

Beyond Search

LLM Adoption and Web Traffic Concentration

Samira Gholami, *Stanford University*

Cristiana Firullo, *Cornell University*

Cristobal Cheyre, *Cornell University*

Alessandro Acquisti, MIT Sloan

Prepared for the Digital Competition Conference at KGI, February 2026

This work was supported by the Alfred P. Sloan Foundation, Carnegie Mellon University Block Center for Technology and Society, Carnegie Mellon University CyLab Security and Privacy Institute, the Institute of Humane Studies and the Program on the Economics & Privacy at George Mason University, the University of Pennsylvania Center for Technology, Innovation and Competition and the Warren Center for Network & Data Sciences, the Pittsburgh Supercomputing Center, the NET Institute, the Cornell Center for Social Sciences, and the National Science Foundation (awards 2237327/2237328/2237329).

Motivation

- With LLMs available to the public, there is concern they might reshape how people navigate the internet.

THE SHIFT

The New York Times

Can Google Give A.I. Answers Without Breaking the Web?

Publishers have long worried that artificial intelligence would drive readers away from their sites. They're about to find out if those fears are warranted.

THE WALL STREET JOURNAL.

TECHNOLOGY | ARTIFICIAL INTELLIGENCE

News Sites Are Getting Crushed by Google's New AI Tools

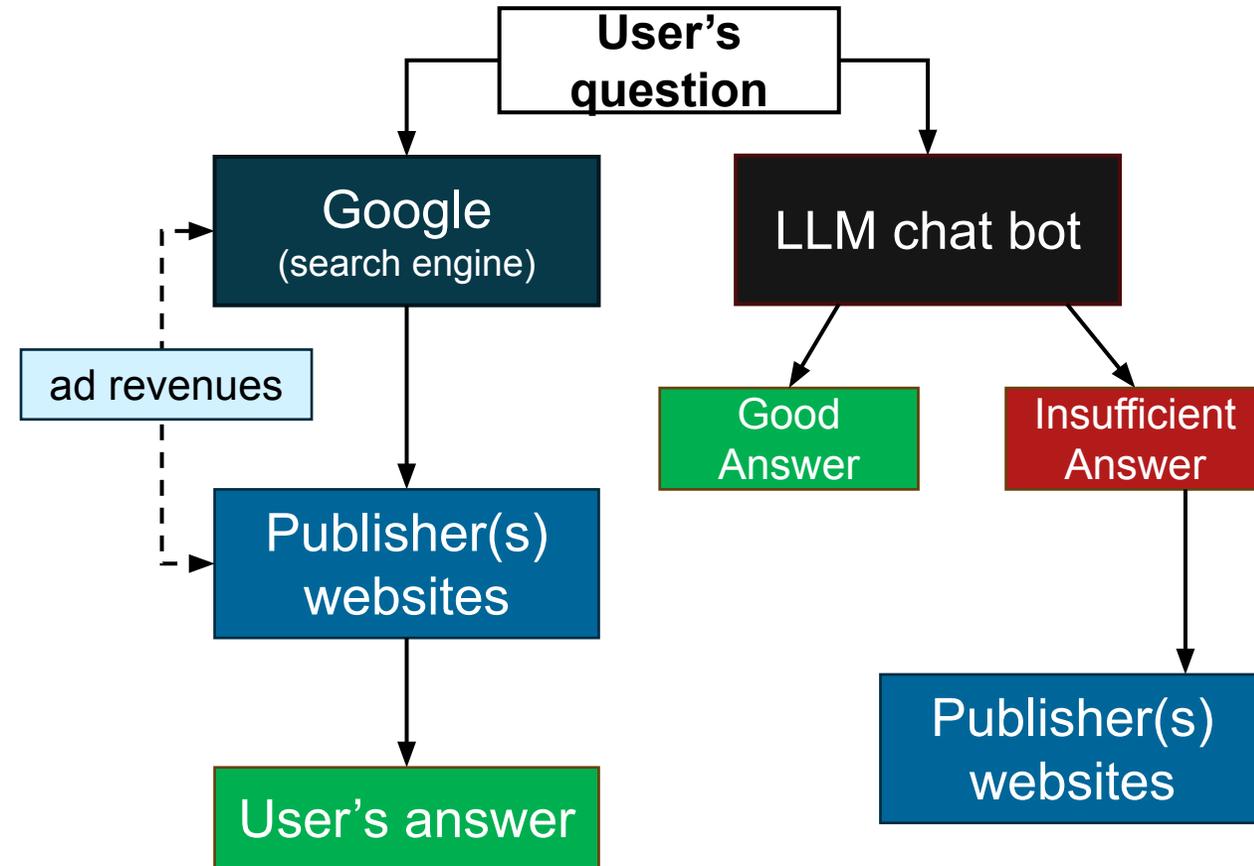
Chatbots are replacing Google's traditional search, devastating traffic for some publishers

MIT
Technology
Review

AI means the end of internet search as we've known it

Motivation

- With LLMs available to the public, there is concern this might reshape how people navigate the internet.
- higher usage of LLMs might:
 - lower visits to publishers
 - lower ad traffic and revenues for publishers / advertisers / networks.
- This threatens the profitability and existence of publishers and smaller players in the internet ecosystem.



This paper

- Analyzes how people use **search engines** and **LLMs**.
- Studies the effect of LLM adoption and usage on browsing activity.
- Proposes a **taxonomy** of LLM conversations to compare with search engine queries □ defines “**user intent**”.
- Studies the effect on browsing activity based on intent and the platform used:
 - *unique domains/URLs visited per session;*
 - *concentration/diversity measures (HHI, Shannon Entropy and Externalization).*

Overview

- **Data**
- Empirical Facts
- Proposed Taxonomy
- Downstream Traffic Analysis

Data

Internet Behavior Experiment (2025 – ongoing)

- Timestamps of visited websites;
 - HTML of selected domains ;
 - **search engines and LLM queries;**
 - ...and more.
-

comScore (2019 – 2024)

- Large representative sample;
- timestamps of visited websites.

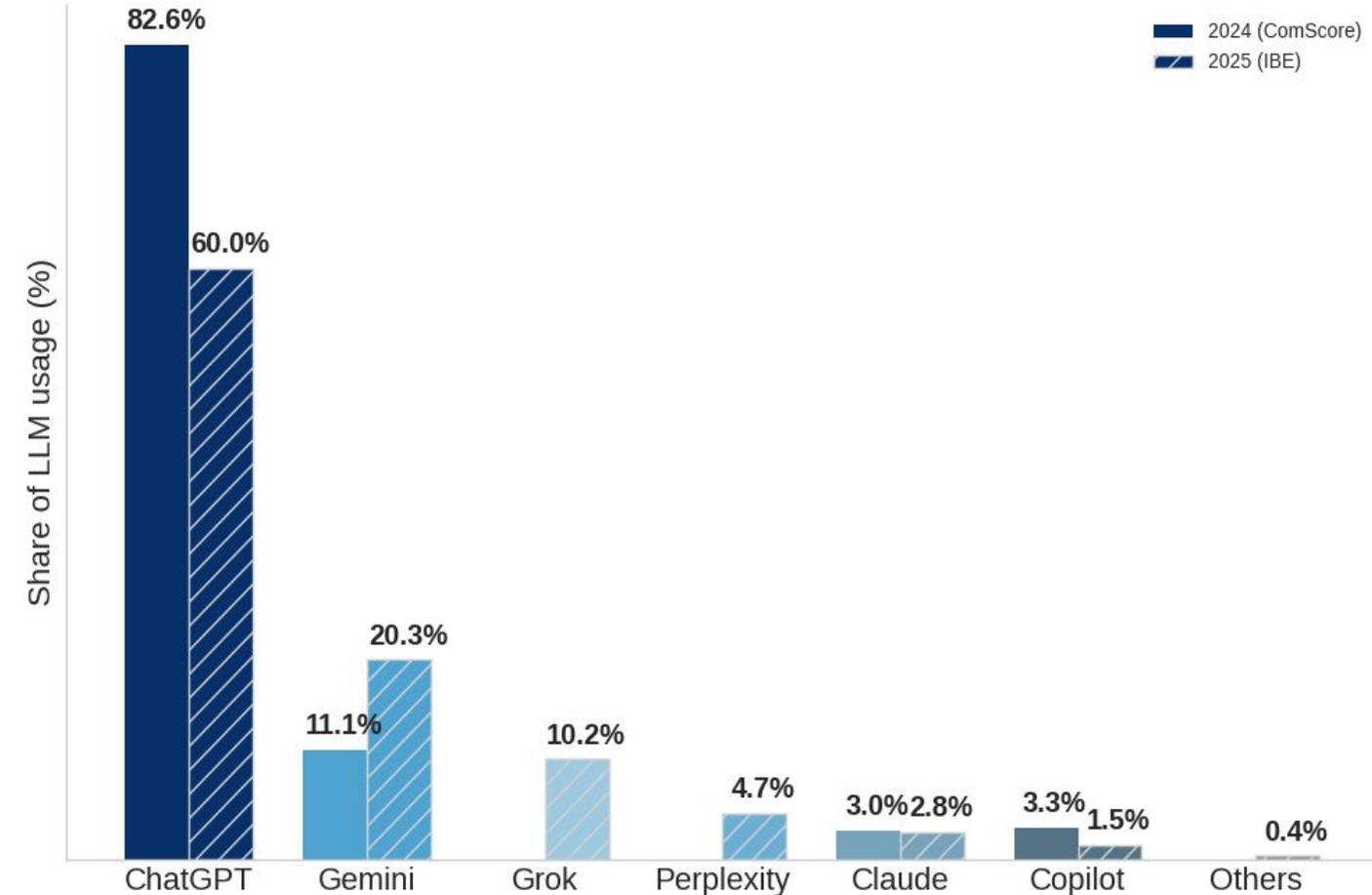
We use a combination of these
two datasets to study online
user behavior in relation to
LLMs

Overview

- Data
- **Empirical Facts**
- Proposed Taxonomy
- Downstream Traffic Analysis

Fact 1: the B2C LLM market is growing

Comparison of LLM services adopted by users in both samples



- ~**50%** of participants in the IBE data use LLMs. Of these, ~**40%** are *multi-homing*.
 - ChatGPT is the first LLM in the **15th website by number of weekly users**.
 - ChatGPT shares ↓ , while smaller players' quota is increasing compared to previous year.
- The B2C segment seems 'healthy' with **entry** and **multi-homing**.

Fact 2: LLM adoption is associated with more diverse browsing

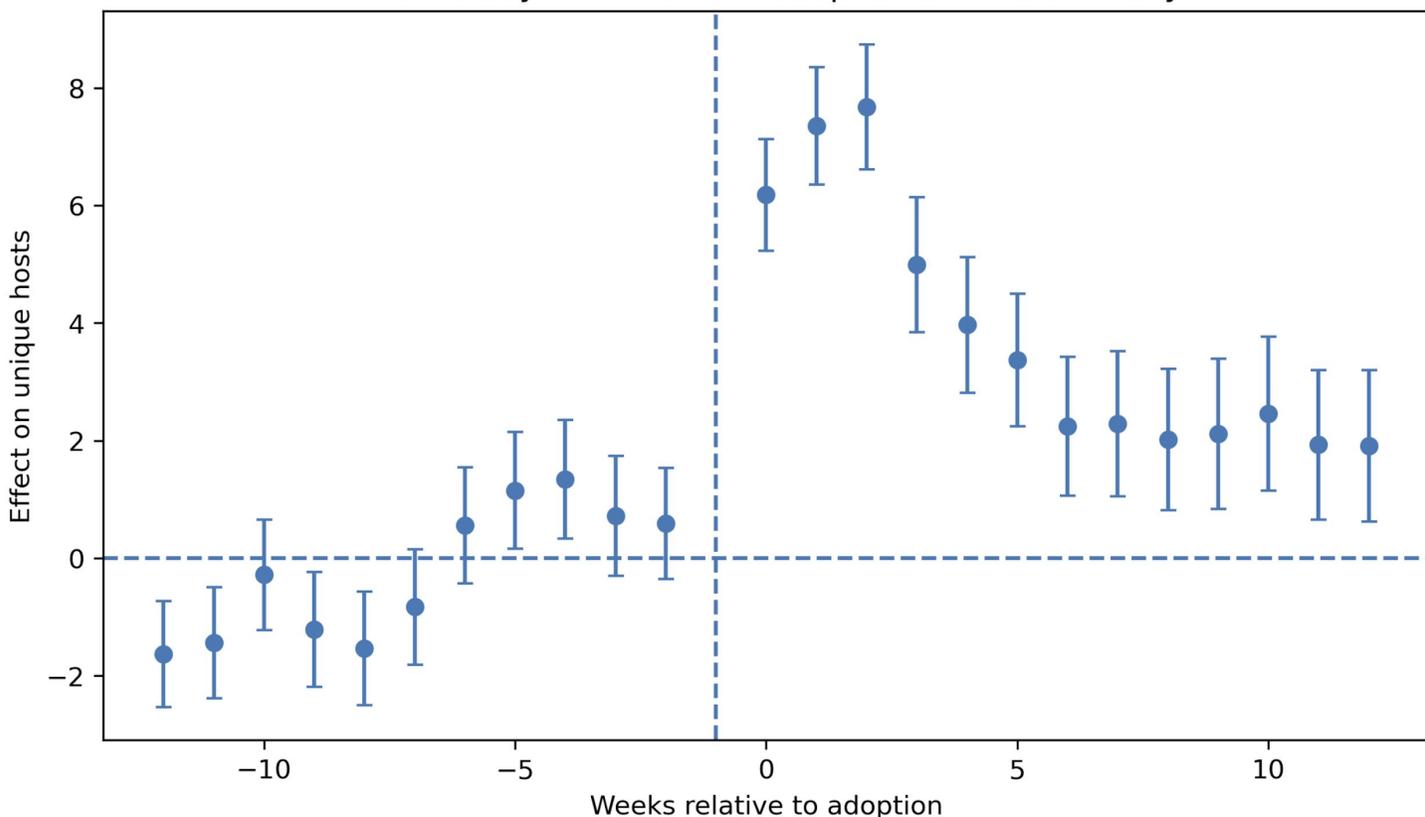
LLMs Adoption (comScore Data, 2022 – 2024)

$$\text{unique hosts (no search/LLMs)}_{i,t} = \sum_k 1\{\tau_{it} = k\} + \gamma n_{i,t} + \alpha_i + \lambda_t + \varepsilon_{i,t}$$

$\tau_{it} = k$
=1 if adopted

$n_{i,t}$ → number of websites

Event study: effect of LLM adoption on web diversity



- Adoption is defined as participant **uses LLMs for at least a 3-week period** (Padilla et al, 2025).
- Regression results show that adopters are **visiting more websites after adoption**.
- These results doesn't seem to support the traditional narrative of LLMs threatening browsing activity.

Fact 3: Adopters are not abandoning search engines

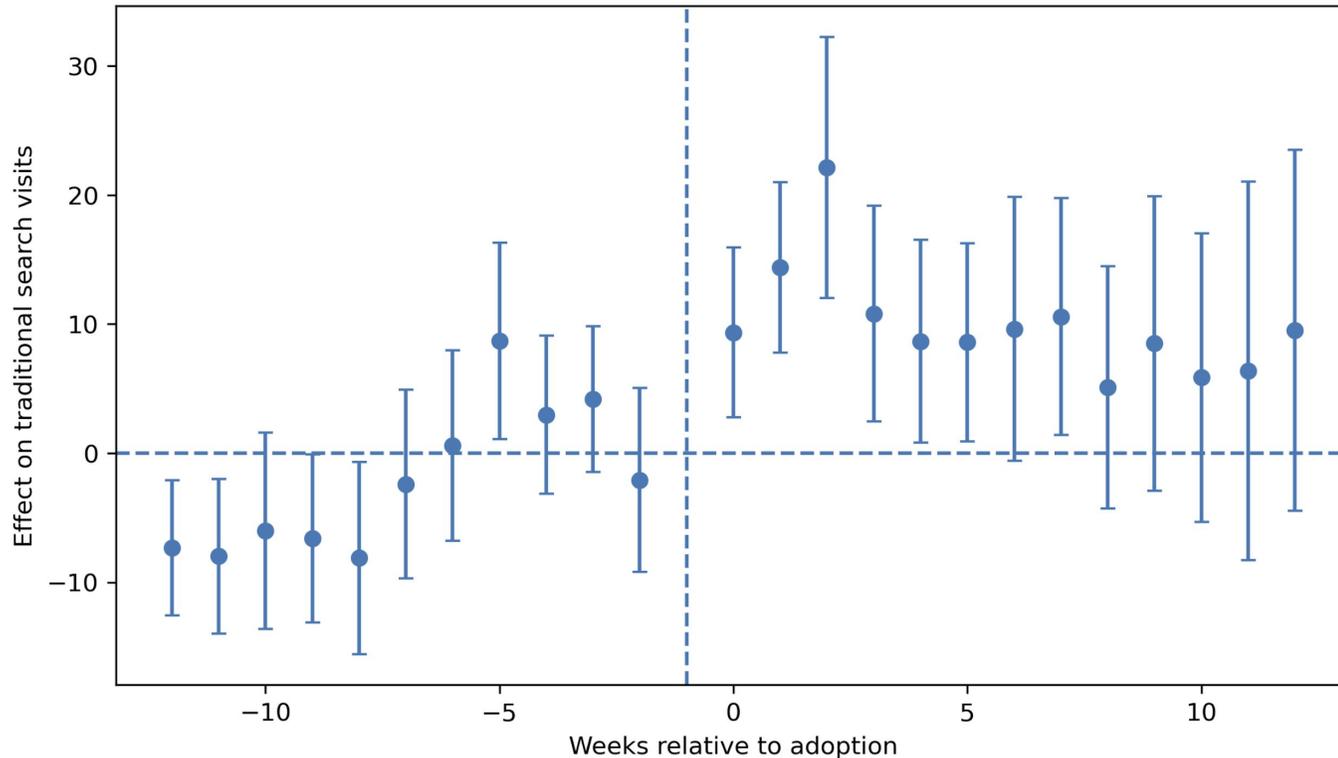
LLMs Adoption (comScore Data, 2022 – 2024)

$$\text{unique visits to search engines}_{i,t} = \sum_k 1\{\tau_{it} = k\} + \gamma n_{i,t} + \alpha_i + \lambda_t + \varepsilon_{i,t}$$

$=1$ if adopted

number of websites

Event study: effect of LLM adoption



- We looked at whether adoption might have changed usages of search engines.
- Analysis suggests there's **no evidence of a reduction in unique visits to search engines.**

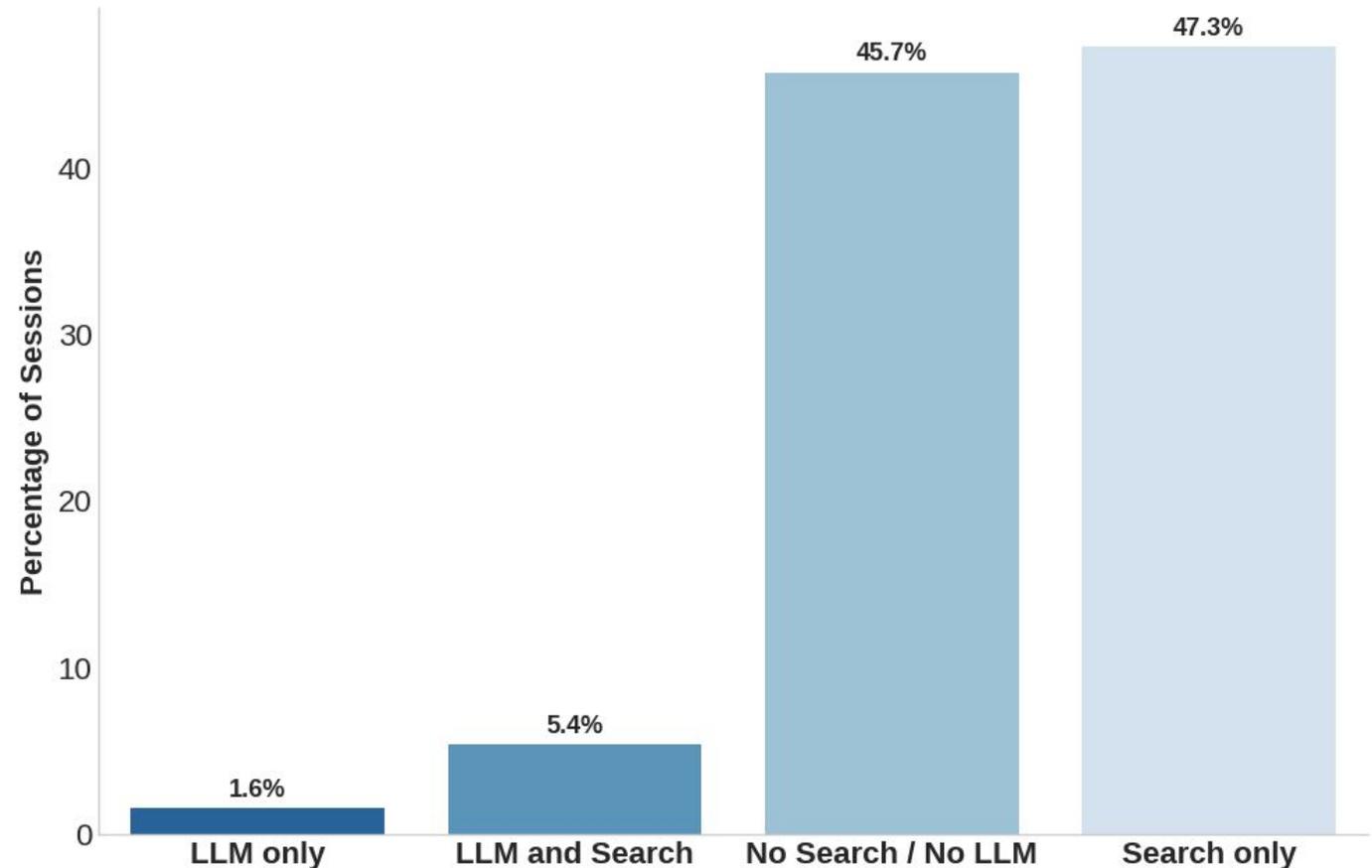
Fact 4: LLM usage is still quite limited (c. 7%)

Session types in the IBE Data

We focus on **browsing sessions**, i.e., browsing activity with <30 min. inactivity.

- Search dominates a great part (**52%**) of sessions.
- In <2% of sessions, users use LLMs without search.
- Overall, LLMs are used in **7%** of sessions.

Distribution of Session Types (Avg. per user)

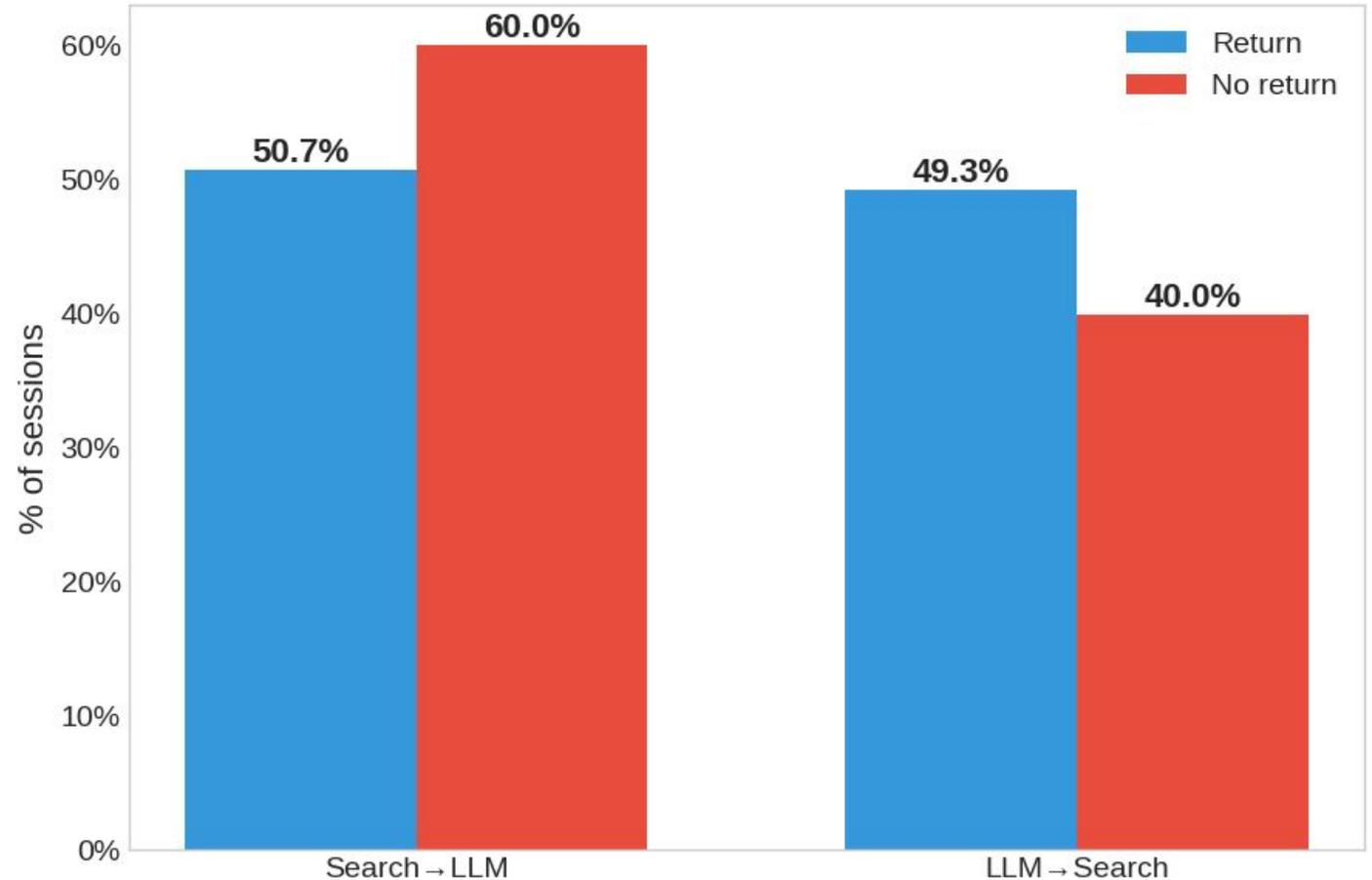


Fact 5: Switches are more frequent from Search to LLMs

Transitions and non-return transitions (IBE Data)

We focus on **mixed browsing sessions** (i.e., Search + LLMs).

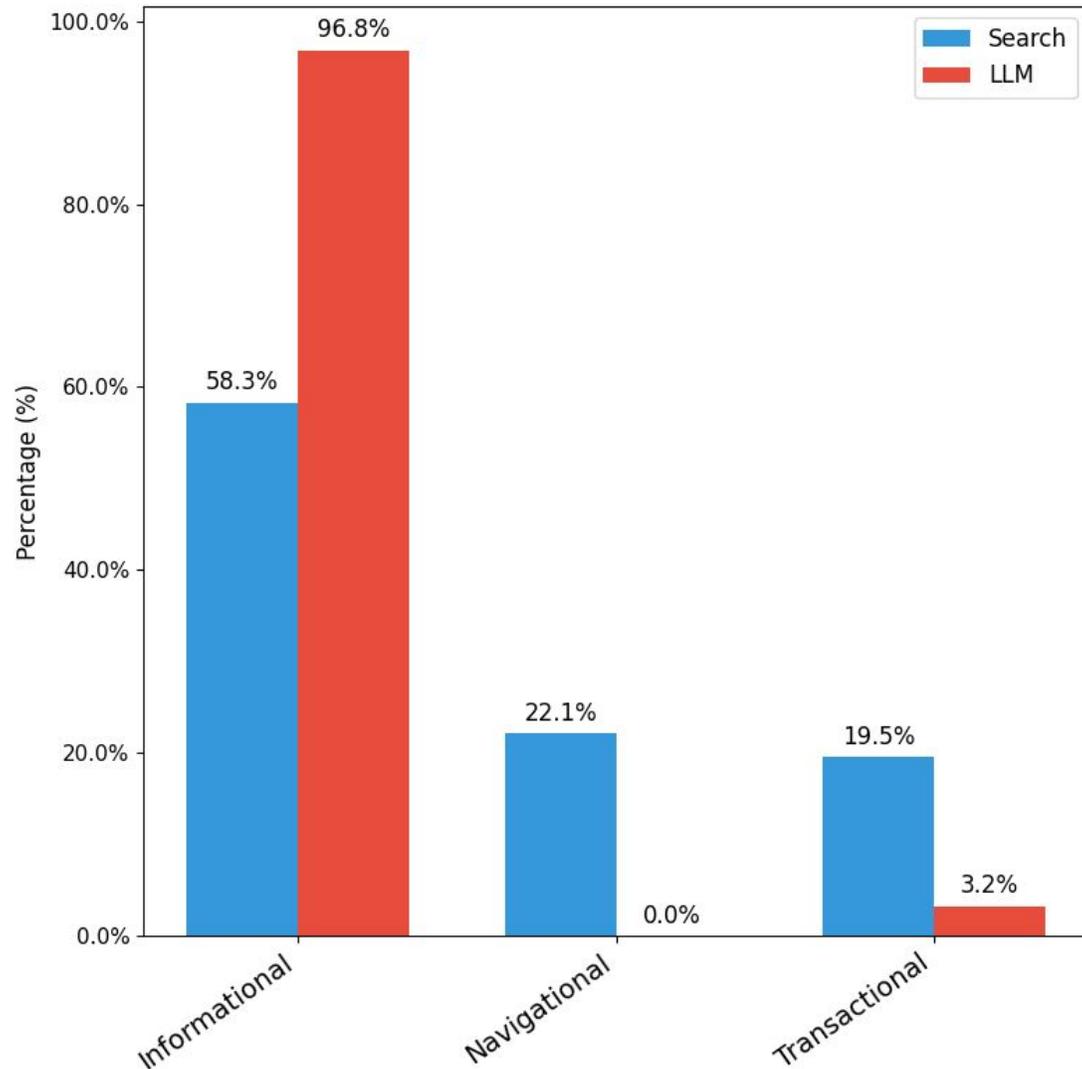
- Users switch frequently from search to LLMs and vice versa.
- **Focus on non-return:** after querying, users tend to switch from search engines to LLMs and stay.



Overview

- Data
- Empirical Facts
- **Proposed Taxonomy**
- Downstream Traffic Analysis

Traditional taxonomy – mostly informative queries (IBE Data)

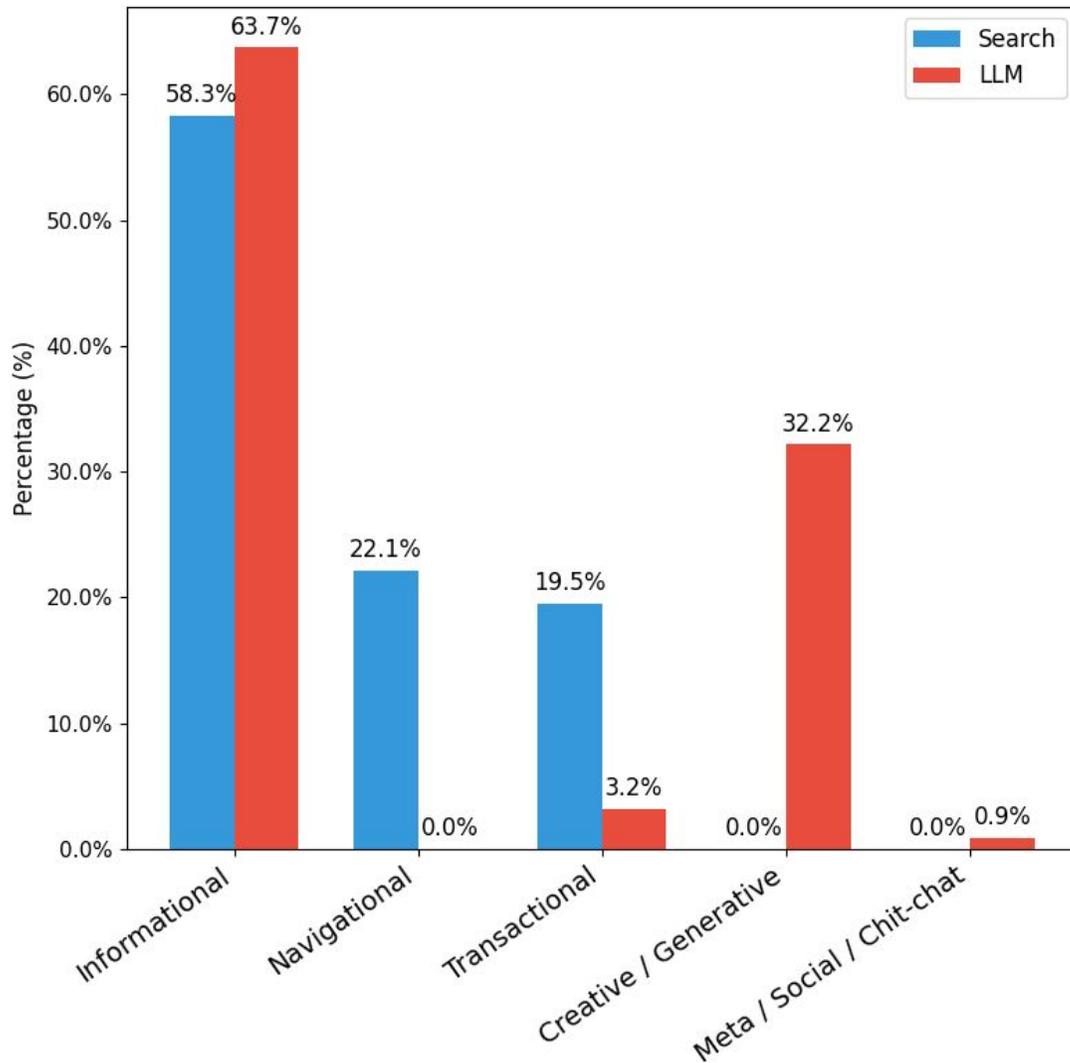


The IBE data includes conversations of participants with LLMs.

Option 1: classify search queries and LLM conversations using Broder (2002)'s taxonomy

- **Informational** – aiming at explanations or definitions
- **Navigational** – aiming at reaching other locations in the web (e.g., Amazon Login)
- **Transactional** – aiming at downloads/purchases

Proposed taxonomy – expand what LLMs can do (IBE Data)

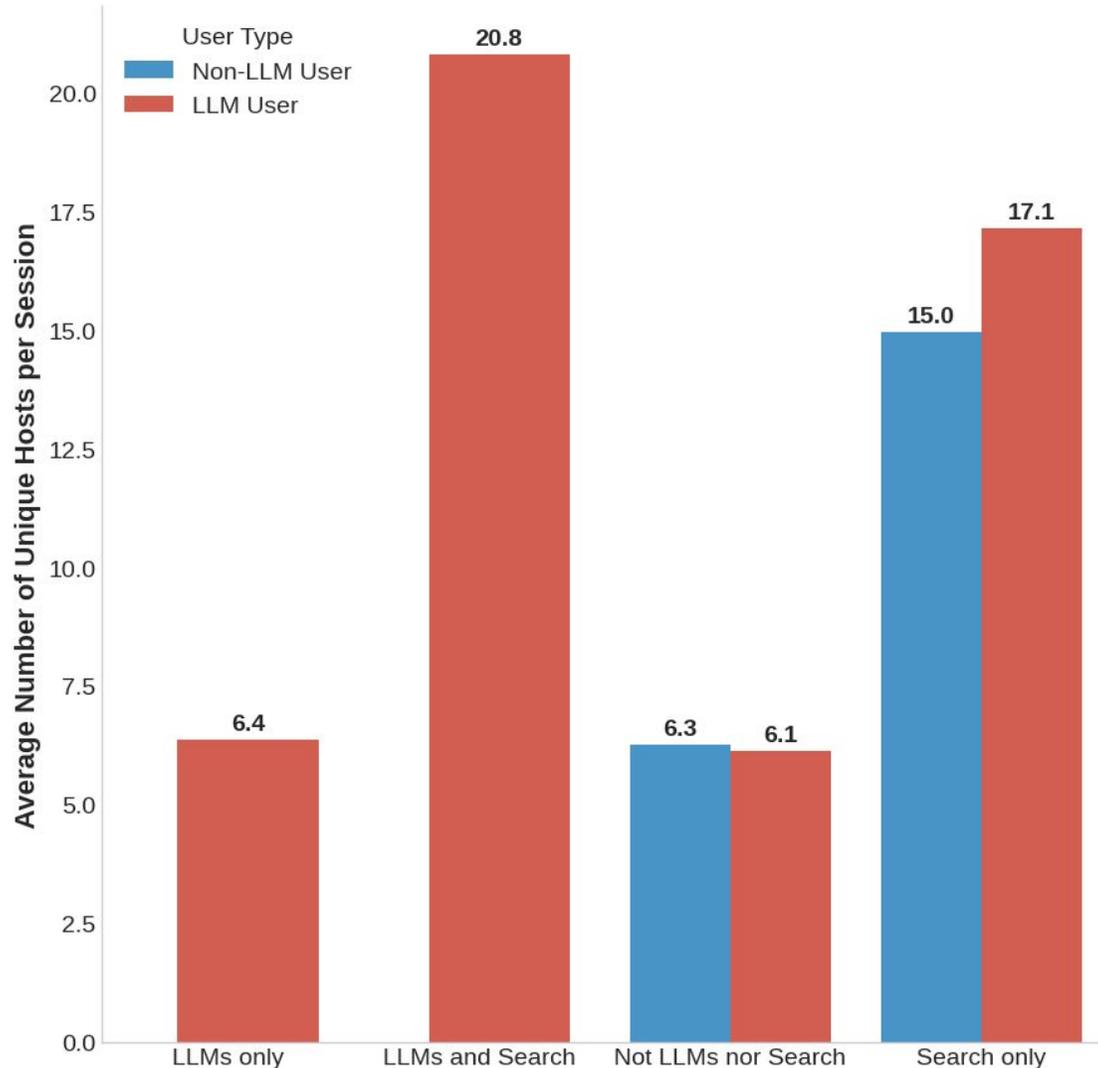


Option 2: ‘Augment’ Broder (2002), by adding two additional categories

- **companionship;**
- **generative / creative tasks.**

- **Informational queries are still dominant.**
- **Creative and Generative is the 2nd most important category**

Research Question: How does web traffic change with LLMs?



We analyze the websites participants visit, in different session types.

Only LLMs: fewer websites (6.4).

Only search: almost 3x more websites

LLMs + Search: highest number of websites.

□ Possible explanation: LLMs augment what people can do and the information they need.

□ Possible issue: *self-selection*

Overview

- Data
- Empirical Facts
- Proposed Taxonomy
- **Downstream Traffic Analysis**

Research Question: How does web traffic change with LLMs?

$$\text{unique domains (no search/LLMs)}_{i,t} = \beta_k \text{Session Type}_{i,t} + \gamma \text{AI Overviews}_{i,t} + \lambda_i + \varepsilon_{i,t}$$

unique domains (no search/LLM)	(I)	(II)
Only LLMs	-0.408 (1.010)	-0.604 (0.987)
LLMs + Search	2.330*** (0.550)	2.141*** (0.544)
AI Overviews Displayed	2.614*** (0.404)	1.663*** (0.405)
<i>Participant FE</i>	Yes	Yes
<i>Category FE</i>		Yes
Observations	20,805	20,805
	0.52	0.53

We regress browsing activity on session type (**baseline = only search**).

- Not significant reduction due to LLMs
- When LLMs are used together with search, users **browse 2 more domains** than they would.
- We control for the presence of AI overviews, which are associated with higher browsing activity.

Research Question:

Are LLMs causing more concentrated/less diverse browsing?

We focus on the **5-minute window** after a user inserts a query. We compute:

- **Externalization:** probability of visiting an external website.
- **Concentration:** $HHI = \sum_d s_d^2$
- **Diversity:** $H = \sum_d s_d \log(s_d)$

	Average HHI	Average Entropy	Average externalization
Search (non-LLM users)	0.674	0.623	0.597
Search (LLM users)	0.644	0.673	0.636
LLMs (LLM users)	0.639	0.711	0.699

- **But** only few (approx. 20%) LLM conversations contains external links.
- **People might just “move on”** with their browsing after LLMs.
 - If so, then **LLMs might compress browsing into fewer websites.**

Conclusions

We study whether the adoption of LLMs might reduce browsing activity, challenging the existence of freely accessible websites.

Findings:

- The adoption of LLMs is gradual, but search engines remain dominant.
- No evidence that usage of LLMs reduce browsing activity.
- LLMs *augment* what users can do with traditional search engines (e.g., *creative queries, coding requests, etc.*)

Next steps:

- Analyze the browsing activity following LLM conversations to study.
- Analyze the overlap in intents between search queries and LLM conversations.