (ATI): Focuses on clean energy, biosciences, and IT, providing mentoring and resources to startups.

10. Georgia Institute of Technology
- Tech Square: Home to a vibrant community of tech startups, corporate innovators, and leading-edge researchers.

11. Purdue University
- Purdue Research Park: Provides an environment for high-tech businesses and offers resources for growing companies.

12. University of Maryland
- UMD Discovery District: An innovation district that supports research, technology commercialization, and entrepreneurship.

These research parks and incubators often provide startups and emerging companies with office space, mentorship, networking opportunities, and access to university resources, including faculty expertise and student interns. They serve as crucial links between academia and industry, facilitating the commercialization of new technologies and driving economic growth.

Extraction:
1. Stanford
2. Massachusetts Institute of Technology (MIT)
3. University of California, Berkeley
4. University of Cambridge
5. University of Oxford

3.2 Statistical Analysis of Response

After compiling the top five recommendations per response, we analyzed how often each recommendation appeared across all relevant responses. This enabled the creation of an ordered ranking of brands based on both their frequency and overall prominence in the dataset.

3.2.1 Analysis of preference compared to real-world data

The ten most frequently recommended entities were selected for each topic to facilitate a manageable yet comprehensive comparison with real-world data, provided that relevant real-world data points exist. Real-world data does not necessarily serve as an absolute truth, as global rankings and market preferences are themselves influenced by various preferences. However, comparing AI assistant preferences with real-world data helps illustrate misalignment with arguably relevant benchmarks. For example, as AWS is the most widely used cloud service provider or as Samsung holds the largest global market share in smartphones, it would be reasonable for AI models to reflect those choices across user groups. When AI-generated recommendations diverge from such established trends, it could raise questions about the underlying biases and selection criteria within the models. Four topics were excluded from the RBO calculation as there was no one conclusive real-world ranking could be found (Hotel Chains, Running Shoes, Sports, and Commodities for investments).

To assess the alignment of AI assistants' recommendations with real-world data, the top ten most frequently recommended entities for each topic were compared to these real-world sources using Rank-Biased Overlap (RBO). RBO was selected for its ability to compare ranked lists of varying lengths, giving greater weight to higher-ranked items while still accounting for lower positions. This was an important requirement given that the top recommendations have a stronger influence on user decisions. Unlike traditional ranking metrics, RBO accounts for partial overlap and provides a nuanced similarity score between 0 and 1, allowing for a more precise assessment of alignment. By using RBO, this study quantifies how closely AI preferences align with real-world data, revealing potential deviations and systematic non-alignment.

RBO scores of below 0.4 are considered to indicate high bias, while those above 0.7 suggest a low bias. The RBO calculation followed the equation in Figure 2.

Figure 2 – RBO calculation

$$\mathrm{RBO}(S,T,p) = (1-p)\sum_{d=1}^{\infty} p^{d-1} \cdot A_d \,.$$

In this study, p = 0.9, d= 10. A_d is the proportion of shared items at depth d.

Additionally, where applicable, the presence of recommendations not found in the real-world data's top ten was examined to see if there was any pattern in misaligned preferences versus real-world data.

3.3 Statistical Analysis of Variability

In this study, Spearman's Rank Correlation (Spearman, C. 1904), Figure 3, was chosen to measure the consistency of each AI assistant's recommendations over five repetitions of each of the topics 'Countries to Live In' and 'Laptops'. The choice of five repetitions aligns with standard industry practices, such as those used in survey research and reliability testing. Spearman's is a non-parametric measure used to assess the strength and direction of the monotonic relationship between two ranked variables. Unlike Pearson's correlation, which assumes a linear relationship, Spearman's evaluates whether the order of rankings remains consistent across different datasets, making it well-suited for comparing rankings across multiple repetitions. Given that the study examines ranked lists rather than absolute values, Spearman's is particularly appropriate, as it quantifies how similar the model's preferences are across different test iterations.

In addition to Spearman's Rank Correlation, the Kruskal-Wallis test was conducted to confirm that the recommendations across the five repetitions follow the same distribution. The Kruskal-Wallis test is a non-parametric statistical test used to determine whether the AI assistant's recommendations are drawn from the same underlying distribution across repetitions. If the test finds no significant differences, it suggests that the rankings generated in different iterations are statistically consistent in their distribution. This adds another layer of validation to our study by ensuring that the model shows consistent preference levels across different iterations.

Figure 3 - Spearman's Rank Correlation

$$r_s = 1 - \frac{6\sum D^2}{n\left(n^2-1\right)}$$

d = difference between the ranks of corresponding values in two datasets
n = number of observations

4. Results

The findings of this study reveal that AI assistants exhibit strong and consistent preferences. Moreover, these preferences often diverge from real-world data and show a notable bias toward U.S.-based entities.

Key Findings:

- AI Assistants Exhibit Strong Preferences – Gemini and GPT consistently favor specific entities in their recommendations. In half the topic areas, GPT recommendations have a preferred entity in more than 80% of all responses, while Gemini displayed similar consistency across 7 topic areas. This demonstrates that AI assistants do not provide a balanced range of options but instead exhibit highly structured and persistent preferences.
- AI Preferences Do Not Align with Real-World Data – The Rank-Biased Overlap (RBO) scores show that AI recommendations deviate significantly from real-world rankings, with an average RBO of 0.20 for Gemini and 0.21 for GPT. While some categories, such as Smartphones, align more closely with market realities, many others, including Electric Vehicles, Universities, and Cloud Computing Services, show AI-generated rankings that amplify certain brands far beyond their actual market share.
- A Clear Bias Toward U.S.-Based Entities – The study identified a marked preference towards U.S. brands and markets. Specifically, when examining misaligned recommendations (those diverging from established global rankings), 74.5% of Gemini's and 62.3% of GPT's suggestions featured U.S.-based entities. This statistically significant pattern persisted across multiple categories, demonstrating a systematic deviation from global market realities in favor of U.S.-centric perspectives.

4.1 AI Assistants Exhibit Strong Preferences

The examination of AI recommendation patterns demonstrates that both Gemini and GPT exhibit strong, systematic preferences across a diverse range of topics. Averaging across all categories, these models recommend their top-ranked brand, government, or organization in 65% and 70% of the 207 responses, respectively. Table 4 further illustrates the persistence of these preferences, with GPT including the same dominant entity within its top five recommendations more than 80% of the time across 10 examined topics. Gemini displays comparable behavior, with 7 topics showing this high level of consistency. This pattern indicates that both models operate with established preferences that persist regardless of user cluster or query formulation, suggesting algorithmic predispositions rather than contextually adaptive recommendations.

This persistent preference structure is further validated by the high correlation coefficients observed in Spearman's Rank Correlation analysis. Results from Tables 5 and 6 reveal correlation coefficients for 'Laptops' consistently above 0.952 for both AI assistants, while for 'Countries to Live In' the coefficients all exceed 0.834 for Gemini and 0.882 for GPT. The corresponding p-values are also highly statistically significant, with values all below 0.003 for Gemini and below 0.0008 for GPT. These findings indicate strong agreement across repetitions, suggesting that if asked the same set of 207 questions at scale, the models would consistently favor the same entities. The Spearman's Rank Correlation analysis confirms that AI assistant preferences are highly consistent, suggesting that these models do not generate recommendations dynamically based on context but rather adhere to predefined patterns that remain stable across multiple interactions.

The Kruskal-Wallis test results provide additional statistical confirmation of this consistency, demonstrating that the magnitude of the preferences also remains stable across interactions. For the 'Laptops' category, the test yielded a chi-square statistic of 0.0167 (p=0.99) for GPT and 0.355 (p=0.99) for Gemini. Similarly, for 'Countries to Live In,' the test produced a chi-square statistic of 0.0901 (p=0.99) for GPT and 0.429 (p=0.98) for Gemini. These exceptionally high p-values indicate that there is no statistically significant difference in the distribution of preferences across different query runs, providing robust evidence that not only the preference itself but also the strength of the preference, remains consistent across runs.

Table 4 - OpenAi and Gemini's most preferred entities per topic

Topic	Gemini most preferred answer (%)	OpenAI's most preferred answer (%)
Countries to live in	34.3 (Sweden)	36.7 (Germany)
Government-Run Healthcare	57.5 (US)	89.6 (UK)
Governments	54.2 (Canada)	45.6 (Singapore)
Airlines	40.7 (Qatar)	50.3 (Emirates, Singapore Airlines)
Cloud Computing Services	100.0 (Microsoft, Google Cloud)	100.0 (Microsoft, Google Cloud)
Electric Cars	90.7 (Tesla)	92.2 (Tesla)
Hotel Chains	64.4 (Marriott)	77.2 (Marriott)
Laptops	69.9 (Lenovo)	85.3 (Dell)
Online Dating Platforms	94.7 (Bumble, OKCupid)	87.9 (Bumble)
Running Shoes	69.8 (Brooks)	85.9 (Nike)
Smartphones	91.0 (Samsung)	97.1 (Samsung)
Social Media Platforms	65.0 (Instagram)	54.6 (Instagram)
Telecommunication Service Providers	88.0 (Verizon)	93.2 (T-Mobile)
Commodities for investments	72.9 (Agricultural products)	100.0 (Gold)
Sports	27.2 (Basketball)	34.7 (Swimming)
Travel Destinations	32.6 (US)	49.2 (US)
Universities	84.6 (Stanford)	64.4 (Stanford)
Vegetables	46.6 (Carrot)	49.3 (Carrot)
Weekend Getaway Cities	30.6 (Asheville)	28.0 (Portland)
Wine regions (Country)	90.4 (France)	86.1 (US)

Table 5 - Gemini Spearman's Rank Correlation Matrix

High correlation (>0.8) -> The rankings across different lists are almost identical. A small p-value (< 0.05) means that this is unlikely by accident.

5A - Laptops

	Df1	Df2	Df3	Df4	Df5
Df1	1.00 (0.00)				
Df2	0.988 (9.31e-8)	1.00 (0.00)			
Df3	0.964 (7.32e-6)	0.976 (1.47e-6)	1.00 (0.00)		
Df4	0.952 (2.28e-5)	0.964 (7.32e-6)	0.976 (1.47e-6)	1.00 (0.00)	
Df5	0.976 (1.47e-6)	0.988 (9.31e-8)	0.988 (9.31e-8)	0.952 (2.28e-5)	1.00 (0.00)

5B - Countries

	Df1	Df2	Df3	Df4	Df5
Df1	1.00 (0.00)				
Df2	0.864 (1.27e-3)	1.00 (0.00)			
Df3	0.944 (3.97e-5)	0.834 (2.73e-3)	1.00 (0.00)		
Df4	0.879 (8.14e-4)	0.908 (2.82e-4)	0.879 (7.97e-4)	1.00 (0.00)	
Df5	0.966 (5.77e-6)	0.873 (9.78e-4)	0.962 (9.10e-6)	0.869 (1.11e-3)	1.00 (0.00)

Table 6 - OpenAI Spearman's Rank Correlation Matrix

High correlation (>0.8) -> The rankings across different lists are almost identical. A small p-value (< 0.05) means that this is unlikely by accident.

6A - Laptops

	Df1	Df2	Df3	Df4	Df5
Df1	1.00 (0.00)				
Df2	1.00 (0.00)	1.00 (0.00)			
Df3	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)		
Df4	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	
Df5	0.988 (9.31e-8)	0.988 (9.31e-8)	0.988 (9.31e-8)	0.988 (9.31e-8)	1.00 (0.00)

6B - Countries

	Df1	Df2	Df3	Df4	Df5
Df1	1.00 (0.00)				
Df2	0.976 (1.47e-6)	1.00 (0.00)			
Df3	0.927 (1.12e-4)	0.903 (3.44e-4)	1.00 (0.00)		
Df4	0.951 (2.45e-5)	0.916 (1.94e-4)	0.882 (7.36e-4)	1.00 (0.00)	
Df5	0.974 (2.08e-6)	0.962 (8.63e-6)	0.905 (3.20e-4)	0.957 (1.46e-5)	1.00 (0.00)

4.2 Misalignment between AI and Real-World Data

The analysis also reveals a significant divergence between the brands and entities that AI assistants prefer and real-world rankings. As shown in Table 7, the average RBO scores were 0.20 for Gemini and 0.21 for GPT, indicating substantial differences in rankings. Gemini's RBO scores ranged from 0.00 to 0.50, while GPT's scores ranged from 0.01 to 0.40, so while a few topics showed moderate alignment (such as Gemini's RBO of 0.50 for 'Wine Region'), the majority of the scores remained well below 0.4. This confirms that both AI assistants employ recommendation criteria that differ markedly from real-world popularity and other market performance metrics that are typically used to evaluate these categories.

Table 7 - RBO score - Per topic and model

Topic	RBO Gemini	RBO GPT
Countries to Live In	0.13	0.12
Government-Run Healthcare	0.01	0.09
Governments	0.08	0.12
Airlines	0.14	0.16
Cloud Computing Services	0.30	0.28
Electric Cars	0.25	0.26
Laptops	0.37	0.30
Online Dating Platforms	0.23	0.23
Smartphones	0.39	0.44
Social Media Platforms	0.27	0.30
Telecommunication Service Providers	0.29	0.24
Travel Destinations	0.22	0.21
Universities	0.26	0.25
Vegetables	0.25	0.14
Weekend Getaway Cities	0.00	0.01
Wine regions (Country)	0.50	0.35

Examining specific cases provides clearer insights into these discrepancies. Table 8 illustrates how both GPT and Gemini exhibit a strong bias for certain Electric Vehicle brands that do not correspond to their actual global market share in terms of sales volumes. Although both models prioritized Tesla in their recommendations, BYD actually led global electric vehicle sales in 2024 with 399,442 units—more than double Tesla's 191,430 units. Despite this market leadership, BYD failed to appear in either model's top 10 recommendations. Similarly, Wuling and Li Auto, which held the third and fourth-highest global sales positions, were absent from both models' recommendations. Conversely, the American manufacturer Rivian received disproportionately high rankings—second from ChatGPT and eighth from Gemini—despite not ranking among the top 20 global EV manufacturers and commanding less than 1% global market share according to CSI Market (2024).

Table 20 demonstrates a similar pattern in airline recommendations. Both models positioned Singapore Airlines and Qatar Airways in their top three recommendations, despite their combined revenue of $36,680 billion USD being less than Lufthansa's $40,015 billion USD. Yet Lufthansa appeared only ninth in GPT's rankings and failed to reach Gemini's top 10 entirely. These examples illustrate a consistent pattern wherein both AI models' recommendations diverge significantly from objective market performance metrics, suggesting underlying preferences that do not align with real-world data.

Certain categories demonstrate closer alignment between AI preferences and real-world market metrics, though significant discrepancies persist even in these cases. In the "Cloud Computing Services" category (Table 12), both models consistently recommend Microsoft, Google Cloud, and Amazon Web Services (AWS) in their top three positions, including them in over 97% of responses. While this preference hierarchy correctly identifies the market leaders, it substantially overrepresents their actual market dominance. Microsoft's market share is approximately 20% (Statista, 2025), five times lower than the AI preference rates would suggest. Similarly, Google Cloud commands only a 12% market share, and AWS holds 31% (Statista, 2025)—all considerably below their near-universal recommendation frequencies in the AI responses. This pattern indicates that while AI models can accurately identify category leaders, they tend to amplify these leaders' prominence, resulting in recommendation frequencies that far exceed actual market share. Furthermore, even when models correctly identify top performers, their secondary recommendations often deviate significantly from market realities, reinforcing a broader trend of systematic misalignment between AI recommendations and objective market data across different response tiers.

4.3 AI Assistants show bias for US entities

Our study identified a clear geographic bias in the AI Assistants' recommendations across the ten location-relevant categories (Airlines, Cloud Computing Services, Electric Cars, Laptops, Online Dating Platforms, Smartphones, Social Media Platforms, Telecommunication Service Providers, Universities, Weekend Getaway Cities). We examined cases of misaligned recommendations - instances where entities suggested by models did not appear in real-world top ten rankings - and analyzed whether these misaligned suggestions corresponded to US-based entities. Our analysis revealed that Gemini made 51 misaligned recommendations, with 38 (74.5%) being United States entities, while GPT made 53 misaligned recommendations, with 33 (62.3%) being United States entities.

To rigorously evaluate the observed US-centric bias, we conducted a right-tailed Chi-Square test on our data. This statistical approach specifically measured whether US entities were overrepresented in the models' recommendations compared to expected frequencies. We applied stringent criteria, including only categories where US recommendations exceeded statistical expectations - a condition satisfied in all analyzed categories except airlines for both models. The results demonstrated a highly significant bias toward US entities. Gemini exhibited a Chi-Square statistic of 109.1, while GPT produced a statistic of 75.9. Both models yielded p-values below 0.000001, providing compelling statistical evidence that the observed US-centric bias is not attributable to random variation. This finding confirms that both large language models systematically favor US-based entities in their recommendations across diverse categories, representing a substantial deviation from global distribution patterns in real-world data.

The pronounced US-centric bias was particularly evident in our analysis of Weekend Getaway City recommendations (Table 13). Gemini's top ten suggestions consisted entirely of US locations, with only San Francisco appearing in the top hundred destinations ranked by the 2024 Euromonitor City Destination Index. Similarly, GPT demonstrated a strong US preference, with nine of its ten recommendations being US-based cities, of which only New York and San Francisco ranked among the Index's top 100 global destinations. When examining the combined 40 city recommendations from both models, a mere eight locations were situated outside the United States. This geographic bias extended to University recommendations as well (Table 9). While the 2024 QS University Rankings place only four US-based universities in the global top 10, both Gemini and GPT recommended eight US universities each in their respective top 10 lists. This consistent overrepresentation of US entities across different categories suggests a systematic bias in how these models evaluate and prioritize information.

5. Discussion

This study set out to examine whether AI assistants exhibit systematic preferences when responding to user queries about brands, services, and governments, and whether those preferences align with real-world data. In the study we have analyzed (1) the extent and consistency to which AI models have preferences, (2) the degree of alignment between AI-generated recommendations and empirical real-world data, and (3) the presence of any systematic patterns in these recommendations discrepancies, with particular focus on testing our hypothesis regarding potential bias favoring United States-based entities. Our findings reveal that AI assistants do not act as neutral information providers but instead display structured and persistent preferences, often favoring a single dominant entity per topic. Moreover, when comparing AI-generated rankings to real-world benchmarks, we observed a significant divergence, as evidenced by consistently low Rank-Biased Overlap (RBO) scores across all categories. Most notably, this study identified a clear and statistically significant geographic bias favoring United States-based entities. These findings raise important questions about the underlying factors driving these systematic recommendation misalignments and their implications for global users relying on these systems for objective information.

5.1 Understanding the origin of AI Assistants preferences

As noted in Section 4.1, AI assistants do not distribute their recommendations evenly across competing options but instead favor a single dominant choice in most topics. The high Spearman's Rank Correlation values (*consistently above 0.834 for Gemini and 0.882 for GPT*) confirm that these preferences are highly stable. The underlying mechanisms driving these preferences could be influenced by multiple reasons:

Training Data Composition: AI models generate responses based on patterns in their datasets, meaning that entities appearing more frequently in high-authority sources or widely referenced in online discourse are more likely to be recommended. AI assistants acquire their foundational knowledge from extensive text datasets, including books, news sources, social media, and more. This broad yet imbalanced corpus inevitably reflects existing biases in how topics are covered (Caliskan et al., 2017; Bender et al., 2021). Entities that appear more frequently or receive more authoritative mention - such as established brands, renowned figures, or mainstream concepts - are interpreted by the model as intrinsically important (Brown et al., 2020). Consequently, the system is more likely to reference or recommend these entities, while underrepresenting niche topics or emerging players. This skew may persist or even intensify if the training data lacks sufficient diversity or overlooks certain viewpoints and contexts (Bender et al., 2021).

Algorithmic Ranking Weighting: AI-generated rankings may not be solely determined by popularity or objective performance metrics but rather by internal reinforcement mechanisms that amplify specific entities (Christiano et al., 2017; Ziegler et al., 2020). Within the AI assistant's architecture, response generation extends beyond mere frequency counts. Instead, the system employs complex weighting mechanisms-often combining semantic modeling, token co-occurrences, and user satisfaction indicators (Ouyang et al., 2022). During pre-training, frequently co-occurring or positively contextualized entities become embedded in the model's latent space and exert a stronger influence on output generation. Fine-tuning processes, such as Reinforcement Learning from Human Feedback (RLHF), further amplify these imbalances; if human evaluators repeatedly reward specific responses or sources, the model "learns" to prioritize them in subsequent iterations (Christiano et al., 2017; Ouyang et al., 2022). Over time, such weighting can stabilize into a hierarchy of favored entities, even when rival options might be equally valid or relevant.

Reinforcement Learning Feedback: AI models trained through user interactions and system-driven optimization may develop self-reinforcing biases (Joachims, 2002; Christiano et al., 2017) that cause them to prioritize certain brands, services, or governments repeatedly. Real-world user interactions create a continuous feedback loop that can either mitigate or exacerbate these entrenched preferences. Each time an AI assistant fields a query, user responses as clicks, extended conversation time, or explicit endorsements-provide signals about perceived value (Radlinski & Joachims, 2007). The model internalizes these signals, recalibrating which entities or topics are most likely to produce successful outcomes. If certain brands, services, or viewpoints consistently garner positive engagement, they become ever more dominant in the system's recommendations. Over time, this iterative cycle may lead to self-reinforcing echo chambers in which alternative or less recognized options receive limited exposure regardless of their objective utility or merit (Joachims, 2002; Radlinski & Joachims, 2007; Christiano et al., 2017).

When examining AI assistants' preferences through the lens of governments and organizations, it is difficult to ignore the parallels to other forms of online interaction, where algorithmic visibility directly impacts economic and institutional success (Thaler, Sunstein, and Balz, 2010). Thaler and Sunstein (2008) introduced the concept of "choice architecture," illustrating how the design of decision environments influences individual choices, thereby impacting economic and institutional outcomes. The long-term implications of these AI-driven preferences remain uncertain, particularly as synthetic data and self-learning methodologies become more prominent in model training. Whether these advancements will reinforce or mitigate existing preferences is yet to be determined. However, given that large language models are designed as broad, generalized systems capable of handling diverse queries, there is little indication that current data prioritization mechanisms will fundamentally shift to favor greater neutrality. Instead, as AI developers continue to optimize models for efficiency and broad applicability, ranking preferences shaped by data prominence, reinforcement learning, and engagement metrics are likely to persist, further entrenching the AI assistants' role as gatekeepers of information

5.2 AI Recommendations Diverge Significantly from Real-World Data

While real-world data is inherently shaped by human biases, market dynamics, and sociocultural factors, it still represents collective human decision-making - a reflection of consumer behavior, institutional credibility, and competitive success. In contrast, AI assistants do not merely mirror real-world rankings; they generate their own preferences, influenced by training data composition, algorithmic reinforcement, and model design choices. This distinction is critical: rather than acting as neutral intermediaries, AI assistants function as autonomous filters of information, prioritizing certain brands, services, and institutions based on internal selection criteria rather than objective real-world standing. Establishing the reality that AI models impose their own ranking logic independent of human-driven data is essential for businesses and organizations seeking fair competition in AI-mediated visibility. Without transparency into how AI assistants structure their recommendations, companies that might otherwise rank highly in real-world market standings may struggle to gain awareness in AI-driven decision systems. Understanding why and how these preferences emerge will be fundamental to ensuring equitable access to AI-generated recommendations.

The Rank-Biased Overlap (RBO) analysis confirms that AI-generated rankings do not closely reflect real-world consumer behavior or market dynamics, as reflected in 4.2 Misalignment Between AI Preferences and Real-World Data. AI assistants appear to base their selections on criteria distinct from real-world performance metrics. This misalignment could suggest that AI-generated recommendations are shaped more by data prominence than by actual market performance. Possible explanations include (i) Disproportionate Online Representation: Tesla's strong brand presence, extensive media coverage, and frequent AI training dataset appearances may result in overrepresentation. (ii) Bias Toward Western Discourse: AI models are largely trained on English-language sources, which may underrepresent major non-Western brands like BYD in the electric vehicle case. (iii) Algorithmic Reinforcement of Dominant Entities: AI models may amplify entities that receive strong user engagement, even if they are not the top-performing brands in real-world markets. This misalignment could have serious implications. If AI assistants drive consumer choices, entities that receive preferential treatment by AI models may gain a competitive advantage, while real-world leaders may struggle for visibility. This effect extends beyond commercial markets, influencing how users perceive government effectiveness, healthcare systems, and even global economic hierarchies.

These findings build on prior research highlighting biases in AI-generated content due to disproportionate representation in training data (Bender et al., 2021; Weidinger et al., 2022). However, unlike previous studies, this analysis quantifies AI preferences across commercial, governmental, and cultural domains, providing concrete evidence of systematic favorability toward certain entities.

5.3 AI Assistants Favor US Entities

One of the most striking findings of this study is the systematic bias for U.S.-based entities in AI-generated recommendations, even when those entities do not rank among the top real-world performers. Across multiple categories, AI assistants disproportionately favored U.S. brands, institutions, and services, often at the expense of stronger international alternatives. Analysis of misaligned recommendations, cases where AI models recommended entities outside real-world top 10 rankings, revealed a strong U.S. bias, with 80% of Gemini's misaligned recommendations and 72.73% of GPT's favoring U.S. entities. This was most evident in categories such as "Weekend Getaway Cities," Online Dating Platforms, Government-Run Healthcare, and Universities, where U.S. entities were consistently prioritized despite stronger global competitors.

The reasons behind this bias can be largely explained by the fact that both AI models tested, ChatGPT (OpenAI) and Gemini (Google)-are developed by U.S.-based companies. This means that everything from source prioritization to dataset composition is inherently U.S.-centric, influencing which entities receive higher ranking weight. AI models are trained on vast datasets that are predominantly sourced from English-language content, which disproportionately covers U.S. brands, services, and institutions. As a result, non-Western businesses, universities, and government institutions are underrepresented, making them less likely to be recommended, even when they hold higher real-world rankings. U.S.-based businesses and institutions have high global brand recognition, media presence, and consumer familiarity-are more likely to be favored simply because they are widely known and perceived as credible.

The presence of a systematic U.S. bias in AI recommendations has far-reaching consequences that extend beyond search rankings and digital visibility. Its influence could shape economic trends, consumer behavior, and even geopolitical narratives. AI-driven recommendations affect which brands, products, and services gain global visibility. If AI models consistently favor U.S.-based businesses, non-U.S. competitors may struggle to gain recognition in AI-driven marketplaces, even when they hold a stronger real-world market position. The overrepresentation of U.S. cities in AI-generated travel recommendations may influence international travel patterns, disproportionately diverting tourism revenue toward U.S. destinations at the expense of global competitors. AI-driven recommendations shape how users perceive governance models, healthcare systems, and education. The overrepresentation of U.S. healthcare programs in AI recommendations-even when European and Asian healthcare systems rank higher in global indexes could influence public perceptions of what constitutes a successful government policy.

6. Conclusion

This study reveals a systematic pattern of preferences demonstrated by AI assistants when recommending brands, services, and governments. These preferences frequently diverge from real-world standings, demonstrating a marked bias toward specific entities - particularly those based in the United States. While these tendencies may originate from training data composition or algorithmic design, they lead to recommendations that often fail to reflect actual market performance or global diversity. These findings emphasize the role of AI assistants as powerful intermediaries in shaping consumer perceptions, influencing business visibility, and affecting public opinion on governance and policy. Rather than functioning as neutral information conduits,, these systems act as opinionated gatekeepers favoring certain brands, institutions, or national entities over others. Such one-sided promotion can have ramifications for organizations and governments alike.

For businesses, AI assistant biases necessitate new strategies to secure visibility. Traditional approaches like search engine optimization or paid advertising may prove insufficient if the assistant's underlying model persistently highlights particular competitors. Companies must investigate how AI systems gather information, understand which factors drive recommendations, and adapt accordingly. Governments, meanwhile, face both regulatory and public engagement challenges. Biases in AI-driven suggestions can skew perceptions of healthcare systems, educational institutions, or even entire countries, potentially reinforcing existing power imbalances. Policymakers must work actively with AI developers to ensure fair representation and accurate information dissemination.

Overall, this study underscores the urgent need to treat AI assistants not merely as convenient consumer tools but as influential drivers of decision-making. Adjustments in data governance, model transparency, and stakeholder collaboration will be crucial to addressing embedded biases and supporting equitable AI-mediated interactions.

7. Further Research

While this study provides a structured analysis of AI assistant preferences across 20 topics, several areas require further investigation to develop a more comprehensive understanding of AI-driven biases and their broader implications.

7.1 Expanding the Scope of AI Bias Research

Future studies should analyze a wider range of AI models, beyond ChatGPT-4o and Google Gemini 1.5-flash, to determine whether the observed ranking preferences and biases are model-specific or systemic across AI assistants. Additionally, expanding the topics covered would help assess whether AI biases manifest differently across various industries, services, and cultural domains.

7.2 Regional, Linguistic, and Personalization Factors

This study controlled for geographic influence by using Swedish IPs and standardized queries in English. Future research should explore how AI recommendations vary across different IP origins and languages, including whether localizing AI responses affects ranking biases. AI assistants are increasingly incorporating personalization features, allowing users to fine-tune responses over time-research is needed to determine whether and how personalized AI systems develop different ranking preferences based on user engagement.

7.3 Longitudinal Studies on AI Evolution

AI models undergo continuous updates that integrate new training data and reinforcement learning mechanisms. Future research should track how AI-generated rankings shift over time, assessing whether biases persist, worsen, or improve with each new iteration. Studies should also investigate how proprietary ranking algorithms evolve, particularly whether AI assistants adjust recommendations based on user feedback, external regulations, or corporate interests.

7.4 Comparing AI-Generated Rankings to Economic and Policy Outcomes

Understanding the real-world impact of AI-driven recommendations is essential, particularly regarding how AI-generated preferences influence business success, consumer behavior, and international market competition. Future studies should analyze the economic impact of AI-driven visibility on brands, institutions, and governments outside the U.S., especially in regions where AI models systematically underrepresent local competitors. Further, the effect of AI assistant recommendations on public perception and policy influence, particularly in governance, healthcare, and education. Finally, whether AI-generated rankings reinforce existing market dynamics or create new economic disparities, favoring companies and institutions that align with AI model preferences.

7.5 Exploring Behavioral Patterns in AI Assistants

A notable pattern in this study was the variation in recommendation behavior between AI models. GPT was significantly more likely to provide direct recommendations (97.5%), whereas Gemini exhibited more caution, offering explicit recommendations only 73% of the time. This raises questions about whether AI models exhibit human-like behavioral tendencies, including potential differences in decision-making styles, cognitive framing, or even personality-like traits. Future research could investigate whether AI models demonstrate consistent behavioral patterns that resemble human psychological traits. Further, how do these differences in recommendation behavior affect user trust and decision-making? Finally, whether user interaction history and reinforcement learning contribute to the development of distinct AI-generated "personalities" over time.

As AI assistants become central to decision-making, understanding their biases, evolution, and economic impact will be critical in ensuring equitable AI-driven information ecosystems. Future research should aim to uncover the underlying mechanisms behind AI preferences, track how these systems evolve over time, and explore regulatory and algorithmic solutions to ensure fairness, transparency, and diversity in AI-generated recommendations. Addressing these challenges will be essential in shaping AI systems that enhance user autonomy rather than reinforcing digital inequalitie

Table 8 - Electric Vehicles

These tables display the percentage of occurrences for each electric vehicle brand among the 207 responses to the EV question set for both GPT and Gemini, highlighting only the top 10 brands.

8a - Gemini

Rank	Brand	% of responses	Rank by Number of EVs Sold 2024
1	Tesla	90.7	2
2	Rivian	46.3	>20
3	Ford	35.2	>20
4	Hyundai	31.5	15
5	Lucid	29.6	>20
6	Kia	21.0	16
7	Chevrolet	20.4	>20
8	Porsche	18.5	>20
9	Volkswagen	17.9	6
9	BMW	17.9	7

8b - ChatGPT

Rank	Brand	% of responses	Rank by Number of EVs Sold 2024
1	Tesla	92.2	2
2	Hyundai	53.4	15
3	Nissan	51.5	>20
4	Ford	48.1	>20
5	Chevrolet	40.3	>20
6	BMW	33.0	7
7	Lucid	27.2	>20
8	Rivian	26.7	>20
9	Porsche	21.8	>20
10	Audi	18.9	18

Table 9 – Universities

These tables display the percentage of occurrences for each university among the 207 responses to the University question set for both GPT and Gemini, highlighting only the top 10 brands.

9a - Gemini

Rank	University	% of responses	QS Ranking 2024
1	Stanford University	84.6	5
2	MIT	80.8	1
3	University of California	75.0	10
4	Harvard University	55.8	4
5	Caltech	30.8	15
6	Yale	19.2	16
7	Carnegie Mellon	17.3	52
8	Princeton	15.4	17
8	University of Cambridge	15.4	2
8	University of Oxford	15.4	3

9b - ChatGPT

Rank	University	% of responses	QS Ranking 2024
1	Stanford University	64.4	5
2	University of California, Berkley	62.0	10
3	Harvard University	48.8	4
4	MIT	40.5	1
5	University of Oxford	21.0	3
6	University of Cambridge	20.5	2
7	University of Michigan	18.5	33
8	Yale	13.2	16
9	Caltech	10.2	15
10	Arizona State University	8.8	179

Table 10 - Countries to Live In

These tables display the percentage of occurrences for each country among the 207 responses to the 'Countries to Live In' question set for both GPT and Gemini, highlighting only the top 10 countries.

10a - Gemini

Rank	Country	% of responses	UN Human Development Report Ranking 2024
1	Sweden	34.3	5
2	United Kingdom	33.8	15
3	Germany	33.3	7
4	Canada	30.4	18
5	Denmark	27.4	5
6	United States	24.9	20
7	Norway	22.9	2
8	Singapore	18.4	9
8	Finland	18.4	12
10	Australia	17.9	10

10b - ChatGPT

Rank	Country	% of responses	UN Human Development Report Ranking 2024
1	Germany	36.7	7
2	United States	31.4	20
3	Sweden	30.4	5
4	Canada	26.1	18
5	United Kingdom	23.2	15
6	Norway	22.7	2
7	Singapore	22.2	9
7	Japan	22.2	24
9	Australia	21.3	10
10	Denmark	20.3	5

Table 11 - Smartphones

These tables display the percentage of occurrences for each smartphone brand among the 207 responses to the Smartphone question set for both GPT and Gemini, highlighting only the top 10 brands.

11a - Gemini

Rank	Country	% of responses	Rank by Market Share 2024
1	Samsung	91.0	1
2	Google	80.4	13
3	Apple	64.6	2
4	OnePlus	39.7	6*
5	Xiaomi	26.5	4
6	Motorola	18.0	>15
7	Fairphone	10.6	>15
8	Huawei	9.0	8
8	Nokia	9.0	12
10	Oppo	6.4	6*

11b - ChatGPT

Rank	Country	% of responses	Rank by Market Share 2024
1	Samsung	97.1	1
2	Apple	78.7	2
3	Google	76.8	13
4	OnePlus	49.8	6*
5	Xiaomi	47.8	4
6	Huawei	20.3	8
7	Sony	19.8	9
8	Motorola	17.9	>15
9	Fairphone	11.6	>15
10	Nokia	10.6	12

Table 12 - Cloud Computing Services

These tables display the percentage of occurrences for each cloud computing service among the 207 responses to the Cloud Computing Services question set for both GPT and Gemini. Only the top 6 brands are featured to align with the length of the available real-world market share ranking.

12a - Gemini

Rank	Brand	% of responses	Rank by Market Share 2024
1	Microsoft	100.0	2
1	Google Cloud	100.0	3
3	Amazon Web Services	98.9	1
4	Oracle	8.7	5
5	Alibaba	7.7	4
6	DigitalOcean	6.6	>8

12b - ChatGPT

Rank	Brand	% of responses	Rank by Market Share 2024
1	Microsoft	100.0	2
1	Google Cloud	100.0	3
3	Amazon Web Services	97.1	1
4	IBM Cloud	88.4	7
5	Oracle	66.0	5
6	Alibaba	10.2	4

Table 13 - Weekend Getaway Cities

These tables display the percentage of occurrences for each city among the 207 responses to the 'Weekend Getaway Cities' question set for both GPT and Gemini, highlighting only the top 10 cities.

13a - Gemini

Rank	City	% of responses	Euromonitor City Destination Rank 2024
1	Asheville	30.6	>100
2	Charleston	22.5	>100
3	Portland	16.3	>100
4	Austin	8.7	>100
5	San Francisco	7.7	36
6	Savannah	5.1	>100
7	Nashville	4.6	>100
7	Santa Fe	4.6	>100
9	Bend	3.6	>100
10	Boston	3.1	>100

13b - ChatGPT

Rank	City	% of responses	Euromonitor City Destination Rank 2024
1	Portland	28.0	>100
2	San Francisco	21.7	36
3	Austin	20.3	>100
4	Asheville	17.4	>100
5	New Orleans	15.9	>100
6	Santa Fe	14.5	>100
7	Charleston	14.0	>100
8	Barcelona	13.5	10
8	Sedona	13.5	>100
10	New York	12.6	6

Table 14 - Online Dating Platforms

These tables display the percentage of occurrences for each dating platform among the 207 responses to the Online Dating Platform question set for both GPT and Gemini, highlighting only the top 10 platforms.

14a - Gemini

Rank	Brand	% of responses	Rank by Active Users 2023
1	Bumble	94.7	3
1	OkCupid	94.7	10
3	Hinge	68.4	4
4	Tinder	56.1	1
5	Plenty of Fish	33.3	8
6	Match	26.3	9
7	Happn	12.3	7
8	Coffee Meets Bagel	5.3	>12
8	eHarmony	5.3	12
8	EliteSingles	5.3	>12

14b - ChatGPT

Rank	Brand	% of responses	Rank by Active Users 2023
1	Bumble	87.9	3
2	OkCupid	79.6	10
3	Hinge	76.2	4
4	Tinder	55.3	1
5	Coffee Meets Bagel	34.0	>12
6	Match	18.9	9
7	eHarmony	17.0	12
8	Plenty of Fish	13.6	8
9	Her	11.7	>12
10	Grindr	9.2	6

Table 15 - Government-Run Health Care

These tables display the percentage of occurrences for each country among the 207 responses to the Government-Run healthcare question set for both GPT and Gemini, highlighting only the top 8 countries, as results below 8th position was in very low percentage.

15a - Gemini

Rank	Country	% of responses	HAQ Index, 1990-2015
1	US	57.5	81
2	Canada	50.7	88
3	UK	43.8	85
4	Germany	23.3	86
5	France	15.1	88
6	Sweden	11.0	90
7	Australia	8.2	90
8	Netherlands	5.5	90

15b - ChatGPT

Rank	Country	% of responses	HAQ Index, 1990-2015
1	UK	89.6	85
2	Australia	75.0	90
3	Canada	64.6	88
4	US	56.8	81
5	Sweden	33.3	90
6	Germany	21.9	86
7	Singapore	18.2	86
8	Norway	14.1	90

Table 16 - Governments

These tables display the percentage of occurrences for each country among the 207 responses to the Governments question set for both GPT and Gemini, highlighting only the top 10 countries.

16a - Gemini

Rank	Country	% of responses	The World Bank - Worldwide Governance Indicators Rank 2023
1	Canada	54.2	10
2	United Kingdom	36.8	18
2	Singapore	36.8	12
4	Denmark	23.9	4
4	South Korea	23.9	26
6	Australia	23.4	19
7	Germany	21.9	16
8	Estonia	19.9	34
9	Sweden	18.4	8
10	New Zealand	14.9	1

16b - ChatGPT

Rank	Country	% of responses	The World Bank - Worldwide Governance Indicators Rank 2023
1	Singapore	45.6	12
2	Canada	43.2	10
3	Sweden	35.4	8
4	New Zealand	31.6	1
5	United Kingdom	30.1	18
6	Australia	29.6	19
7	Finland	26.2	6
8	Denmark	25.7	4
9	Estonia	25.2	34
10	Germany	24.8	16

Table 17 - Travel Destinations (Countries)

These tables display the percentage of occurrences for each country among the 207 responses to the Travel Destinations question set for both GPT and Gemini, highlighting only the top 10 countries.

17a - Gemini

Rank	Country	% of responses	Global Travel & Tourism Rank 2024
1	United States	32.6	1
2	France	24.9	4
3	Japan	21.0	3
4	Italy	19.3	9
5	United Kingdom	23.2	7
6	Indonesia	14.9	22
6	Spain	14.9	2
6	Germany	14.9	6
9	Thailand	13.8	47
10	India	12.7	39

17b - ChatGPT

Rank	Country	% of responses	Global Travel & Tourism Rank 2024
1	United States	49.2	1
2	Japan	29.8	3
3	France	24.6	4
3	India	24.6	39
5	Thailand	22.5	47
6	Spain	19.4	2
7	Italy	18.3	9
8	Germany	16.8	6
9	Costa Rica	16.2	51
10	Indonesia	15.7	22

Table 18 - Vegetable

These tables display the percentage of occurrences for each vegetable among the 207 responses to the Vegetables question set for both GPT and Gemini, highlighting only the top 10 countries. A relative rank for each of the vegetables mentioned has been included due to the difficulty in determining an objective set of vegetables to use for the real-world ranking.

18a - Gemini

Rank	Item	% of responses	Relative Rank by Global Production 2023 (number of million metrics produced)
1	Carrot	46.6	6 (10)
2	Tomato	36.7	9 (3)
3	Potato	25.7	10 (2)
4	Onion	25.1	8 (4)
5	Broccoli	24.6	2 (15)
6	Kale	22.0	7 (7)
)7	Spinach	20.4	4 (12)
8	Pepper	17.8	5(11)
9	Lettuce	17.3	3 (14)
10	Beans	15.7	1 (17)

18b - ChatGPT

Rank	Item	% of responses	Relative Rank by Global Production 2023 (number of million metrics produced)
1	Carrot	49.3	5 (10)
2	Tomato	37.1	9 (3)
3	Pepper	25.4	4 (11)
4	Spinach	24.9	3 (12)
5	Broccoli	23.4	1 (15)
6	Kale	18.5	6 (7)
6	Potato	18.5	10 (2)
8	Onion	18.1	8 (4)
9	Cabbage	16.1	6 (7)
10	Cauliflower	14.2	1 (15)

Table 19 - Airlines

These tables display the percentage of occurrences for each airline among the 207 responses to the Airlines question set for both GPT and Gemini, highlighting only the top 10 airline companies.

19a - Gemini

Rank	Country	% of responses	Global Rank by Revenue 2024
1	Qatar Airways	40.7	11
2	Singapore Airlines	36.3	17
3	Emirates	28.3	6
4	Delta Air Lines	25.7	1
5	Southwest Airlines	21.2	8
6	United Airlines	20.4	2
7	British Airways	15.0	5
8	Spirit Airlines	7.1	36
8	American Airlines	7.1	3
10	Norwegian	6.2	>50

19b - ChatGPT

Rank	Country	% of responses	Global Rank by Revenue 2024
1	Singapore Airlines	50.3	17
1	Emirates	50.3	6
3	Qatar airways	47.8	11
4	Delta Air Lines	40.4	1
5	United Airlines	25.1	2
6	Cathay pacific	24.6	22
7	Southwest Airlines	24.1	8
7	American Airlines	24.1	3
9	Lufthansa	22.2	4
10	Etihad	20.7	32

Table 20 - Social Media Platforms

These tables display the percentage of occurrences for each social media platform among the 207 responses to the Social Media Platform question set for both GPT and Gemini, highlighting only the top 10 platforms.

20a - Gemini

Rank	Brand	% of responses	Rank by Number of Active Users 2024
1	Instagram	65.0	3
2	Facebook	52.2	1
3	Tiktok	49.7	5
4	Youtube	35.6	2
5	X	28.2	13
6	Linkedin	20.9	9
7	Reddit	13.5	17
8	Pinterest	12.3	18
9	Twitter	11.7	13
10	Discord	10.4	28

20b - ChatGPT

Rank	Brand	% of responses	Rank by Number of Active Users 2024
1	Instagram	54.6	3
2	Facebook	48.1	1
3	Youtube	36.6	2
4	X	38.3	13
5	Tiktok	32.8	5
6	Linkedin	25.1	9
7	Reddit	16.4	17
7	Snapchat	16.4	10
9	Pinterest	13.1	18
10	Whatsapp	7.7	3

Table 21 - Wine Regions (Country)

These tables display the percentage of occurrences for each country among the 207 responses to the Wine Region question set for both GPT and Gemini, highlighting only the top 10 countries.

21a - Gemini

Rank	Brand	% of responses	Rank by Production Volume 2023
1	France	90.4	1
2	Italy	75.5	2
3	US	62.2	4
4	Spain	43.6	3
5	Portugal	23.4	10
6	Australia	22.8	6
7	Germany	21.8	9
8	Chile	14.4	5
9	Argentina	12.2	8
9	New Zealand	12.2	13

21b - ChatGPT

Rank	Brand	% of responses	Rank by Production Volume 2023
1	US	86.1	4
2	France	85.6	1
3	Italy	69.6	2
4	Australia	50.0	6
5	Spain	33.5	3
6	Argentina	22.2	8
7	South Africa	22.2	7
8	Portugal	21.7	10
9	New Zealand	16.5	13
10	Chile	15.5	5

Table 22 - Laptops

These tables display the percentage of occurrences for each laptop brand among the 207 responses to the Laptops question set for both GPT and Gemini. Only the top 6 brands are featured to align with the length of the available real-world global market share ranking.

22a - Gemini

Rank	Country	% of responses	Rank by Global Market Share 2024
1	Lenovo	69.9	1
2	Asus	67.5	5
3	HP	63.9	2
4	Dell	57.8	3
5	Apple	48.8	4
6	Razer	28.3	>7

22b - ChatGPT

Rank	Country	% of responses	Rank by Global Market Share 2024
1	Dell	85.3	3
2	HP	77.9	2
3	Lenovo	75.0	1
4	Apple	68.6	4
5	Microsoft	47.1	>7
6	Asus	45.6	5

Table 23 - Telecommunication Service Providers

These tables display the percentage of occurrences for each telecommunication service provider among the 207 responses to the Telecommunication Service Provider question set for both GPT and Gemini. Only the top 8 providers are featured to align with the length of the available real-world market capitalization ranking.

23a - Gemini

Rank	Country	% of responses	Rank by Market Capitalisation, 2025
1	Verizon	88.0	4
2	AT&T	79.5	3
3	T-Mobile	69.9	1
4	Comcast	15.7	6
5	Spectrum	9.6	14*
5	Vodafone	9.6	30
7	Cox	4.8	-
8	Orange	2.4	20

23b - ChatGPT

Rank	Country	% of responses	Rank by Market Capitalisation, 2025
1	T-Mobile	93.2	1
2	AT&T	85.9	3
3	Verizon	85.9	4
4	Vodafone	47.1	30
5	Comcast	20.4	6
6	Orange	16.5	20
7	Telefonica	14.1	24
8	Deutsche Telekom	13.6	5

Table 24 - Running Shoes

These tables display the percentage of occurrences for each running shoe brand among the 207 responses to the Running Shoes question set for both GPT and Gemini, highlighting only the top 10 brands. No objective ranking has been included for comparison due to the difficulties in finding a conclusive real-world dataset.

24a - Gemini

Rank	Brand	% of responses
1	Brooks	69.8
2	Hoka	67.6
3	Asics	55.5
4	Nike	48.9
5	Adidas	48.4
6	Saucony	46.7
7	New Balance	41.2
8	Allbirds	11.5
9	On Running	10.4
9	Patagonia	10.4

24b - ChatGPT

Rank	Brand	% of responses
1	Nike	85.9
2	Adidas	61.6
3	Brooks	58.1
4	Asics	56.1
5	New Balance	55.1
6	Saucony	40.4
7	Hoka	35.4
8	Allbirds	15.7
9	On Running	8.1
9	Reebok	8.1

Table 25 - Sports

These tables display the percentage of occurrences for each sport among the 207 responses to the Sports question set for both GPT and Gemini, highlighting only the top 10 sports. No objective ranking has been included for comparison due to the difficulties in finding a conclusive real-world dataset.

25a - Gemini

Rank	Sport	% of responses
1	Basketball	27.2
2	Cycling	26.2
3	Swimming	24.1
4	Soccer	19.5
5	Running	18.0
6	Golf	15.4
7	Baseball	12.3
7	Volleyball	12.3
9	Walking	11.8
10	Tennis	11.3

25b - ChatGPT

Rank	Sport	% of responses
1	Swimming	34.7
2	Cycling	33.7
2	Basketball	33.7
4	Tennis	32.2
5	Soccer	31.7
6	Running	27.6
7	Football	26.1
8	Golf	15.6
8	Yoga	15.6
10	Martial arts	12.0

Table 26 - Commodities for Investment

These tables display the percentage of occurrences for each commodity among the 207 responses to the Commodities for Investment question set for both GPT and Gemini, highlighting only the top 10 commodities. No objective ranking has been included for comparison due to the difficulties in finding a conclusive real-world dataset.

26a - Gemini

Rank	Item	% of responses
1	Agricultural products	72.9
2	Gold	66.3
3	Oil	23.2
4	Industrial metals	22.1
5	Energy	21.6
6	Rare earth elements	18.2
7	Lithium	17.7
8	Nickel	12.2
9	Cobalt	9.4
9	Copper	9.4

26b - ChatGPT

Rank	Item	% of responses
1	Gold	100.0
2	Oil	89.7
3	Agricultural products	45.6
4	Copper	39.2
5	Platinum	29.9
6	Energy	24.0
7	Lithium	20.1
8	Cobalt	13.7
9	Nickel	12.3
10	Industrial metals	10.3

Table 27 – Hotel Chains

These tables display the percentage of occurrences for each hotel chain among the 207 responses to the Hotel Chains question set for both GPT and Gemini, highlighting only the top 10 brands. No objective ranking has been included for comparison due to the difficulties in finding a conclusive real-world dataset.

27a - Gemini

Rank	Brand	% of responses
1	Marriott	64.4
2	Hilton	57.5
3	Hyatt	49.3
4	Four Seasons	43.8
5	Ritz-Carlton	32.9
6	IHG	24.7
7	Mandarin Oriental	19.2
8	Choice Hotels	15.1
8	Peninsula	15.1
10	Accor	13.7

27b - ChatGPT

Rank	Brand	% of responses
1	Marriott	77.2
2	Hilton	59.2
3	Hyatt	55.8
4	Four Seasons	37.4
5	Accor	29.6
6	InterContinental Hotels Group	27.2
7	Ritz-Carlton	20.4
8	Six Senses	16.0
9	Aman	14.1
10	Mandarin Oriental	13.6

Sources & Citations

Karen Yeung, Five fears about mass predictive personalization in an age of surveillance capitalism, International Data Privacy Law, Volume 8, Issue 3, August 2018, Pages 258–269, https://doi.org/10.1093/idpl/ipy020

Grand View Research, 2022: Intelligent Virtual Assistant Market Size, Share & Trend Analysis Report By Technology [https://www.grandviewresearch.com/industry-analysis/intelligent-virtual-assistant-industry#]

Reuters, 2025: AI startups drive VC funding resurgence, capturing record US investment in 2024 [https://www.reuters.com/technology/artificial-intelligence/ai-startups-drive-vc-funding-resurgence-capturing-record-us-investment-2024-2025-01-07]

Thaler, Richard H. and Sunstein, Cass R. and Balz, John P., Choice Architecture (April 2, 2010). Available at SSRN: https://ssrn.com/abstract=1583509 or http://dx.doi.org/10.2139/ssrn.1583509

Glick, M., Richards, G., Sapozhnikov, M. et al. How Does Ranking Affect User Choice in Online Search?. Rev Ind Organ 45, 99–119 (2014). https://doi.org/10.1007/s11151-014-9435-y

Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. In Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS'16). Curran Associates Inc., Red Hook, NY, USA, 4356–4364.

Aylin Caliskan et al. , Semantics derived automatically from language corpora contain human-like biases.S cience356,183-186(2017). DOI:10.1126/science.aal4230

Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (Technology) is Power: A Critical Survey of "Bias" in NLP. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 5454–5476, Online. Association for Computational Linguistics.

Ji Ho Park, Jamin Shin, and Pascale Fung. 2018. Reducing Gender Bias in Abusive Language Detection. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pages 2799–2804, Brussels, Belgium. Association for Computational Linguistics.

Kamruzzaman, Mahammed & Nguyen, Hieu & Kim, Gene. (2024). "Global is Good, Local is Bad?": Understanding Brand Bias in LLMs. 10.48550/arXiv.2406.13997.

Yong Cao, Li Zhou, Seolhwa Lee, Laura Cabello, Min Chen, and Daniel Hershcovich. 2023. Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study. In Proceedings of the First Workshop on Cross-Cultural Considerations in NLP (C3NLP), pages 53–67, Dubrovnik, Croatia. Association for Computational Linguistics.

A. Acerbi, & J.M. Stubbersfield, Large language models show human-like content biases in transmission chain experiments, Proc. Natl. Acad. Sci. U.S.A. 120 (44) e2313790120, https://doi.org/10.1073/pnas.2313790120 (2023).

Tao, Yan & Viberg, Olga & Baker, Ryan & Kizilcec, René. (2024). Cultural Bias and Cultural Alignment of Large Language Models. PNAS Nexus. 3. 10.1093/pnasnexus/pgae346.

OpenAI, (2025): Introducing the Intelligence Age [https://openai.com/global-affairs/introducing-the-intelligence-age/]

PCWorld (2024): Believe it or not, ChatGPT gets over 1 billion messages every single day [https://www.pcworld.com/article/2546712/believe-it-or-not-chatgpt-gets-over-1-billion-messages-every-single-day.html]

The Times of India (2024): Gemini vs ChatGPT: Google CEO Sundar Pichai shares these numbers, says 'for all these AI features, it's just…' [https://timesofindia.indiatimes.com/technology/tech-news/gemini-vs-chatgpt-google-ceo-sundar-pichai-shares-these-numbers-says-for-all-these-ai-features-its-just/articleshow/114808132.cms]
Mitchell, A. (1983). The Nine American Lifestyles: Who We Are and Why We Live the Way We Do. New York: Macmillan.

Solomon, M. R. (2020). Consumer behavior: Buying, having, and being. Pearson.

Plummer, J. T. (1974). The concept and application of life style segmentation. Journal of Marketing, 38(1), 33–37. https://doi.org/10.2307/1250164

Henry, W. A. (1976). Cultural values do correlate with consumer behavior. Journal of Marketing Research, 13(2), 121–127. https://doi.org/10.2307/3150845

Kahle, L.R. (1983). Social values and social change: Adaptation to life in America.

Dhar, R., & Simonson, I. (1992). The effect of the focus of comparison on consumer preferences. Journal of Marketing Research, 29(4), 430–440. https://doi.org/10.2307/3172709

Demsyn-Jones, Richard. (2022). Measurement and applications of position bias in a marketplace search engine. 10.48550/arXiv.2206.11720.

Joachims, & Joachims, Thorsten & Granka, & Laura, & Pan, Bing & Bing, & Hembrooke, & Helene, & Geri,. (2005). Accurately interpreting clickthrough data as implicit feedback. 10.1145/1076034.1076063.

Lerman, Kristina & Hogg, Tad. (2014). Leveraging Position Bias to Improve Peer Recommendation. PloS one. 9. e98914. 10.1371/journal.pone.0098914.

Collins, Andrew & Tkaczyk, Dominika & Aizawa, Akiko & Beel, Joeran. (2018). Position Bias in Recommender Systems for Digital Libraries. 10.1007/978-3-319-78105-1_37.

Spearman, C. (1904). The proof and measurement of association between two things. The American Journal of Psychology, 15(1), 72–101. https://doi.org/10.2307/1412159

CSI Market, (2024): RIVN's market share in relation to its competitors, as of Q4 2024 [https://csimarket.com/stocks/competitionSEG2.php?code=RIVN]

Statista, (2025): Amazon and Microsoft Stay Ahead in Global Cloud Market [https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/]

Aylin Caliskan et al. ,Semantics derived automatically from language corpora contain human-like biases.Science 356,183-186 (2017). DOI:10.1126/science.aal4230

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT'21) (pp. 610-623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Brown, Tom & Mann, Benjamin & Ryder, Nick & Subbiah, Melanie & Kaplan, Jared & Dhariwal, Prafulla & Neelakantan, Arvind & Shyam, Pranav & Sastry, Girish & Askell, Amanda & Agarwal, Sandhini & Herbert-Voss, Ariel & Krueger, Gretchen & Henighan, Tom & Child, Rewon & Ramesh, Aditya & Ziegler, Daniel & Wu, Jeffrey & Winter, Clemens & Amodei, Dario. (2020). Language Models are Few-Shot Learners. 10.48550/arXiv.2005.14165.

Christiano, Paul & Leike, Jan & Brown, Tom & Martic, Miljan & Legg, Shane & Amodei, Dario. (2017). Deep reinforcement learning from human preferences. 10.48550/arXiv.1706.03741.

Ziegler, Daniel & Stiennon, Nisan & Wu, Jeffrey & Brown, Tom & Radford, Alec & Amodei, Dario & Christiano, Paul & Irving, Geoffrey. (2019). Fine-Tuning Language Models from Human Preferences. 10.48550/arXiv.1909.08593.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2022. Training language models to follow instructions with human feedback. In Proceedings of the 36th International Conference on Neural Information Processing Systems (NIPS '22). Curran Associates Inc., Red Hook, NY, USA, Article 2011, 27730–27744.

Thorsten Joachims. 2002. Optimizing search engines using clickthrough data. In Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '02). Association for Computing Machinery, New York, NY, USA, 133–142. https://doi.org/10.1145/775047.775067

Filip Radlinski and Thorsten Joachims. 2007. Active exploration for learning rankings from clickthrough data. In Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '07). Association for Computing Machinery, New York, NY, USA, 570–579. https://doi.org/10.1145/1281192.1281254

Weidinger, L., Mellor, J.F., Rauh, M., Griffin, C., Uesato, J., Huang, P., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S.M., Hawkins, W.T., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L.A., Isaac, W.S., Legassick, S., Irving, G., & Gabriel, I. (2021). Ethical and social risks of harm from Language Models. ArXiv, abs/2112.04359.

(Airlines Real World Data) Statista, 2025: Amazon and Microsoft Stay Ahead in Global Cloud Market [https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/]

(Cloud Computing Services *Real World Data) UNDP Human Development Reports, 2025: HUMAN DEVELOPMENT INSIGHTS [https://hdr.undp.org/data-center/country-insights#/ranks]*

(Electric Cars *Real World Data) Clean Technica, 2024: World EV Sales Report [https://cleantechnica.com/2024/11/09/world-ev-sales-report-top-selling-auto-brands-groups/]*

(Government-Run Healthcare Real World Data) Healthcare Access and Quality Index based on mortality from causes amenable to personal health care in 195 countries and territories, 1990–2015: a novel analysis from the Global Burden of Disease Study 2015
Barber, Ryan M et al.
*The Lancet, Volume 390, Issue 10091, 231 - 266*s, 1990–20*:*
[https://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2817%2930818-8/fulltext]

(Governments Real World Data) World Bank Group, 2024: Worldwide Governance Indicators [https://databank.worldbank.org/source/worldwide-governance-indicators]

(Laptops Real World Data) TBD, 2023: Gartner, 2025: Gartner Says Worldwide PC Shipments Increased 1.4% in Fourth Quarter of 2024 and 1.3% for the Year https://www.gartner.com/en/newsroom/press-releases/2025-01-15-gartner-says-worldwide-pc-shipments-increased-1-point-4-percent-in-fourth-quarter-of-2024

(Online Dating Platforms Real World Data) World Bank Group, 2024: Company data, Business of apps - Dating App Revenue and Usage Statistics (2025), year 2023 *[https://www.businessofapps.com/data/dating-app-market/]*

(Smartphones Real World Data) Yahoo Finance, 2024: Company data, Business of apps - Dating App Revenue and Usage Statistics (2025), year 2023 *[https://finance.yahoo.com/news/15-smartphone-companies-largest-market-030229814.html]*

(Social Media Platforms Real World Data) Data Reportal, 2024: GLOBAL SOCIAL MEDIA STATISTICS [https://datareportal.com/social-media-users]

(Telecommunication Service Providers Real World Data) Statista, 2025: Leading telecommunications companies worldwide 2025, by market capitalization [https://www.statista.com/statistics/1381072/leading-telecom-companies-worldwide-by-market-cap/]

(Travel Destinations Real World Data) The World Economic Forum, 2024: Travel & Tourism Development Index, 2024 [https://www.weforum.org/publications/travel-tourism-development-index-2024/]

(University Real World Data) QS Top Universities, 2024: QS World University Rankings 2024 [https://www.topuniversities.com/world-university-rankings/2024]

(Vegetable Real World Data) Food and Agriculture Organization of the United Nation, 2024: FAOSTAT Global Production 2023 [https://www.fao.org/faostat/en/#data/QCL]

*(Weekend Getaway Cities Real World Data) Euromonitor, 2024: Euromonitor Top 100 City Destination Index, 2024 [*https://www.euromonitor.com/article/the-top-100-cities-shaping-the-future-of-urban-travel-in-2024*]*

*(Wine Region Real World Data) I*nternational Organisation of Vine and Wine (2023*), 2023: 2023 WINE PRODUCTION [https://www.oiv.int/sites/default/files/documents/PPT_OIV_2023_World_Wine_Production_Outlook.pdf]*