

Penny for Your Prompts: A Measurement Study of Contextual Leakage in AI Chat Platforms

Paschalis Bekos
Stony Brook University
Stony Brook, NY, USA
pbekos@cs.stonybrook.edu

Billy Tsouvalas
Stony Brook University
Stony Brook, NY, USA
vtsouvalas@cs.stonybrook.edu

Nick Nikiforakis
Stony Brook University
Stony Brook, NY, USA
nick@cs.stonybrook.edu

Michalis Polychronakis
Stony Brook University
Stony Brook, NY, USA
mikepo@cs.stonybrook.edu

***Abstract*—AI chatbot platforms have become one of the most popular web interfaces users interact with online, routinely receiving personal and intimate information shared under an expectation of privacy. Yet, on the web application side, these platforms have the ability to embed third-party analytics, advertising, and session replay scripts that operate silently alongside the chatbot, creating an underexamined channel for conversational data leakage towards third parties. We conduct the first systematic large-scale measurement study of contextual leakage in AI chatbots, auditing the web interfaces of 42 general-purpose and domain-specific (health, legal, and education) platforms. Beyond standard network interception, we instrument the browser at the JavaScript engine level to attribute each data flow to the exact user action that triggered it, and the DOM capture mechanism used to retrieve its content—granularity invisible to network-layer analysis alone. We find that 21 of these platforms leak conversation titles to third-party trackers, while nine leak verbatim user prompts. Beyond content, we observe a substantial amount of chatbot platforms sending user emails to third-party trackers, enabling the association of conversational context with a user's identity. Our findings cast traditional web privacy issues under a new light, highlighting the privacy issues that are inherent in how users interact with modern AI chatbots.**

1. Introduction

AI chatbot platforms have become one of the most widely used web-based services, with hundreds of millions of users submitting queries that range from general information requests to intimate personal disclosures [1], [2], [3], [4]. Unlike traditional search engines, where queries are typically short and do not encapsulate content in detail, chatbot interactions are typically far more detailed and contextual: users describe medical symptoms, seek legal advice, disclose personal circumstances, and maintain extended conversations that accumulate sensitive context over time. This shift fundamentally raises the privacy stakes of web activity, as chatbot interactions often contain substantial sensitive personal information.

Yet, from a web-infrastructure standpoint, AI chatbot platforms are built and operated like any other commercial web application, embedding the same third-party scripts found across the broader web. Among these authorized third parties are the same tracking pixels and analytics frameworks that have been extensively studied in the context of traditional websites [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. These frameworks are designed to monitor user interactions, attribute actions to conversions, and infer interests from the content a user engages with. On a traditional website, this amounts to observing which products users browsed or which articles they read. On an AI chatbot, however, the same mechanisms operate over content that is not passively consumed but *actively* disclosed by the user, potentially carrying medical, legal, financial, or deeply personal context—a distinction that has received limited attention.

The gravity of this issue is underscored by class-action complaints recently filed against OpenAI [16], [17] and Perplexity [18], alleging that embedded trackers transmitted conversational context to third parties upon user interaction with these platforms. The research community has only begun to surface the problem. Jazlan et al. [19] audited 20 general-purpose chatbots using single-prompt network capture, and another study [20] examined four prominent platforms through HAR (HTTP Archive) traffic analysis. Both document the presence of third-party trackers and instances of data leakage, but at a limited scale and depth. First, they neglect the interaction-driven nature of tracking pixels and miss the observation of additional leaks that only occur after multiple interactions, such as the leakage of conversation titles that are generated after the first user prompt. Second, they focus primarily on general-purpose chatbots, while there is a growing ecosystem of task-specific chatbots in sensitive domains such as health, legal, and education. These platforms are likely to collect even more sensitive context, yet they have been largely overlooked in prior research. As a result, the scope and mechanics of contextual leakage in these platforms remain largely unexplored.

In this work, we present a systematic large-scale measurement study of *contextual leakage* in AI chatbot platforms. We audit the web interfaces of 42 platforms, including 25 general-purpose chatbots and 17 task-specific agents in sensitive domains (health, legal, and education). Our study combines Puppeteer-based network capture with

VisibleV8 [21] JavaScript (JS) engine instrumentation to attribute each data flow to its causal user interaction and context extraction mechanism. Using Puppeteer, we conduct structured multi-interaction sessions across each platform, capturing all outbound network traffic to identify third-party recipients of conversational context and user identity. We then employ VV8 to isolate five interaction scenarios per platform, attributing each leakage flow to its causal trigger. Beyond interaction attribution, VV8's execution traces allow us to reconstruct the JS-level mechanisms through which third-party scripts retrieve conversational context from the structural components of the web page.

Our study allows us to identify conversational context leakage across AI chatbot platforms. Specifically, we find that 21 platforms leak conversation titles and 9 leak verbatim user prompts to third-party trackers. Moreover, in the general-purpose chatbot platforms we identify that 11 of the 13 context-leaking platforms also expose the user's email to trackers, enabling linkage of conversational context to a named individual. Our analysis of the JavaScript execution traces reveals that the mechanisms deployed by third-party scripts to retrieve descriptive page content on traditional websites, inadvertently capture conversational context on AI chatbot platforms. We find that each chat session is treated as a separate sub-page within the platform and assigned a dynamically generated title derived from the user's prompt, which these scripts then read and transmit as if it were ordinary page metadata. Our key contributions are:

- We audit 42 AI chatbot platforms, covering both general-purpose and task-specific chatbots in sensitive domains (health, legal, and education).
- We identify conversational context leakage alongside identity leakage towards third-party entities, revealing that trackers receiving conversation context can also link it to the identity of the user.
- We attribute each leakage flow to its causal user interaction and reconstruct the JS-level mechanisms trackers use to extract conversational context from the page.
- We show that third-party scripts deployed on AI chatbot platforms inadvertently capture conversational context through the same mechanisms they use on traditional websites, as the dynamic nature of chatbot interfaces causes page-level attributes to encode user intent rather than generic page metadata.

2. Background

The preliminary analysis of the Content Security Policies (CSPs) of some popular general-purpose chatbot platforms, as illustrated in Figure 1, reveal a dense third-party ecosystem. Platforms such as You.com, QuillBot, and Claude reference as many as 66, 57, and 56 distinct third-party domains, respectively, in their CSP headers alone. While a CSP entry does not necessarily imply active data collection, it authorizes third-party scripts to execute within the platform's security context, establishing the potential for data flows between the chatbot interface and external entities. This observation motivates a deeper investigation into how these third parties operate in the context of a user's navigation and use of these platforms, and crucially, what data they are able to collect.

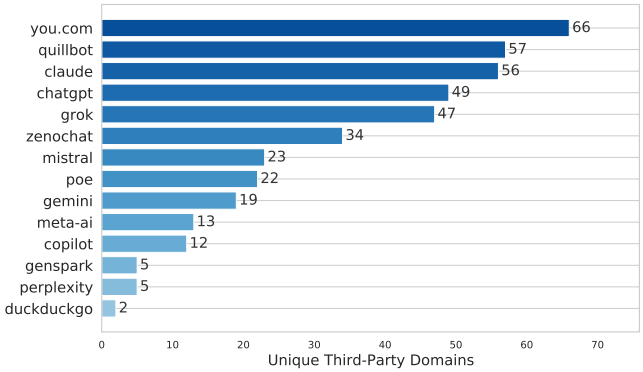


Figure 1. Number of unique third-party domains referenced in the Content Security Policy (CSP) of each general-purpose AI chatbot.

2.1. AI Chatbots

AI chatbots are web-based or application-based conversational interfaces powered by large language models (LLMs) that simulate human conversation and generate natural language responses to user queries. In recent years, users have integrated these platforms into everyday use, submitting prompts ranging from generic questions and work-related inquiries to sensitive personal topics. Platforms such as ChatGPT, Gemini, and Claude have seen rapid consumer adoption [1], [2], [3], [4], serving hundreds of millions of users across a wide range of tasks, including information retrieval, content generation, and personal assistance.

A distinguishing characteristic of AI chatbots compared to traditional web applications, websites, or search engines, is the inherently contextual and potentially sensitive nature of user input. Unlike a typical search engine query, which usually consists of a few keywords, a chatbot prompt often contains detailed disclosures ranging from generic topics (research questions, open-ended inquiries, task assistance) to sensitive ones (medical symptoms, legal situations, financial concerns, or personal inquiries). A further implication arises from the preservation of *history* across chats, which maintains both the context and memory of prior questions. To that end, most platforms automatically generate conversation titles that summarize the user's inquiry, creating a second, platform-derived artifact that encodes the semantic context of the interaction.

Figure 2 illustrates how conversational context can be exposed to third-party entities alongside personally identifiable information (PII), enabling the association of a user's identity with their conversational context. To illustrate the leakage model, we consider a scenario involving three actors: a user X, a web-based AI chatbot platform C, and a third-party tracker T operating under the domain `tracker.net`. We assume that T has no prior information linking user X to any online identity.

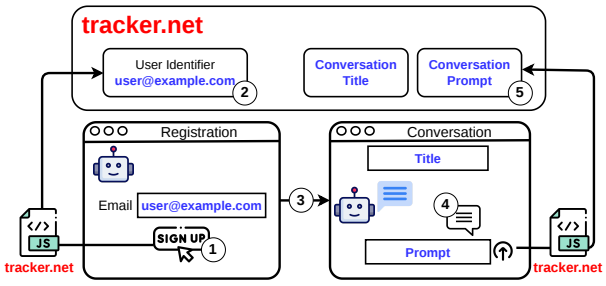


Figure 2. Overview of contextual leakage from an AI chatbot platform to a third-party tracker.

The user first visits the AI chatbot platform C ①. Upon page load, the platform includes a third-party tracking script served by T. The user proceeds to create an account by filling out and submitting a registration form ②. This submission event is captured by an event listener registered by T's script, which extracts the user's email address (`user_email`) from the form and transmits it to T's server ③. At this point, T can associate a persistent identifier with user X. The user then navigates to a new chat window ④ and submits a prompt containing a personal inquiry. The platform processes the query, generates a response, and dynamically assigns a *title* to the conversation that summarizes the user's input ⑤. This title is rendered into the page's DOM, typically as the `<title>` element of the page or within a sidebar navigation entry. A subsequent browser event triggers T's tracking script to read the page state again. The script now accesses the conversation title and, depending on the platform, the prompt text itself, and transmits this contextual information to T's server ⑥. From this point onward, tracker T possesses both the user's email address and the semantic context of their conversation. By joining these two data points, T can associate a previously anonymous user with the substance of their private inquiry, effectively *deanonymizing* the conversational context.

This type of contextual leakage, particularly when coupled with identifiers that link conversation context to an identifiable user, has not received sufficient attention. Recent work has studied data collection within the GPT app ecosystems [22], [23], while a recent report [24] identified malicious extensions exfiltrating conversations across 10 AI platforms. Jazlan et al. [19] presented the first measurement of third-party tracking on chatbot web platforms, finding a subset of chatbots sharing contextual information with at least one third party. However, to date, no study has traced the full leakage chain from user interaction to JavaScript-level extraction mechanism, nor has extended the analysis to task-specific chatbots in sensitive domains.

2.2. Tracking Pixels

The online activity of a user is a rich source of information that can be leveraged to understand behavior, tendencies, and preferences. Website operators routinely embed third-party tracking scripts, commonly referred to as *tracking pixels* [5], [6], [7], [8], [9], to support analytics, advertising attribution, and conversion measurement. These JavaScript-based trackers have been widely used across the web and are deployed by platforms such as Meta, Google, and TikTok to enable targeted advertising at scale. When deployed on a website, these pixels can read contextual signals from the webpage, including the URL, page title, and referrer, and propagate the collected data to their own servers via network requests typically embedded as values to the request's payload. Beyond page-level context, many trackers register event listeners on forms and interactive elements, enabling them to capture user-provided PII, such as email addresses, names, and phone numbers.

Prior work has extensively documented the privacy implications of tracking scripts, revealing their widespread use across the web (including in sensitive categories of websites) [10], [11], [12], [13], [14], [15], studying their data collection from a website's context [25], [26], and their ability to collect PII [27], [28], [29], [30], [31], [32], [33], [34], [35]. The adoption of tracking scripts by AI chatbots introduces additional privacy implications, as these platforms render user-generated content directly into the DOM. Prompts, responses, and dynamically generated conversation titles all become part of the page structure, readable by any executing script. The same DOM read that captures a news headline on a publisher's site can capture a summary of a user's private query on a chatbot. The exposed context is therefore not just *who* the user is, but *what* they are asking about, seeking help with, or disclosing in confidence.

3. Methodology

To identify cases of contextual leakage from AI chatbot platforms to third-party trackers, we design a systematic measurement methodology that combines network capture with JavaScript-level instrumentation. Our approach consists of two main complementary phases: *session capture* and *interaction decomposition*. For session capture, we instrument a Puppeteer-driven browser that allows us to manually interact with each platform across structured sessions, capturing all outgoing traffic. This enables the identification of third-party data collection of conversation context and personally identifiable information (PII). While network interception reveals the full scope of data collection, it cannot attribute a given data flow to a specific user action or reveal *how* trackers extract conversation context from the page. To address this, we implemented an interaction decomposition approach based on VisibleV8 (VV8) [21], an instrumented build of the Chromium browser that logs all JavaScript API calls at the V8 engine level, enabling us to decompose each leakage event across its causal trigger and the underlying extraction mechanism. We use this methodology to audit a wide range of general-purpose and task-specific AI chatbots.

3.1. AI Chat Platform Selection

Our study focuses on AI chatbots that provide a web-based interface, as users typically interact with these plat-

TABLE 1. AI CHATBOT PLATFORMS INCLUDED IN OUR STUDY.

General Purpose		
Microsoft Copilot [37]	DeepSeek [38]	Claude [39]
Mistral Le Chat [40]	ZenoChat [41]	Meta AI [42]
Character.AI [43]	QuillBot [44]	Manus [45]
HuggingChat [46]	Qwen Chat [47]	Grok [48]
Writesonic [49]	You.com [50]	Kimi [51]
Perplexity [52]	ChatGPT [53]	Poe [54]
DuckDuckGo [55]	SeaArt [56]	Pi [57]
Genspark [58]	ChatOn [59]	
PolyBuzz [60]	Gemini [61]	
Education		
Khanmigo [62]	EduBrain [63]	Flexi [64]
SciSpace [65]	Elicit [66]	iAsk [67]
Answer.AI [68]		
Health		
Free AI (not a) Therapist [69]	FeelBetterBot [70]	Wellzy [71]
iAsk (companion mode) [72]	Earkick [73]	Abby [74]
AITherapistFree [75]		
Legal		
LawConnect [76]	NyayGuru [77]	AI Lawyer [78]
TheLawGPT [79]		

forms through a standard browser session. We selected a total of 42 chatbot platforms; 25 general-purpose and 17 task-specific ones. For the general-purpose platforms, our selection was based on publicly-available popularity rankings [36]. The task-specific chatbots were selected to span educational, health-oriented, and legal-assistance platforms; three domains explicitly chosen because of the inherent sensitivity of the conversations users share within them.

Table 1 lists the platforms included in our study. Although several general-purpose platforms target a specific niche, such as, creative writing assistance (Writesonic, QuillBot), image generation (SeaArt, Genspark), or character-based roleplay (Character.AI, PolyBuzz), all selected platforms expose a general conversational interface that accepts arbitrary user input, distinguishing them from task-specific tools that restrict queries to a narrow domain. The task-specific platforms fall under the sensitive category with respect to their primary conversational target, covering (i) *Education*, (ii) *Health*, and (iii) *Legal* assistance. In the case of iAsk we consider it for two categories: (i) Educational, which is the primary category, and (ii) Health, as it offers iAsk companion-mode, a virtual companion focused on emotional and mental well-being support.

Our selection is limited by the requirement for free account subscription and the availability of a web-based interface that accepts arbitrary natural language prompts. Starting from SimilarWeb's "AI Chatbots and Tools" ranking, we selected platforms that expose a chat-based interaction model where users can submit free-form queries and receive AI-generated responses, filtering out coding assistants, app-building environments, API routers, and enterprise platforms without free public access. We focus exclusively on web-based interfaces and do not consider mobile applications, which may employ different tracking configurations and require separate analysis. In total, we audited *42 AI platforms* for both identity and contextual leakage towards third-party entities. All measurements were conducted from a US East vantage point during May 6 and May 20, 2026.

3.2. Session Capture

We instrumented a Puppeteer-driven browser [80] to interact with each platform across multiple browsing sessions, performing our analysis manually to remain faithful to the actual user experience. Each session followed a structured sequence of user interactions:

1) *Account Creation*
2) *Prompt Submission:* We submit a prompt and observe the platform's response ("weather inquiry for a city X").
3) *Prompt Submission:* We submit a different prompt in a new chat window ("dentist recommendation for a neighborhood Y in city X").
4) *Prompt Navigation:* Navigation between the two conversation windows.
5) *Session Termination*

Each session is executed in a clean browser instance with no prior state, cookies, or cached data, and we interact with each platform under its default configuration, accepting any cookie or consent prompts as presented without modifying chat settings or privacy controls. We repeat each session under two separate accounts: one registered with a Gmail address and one with a custom email hosted on our own mail server. This variation is applied only to platforms that support email-based registration. For platforms that allow interaction without account creation, such as DuckDuckGo AI Chat, we capture a no-account baseline session.

Our prompts cover both a generic inquiry (weather) and a sensitive one (medical/dentist) and across all sessions we submit the same exact prompts to enable consistent comparison. Each prompt yields a *dynamically* generated title containing *contextual* information about the user's inquiry, which we record for each session. For the character-based platforms (Character.AI and PolyBuzz), we interact with the *same* set of characters across sessions and record the resulting conversation title, which reveals the name of the character the user is conversing with. Throughout each session we capture all outgoing network requests, recording the full URL, cookies, headers, and POST bodies. Each request is then analyzed for the presence of the seeded prompt or the recorded title. We classify a leakage event when the domain receiving conversation context (the full prompt or prompt title) or PII does not belong to the same parent domain (eTLD+1) or an organizational domain from the same vendor (e.g., anthropic.com with claude.ai). We further exclude CDNs, consent management platforms, payment processors and other infrastructure domains that do not constitute tracking or advertising third parties.

For the task-specific chatbots, we follow the same session structure but limit registration to a Gmail account only, as the goal is to measure whether sensitive conversation context is exposed to third parties rather than to re-evaluate

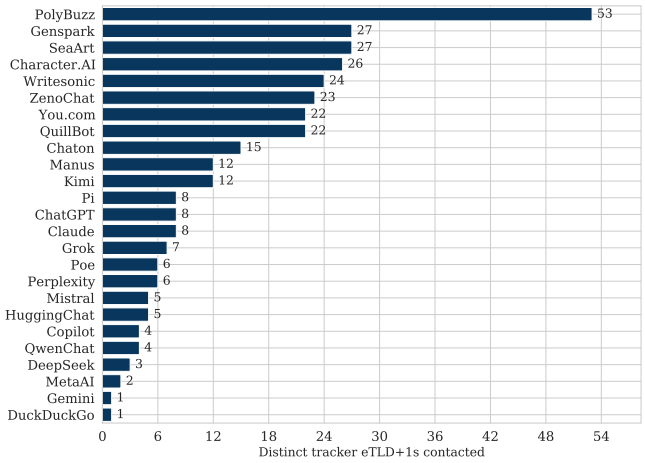


Figure 3. Unique tracker domains (eTLD+1) contacted by each chatbot.

the effect of registration method. For each of the three categories, we submit two prompts: one generic and one carrying sensitive context reflective of the platform's domain. Specifically, for *education* platforms we submit a general academic reference request alongside a query seeking sources on a sensitive personal topic; for *legal* platforms we submit a general criminal law question alongside an immigration status inquiry; and for *health* platforms we submit a general wellness question alongside a query disclosing symptoms of anxiety and interest in medication. Given that these platforms inherently handle sensitive user inquiries by nature of their domain, any contextual leakage observed carries significantly higher privacy implications than that of general-purpose chatbots.

While this phase reveals third-party actors that collect conversation context, the multi-interaction session design cannot attribute which specific user action triggered each observed data flow. A tracker receiving the conversation title during a session that included both prompt submission and chat navigation could have been triggered by either action.

3.3. Interaction Decomposition

To address this attribution gap, we employ VisibleV8 (VV8), an instrumented build of the Chromium browser that logs all JavaScript API calls at the V8 engine level, including property reads and writes, function invocations, and network-initiating operations. VV8 produces execution traces that capture every interaction between JavaScript code and the browser's Web API surface, enabling us to observe not only *that* data was sent to a third party, but *which* script read it, from *which* DOM property, and through *which* network API.

We design five isolated scenarios, each executed after login using the same email address in a separate VV8 session, producing independent execution traces:

1) *Prompt Submission:* The user types a query, submits it, waits for the chatbot's response, and closes the browser. This establishes the baseline: what third-party collection occurs from the act of submitting a prompt alone.
2) *Chat Navigation:* The user submits a prompt, waits for the response, then navigates to a previously created conversation.
3) *Chat Deletion:* The user selects an existing conversation and deletes it, without submitting any new prompt.
4) *Scroll:* The user submits a prompt, waits for a sufficiently long response, and scrolls through it.
5) *Button Click:* The user submits a prompt, then clicks an unrelated UI element (such as a settings icon or sidebar button) that does not navigate to a different conversation.

Each session is executed after the network is idle, and all scenarios use the same seed prompt ("Pharmacies near me CITY_X, STATE_Y") under the same browser configuration. For each scenario, we search all outbound network operations for two taint markers: the raw prompt and the platform-generated conversation title, in their original form as well as URL-encoded variants. By comparing the third-party destinations and leaked values observed in each scenario against the prompt-submission baseline, we determine which user action is the causal trigger for each data flow.

Beyond this cross-platform differential analysis, VV8's execution traces allow us to reconstruct the full JavaScript execution chain leading to each observed leakage flow. For every outbound request containing the prompt or title, we identify which script executed the request, resolve its origin, and enumerate DOM read operations that the script performed before sending the data. This reveals the specific capture mechanism each tracker deploys to retrieve conversation context, for example, a property access on `document.title` or a DOM query for the `<title>` element via `querySelector` or `getElementsByTagName`. These JavaScript-level distinctions are invisible to network-layer capture, where all mechanisms produce identical HTTP request parameters, thus enabling us to study not only *what* context is leaked but *how* it is extracted before propagation to a tracker.

4. Findings

In this section, we present the results of our measurement experiment across 42 AI chatbot platforms. Specifically, we characterize the prevalence and volume of contextual leakage to trackers, identify the destinations receiving conversation data, and examine how leakage patterns differ across account types for general-purpose (25) and task-specific (17) AI chatbot platforms. Then, we attribute each leakage flow to the user interaction it originated from, and describe the underlying mechanism responsible for gathering and transmitting the data to third parties. Our per-prompt quantities reported throughout this section are averaged over the two seeded prompts submitted within each session.

We note that throughout our tracker analysis, we exclude domains related to cloud infrastructure, CDNs,

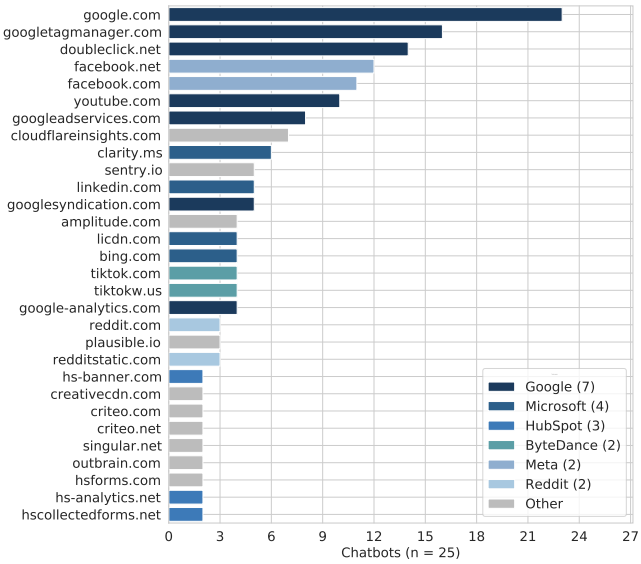


Figure 4. Third-party tracker prevalence across chatbots.

consent management, CAPTCHA, and any other non-tracker service, as well as each platform's own related domains (e.g., `openai.com` for ChatGPT). Both sets were curated manually, with non-obvious platform affiliations confirmed against vendor documentation (e.g., `oaistatic.com` is listed on OpenAI's support pages). Remaining third-party eTLD+1s were attributed to their parent company from public corporate-ownership records (e.g., `doubleclick.net` to Google, `clarity.ms` to Microsoft). We validated this attribution against Disconnect [81] and DuckDuckGo Tracker Radar [82]: 88.9% of assignments are confirmed by both sources and 97.2% by at least one, with no contradictions; the single unconfirmed domain (`gator.volces.com`) was attributed from vendor documentation. This ensures that all reported third-party flows reflect solely analytics, advertising, and behavioral-tracking services.

4.1. Trackers in Generic AI Chatbots

Following the steps outlined in Section 3.2, we first analyze the network traffic generated by each general-purpose chatbot platform. On average, across the 25 general purpose chatbots, each prompt interaction generates 540 outbound requests, of which 118 (22%) are directed to known tracker domains. In terms of tracker domains, we identify 153 unique trackers (eTLD+1) across all chatbots, with each session involving 13 of them on average. In Figure 3, we present the distribution of unique tracker domains across the general-purpose AI chatbots, showing how many distinct trackers each platform embeds.

4.1.1. Tracker Traffic. Beyond their prevalence, we examine where tracker traffic ultimately concentrates. Since a single company operates many distinct domains,

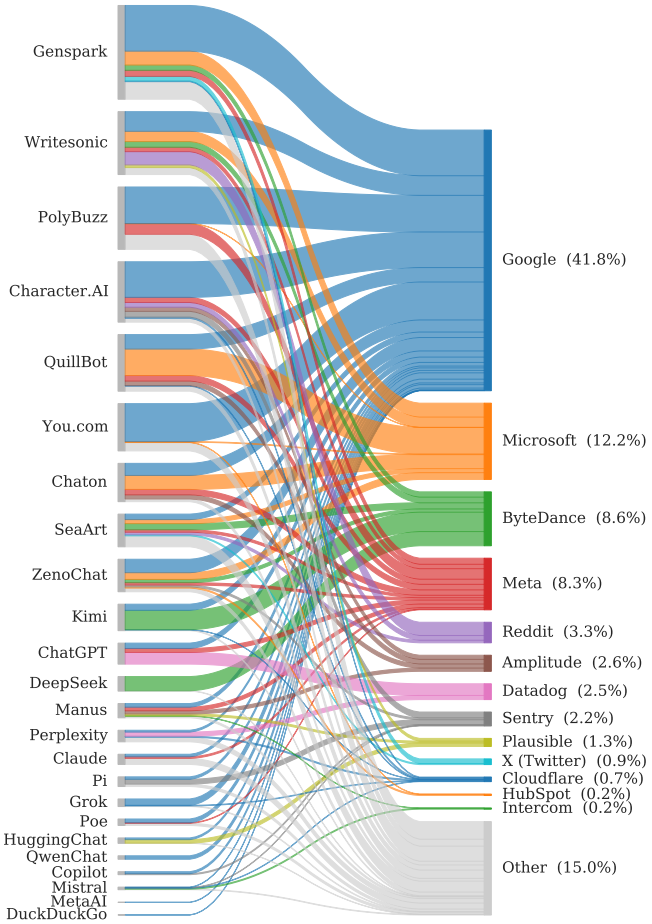


Figure 5. Third-party requests per generic chatbot platform to known tracker domains, aggregated by owner company.

we consolidate them into *owner blocks*, e.g., Google comprises `google.com`, `googletagmanager.com`, `doubleclick.net`, `googleadservices.com`, etc. Counting domains in isolation understates how much data converges on a single recipient, whereas owner blocks reveal the true entity ultimately receiving the requests. Applying this consolidation, in Figure 4, we present each tracker's reach across the corpus grouped by owner block, making clear how a few companies stand behind the many domains and appear on nearly every platform.

Figure 5 maps each of the 25 evaluated chatbot platforms to the tracker owner companies that receive their outbound third-party requests. Each ribbon connects a chatbot on the left to a tracker owner on the right, with ribbon width proportional to request volume. We observe that Google is the dominant recipient, accumulating more than 40% of all tracker traffic, followed by Microsoft (12.2%), ByteDance (8.6%), and Meta (8.3%). Overall, *more than 70%* of all tracker traffic from the 25 chatbot platforms flows to just *four companies*, providing them cross-platform visibility into user interactions across the AI chatbot ecosystem, regardless of which platform a user chooses.

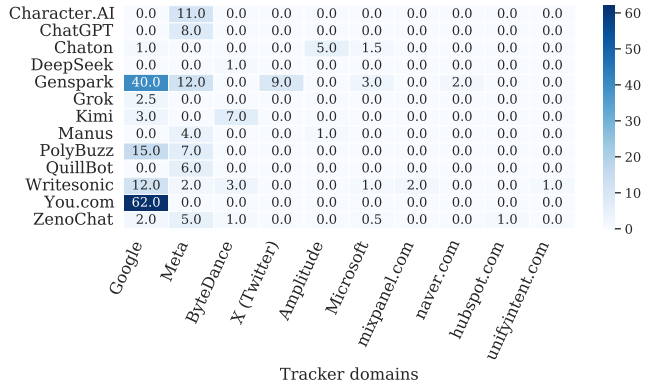


Figure 6. Whole-string title leaks per generic chatbot, averaged per prompt.

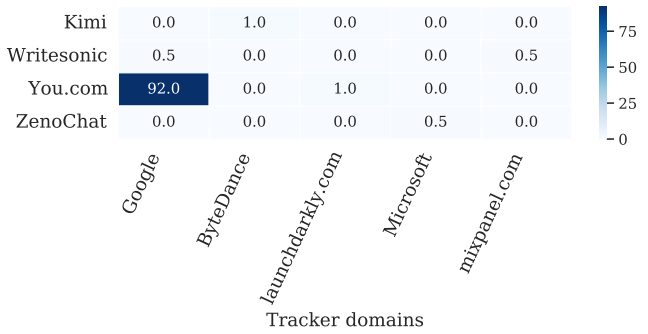


Figure 7. Whole-string prompt leaks per generic chatbot, averaged per prompt.

4.1.2. Conversational Context Leakage. Having established that a significant share of outbound traffic is directed to known trackers, we now examine these requests for conversational context, i.e., chat titles and raw user prompts. The heatmap of Figure 6 presents the instances of *chat title leakage* per chatbot. These are cases where a chatbot makes a third-party request to a known tracker domain, containing the title of the conversation a user has with the chatbot. Each cell corresponds to the average number of leaking requests per prompt towards a domain owned by the respective vendor. Overall, 13 of the 25 evaluated platforms send the title of a user's conversation to a third party tracker. In terms of request volume, You.com and Genspark lead this group, sending 62 and 40 title-bearing requests per prompt to Google, respectively. Moreover, multiple chatbot platforms exhibit substantial leakage to more than one tracker destinations. These findings demonstrate that conversation title leakage is not an isolated phenomenon but a systemic one, with Google and Meta as the primary beneficiaries.

We also identify direct *raw prompt leakage* within the payload of requests propagated to third-party entities, as shown in Figure 7. These are cases where the verbatim user prompt appears in an outbound request towards a known tracker domain. Four platforms exhibit this more severe form of leakage, with You.com being the most prevalent, sending on average 92 prompt-bearing requests to Google every time a user submits a prompt. Notably, Writesonic distributes the raw user prompt to three independent tracker endpoints. This is particularly concerning because unlike conversation titles, prompts are composed entirely by the user and routinely contain sensitive information.

Finally, an interesting observation is the leakage asymmetry based on the email used to register an account. Accounts registered with a Gmail address exhibit higher tracker-bound request rates (24%) compared to accounts using other email providers (19%), and for most platforms the title or verbatim prompt reaches trackers more often under a Gmail account than a non-Gmail one, as we present in Figure 12 in the Appendix. However, with one session per configuration, confirming this asymmetry would require a future dedicated measurement.

4.1.3. Context Leakage per Tracker. Our per-platform analysis reveals which chatbot services leak conversational context and to whom. To understand the aggregate exposure at the ecosystem level, we evaluate which tracker-owning vendors accumulate the most leaking requests.

Figure 8 presents the number of *whole-string title or prompt leakage* requests received by each tracker-owning company. Google accumulates 230 distinct context-bearing requests across eight chatbot platforms and Meta follows with about 55 requests spanning the same number of platforms. The steep volume drop-off after Google and Meta, combined with their equal platform breadth, indicates that conversational context leakage is structurally concentrated in the two largest digital advertising entities, granting them cross-platform visibility into what users discuss with AI chatbots. ByteDance ranks third, driven primarily by Kimi's title leakage, followed by X (Twitter) and Microsoft, each receiving fewer than 15 distinct requests. Figure 13 in the Appendix provides a detailed breakdown of whole-string title and prompt leakage by each tracker destination domain.

4.1.4. Email Leakage and Identity Correlation. So far, we have established *what* leaks and to whom, but not whether any of it can be traced back to a real user. A leaked title or prompt rarely arrives alone: it is accompanied by the persistent identifiers the tracker has already placed on the user, most commonly a long-lived cookie (e.g. Google's *NID* cookie). These identifiers allow the tracker-owning companies to stitch successive conversations into a single profile. Up to this point, this profile remains *pseudonymous*, only referring to a unique cookie value, rather than a real-world identity. However, when the same tracker-owning company obtains a real-world identifier, e.g., the registration email, it is possible to associate the accumulated profile data to a real, named individual.

Therefore, in this section, for each platform, we evaluate which of the 13 AI chatbot platforms that leak the title or the prompt of the conversation also share the registration email with a tracker. Then, we examine whether a *single* tracker-owning company receives both these signals: the registration email (in raw, URL-encoded, or hashed form) and the leaked conversation context. Furthermore, we check whether the

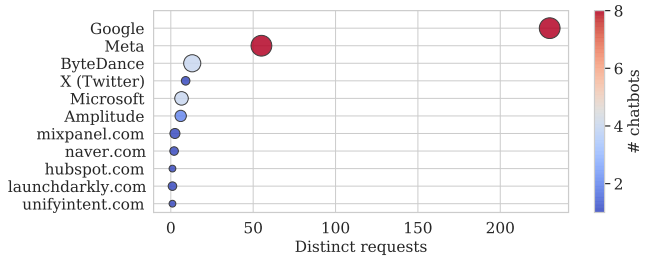


Figure 8. Distinct context-bearing (title or prompt) requests from generic chatbots to top tracker domains.

recipient company can actually join the two, either because they arrive in the same outbound request, or by using a shared, persistent tracking cookie that ties the email and the context requests. Our findings, for the 13 platforms that leak conversational context (the inferred title or the raw prompt), are reported in Figure 9 and in Table 2,

In the funnel diagram of Figure 9, we begin with the 13 platforms that leak the title or prompt on the left, and we progressively evaluate which of those also send the email identity to trackers. Specifically, we observe that out of the 13 that make context-bearing requests to some tracker, 11 also send the user's email. For eight of these, a single company receives *both* the conversation and the email. For two of those eight chatbot platforms (Writesonic, You.com), the email and the conversational context can be found on *the same outbound request*, so we observe the binding directly, while for four others, the respective company collects both signals, across separate requests which share the same unique identifier and can be associated. Last, for Kimi and Polybuzz, we observe that both the email and the conversation reach the same company; however, this occurs over separate requests that bear no common token or identifier. Although we capture no shared token (cookie, URL parameter, etc.) between these requests, the same company still receives both requests which could be probabilistically linked together via the user's IP address or browser fingerprint. The remaining five platforms either split the email and the conversation across different companies (Character.ai, ChatGPT, Quillbot), or never expose the email to any third party altogether (Genspark, Grok).

On Table 2, we attribute each aforementioned email-to-conversation linkage for the general-purpose AI chatbots to the tracker-owning company that receives the relevant requests. In this context, a single entity that receives both the email of a user and the conversational context of their interaction can associate the tracking data to the real user, thus fully identifying them. As such, we find that multiple AI chatbot platforms send both signals to more than one tracker-owning company (Writesonic, Chaton, Manus, ZenoChat), with Writesonic delivering both signals to six distinct companies. In terms of the tracker-owning companies, we find that Google can re-identify users on at least four of the eight platforms, followed by Meta and ByteDance, mirroring the content-leakage concentration reported by Figure 5.

Overall, for eight of the thirteen context-leaking platforms, a single company can correlate the leaked conversation with the registration email, re-identifying the user and tying their chatbot interactions to their real-world identity. On four of these platforms, the exposure is distributed, as multiple tracker-owning companies receive both signals and can re-identify the user independently.

4.2. Trackers in Task-specific AI Chatbots

Having characterized general-purpose chatbot platforms, we now turn to the 17 task-specific chatbots across the education, health, and legal domains. In contrast to general purpose AI chatbots, such platforms are purpose-built for contexts where submitted queries are inherently more likely to be sensitive in nature, making contextual leakage to third-party trackers especially concerning.

4.2.1. Tracker Landscape. Across our interactions with the 17 task-specific platforms, we recorded 5,046 tracker-bound requests spanning 106 distinct tracker domains. On average, each prompt interaction generates 353 outbound requests, of which 140 (39.7%) reach known trackers. This tracker ratio is approximately uniform across the three chatbot categories: education 198/514 (38.5%), health 70/169 (41.6%), and legal 162/393 (41.1%). It holds even though absolute volume varies sharply, with education platforms generating the most requests per prompt and health the fewest. Notably, while task-specific AI platforms issue fewer total requests per prompt than general-purpose ones (353 vs. 540), they devote a markedly larger fraction of that traffic to trackers (39.7% vs. 22%). These smaller, single-purpose tools are proportionally more tracker-dense.

In terms of the tracker domains, traffic is concentrated towards the same companies that dominate the general-purpose ecosystem, as shown in the chatbot-to-owner Figure 14 in the Appendix. Google is present on *every* platform and, together with just Meta and Microsoft account for *more than 70%* of all tracker requests. This reproduces the concentration we observed among general-purpose AI chatbots (Section 4.1.1), with the same handful of large advertising and analytics owners capturing the bulk of the tracker traffic.

4.2.2. Context and Identity Leakage. Task-specific platforms leak conversational context as readily as general-purpose ones: 8 of the 17 send the conversation title to a tracker and five send the verbatim user prompt (Figures 15 and 16 in the Appendix). Because these tools are purpose-built for sensitive domains, the leaked text is often deeply personal. For example, we observed that health chatbot platforms leak verbatim prompts such as "I've been feeling really anxious and having trouble sleeping lately [..]," while law-oriented chatbot platforms transmit prompts explicitly discussing the immigration status of the user. General-purpose chatbot leakage is already harmful, but in the

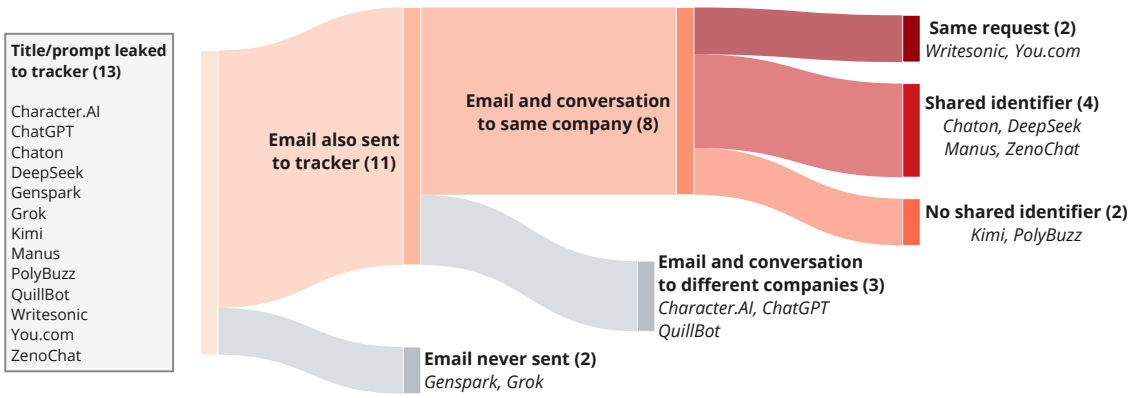


Figure 9. Identity correlation: 13 generic chatbot platforms leak the conversation to a third party, 11 also leak the email, and 8 of them send both to the same company (6 with a directly observed linking identifier).

TABLE 2. EMAIL AND CONTEXTUAL LEAKAGE ACROSS GENERAL-PURPOSE AND TASK-SPECIFIC CHATBOTS. A SINGLE SPANNING ENTRY MEANS EMAIL AND CONVERSATION REACH THE SAME COMPANY. LEAK: T = TITLE, P = PROMPT, E = EMAIL. FOR ENTRIES IN THE "DIFFERENT COMPANIES" GROUP, A / B DENOTES THAT A RECEIVES THE EMAIL AND B RECEIVES THE CONVERSATION.

	Platform	Leak	Receiving Company
Same request	Writesonic	T, P, E	Google, Meta, Microsoft, ByteDance, Mixpanel, Unifyintent
	You.com	T, P, E	Google
	AI Lawyer	T, E	Meta, ByteDance
Shared identifier	Chaton	T, E	Google, Microsoft, Amplitude
	DeepSeek	T, E	ByteDance
	Manus	T, E	Meta, Amplitude
	ZenoChat	T, P, E	Google, Meta
	SciSpace	T, P, E	Google, Amplitude, Intercom
No shared identifier	Kimi	T, P, E	Google
	PolyBuzz	T, E	Google
	EduBrain	T, P, E	Google / Intercom
	Elicit	T, E	Google
	iAsk	T, P, E	Google
	iAsk Companion	T, P, E	Google
	TheLawGPT	T, E	Google
Different companies	Character.AI	T, E	Google / Meta
	ChatGPT	T, E	Google / Meta
	QuillBot	T, E	Google, Microsoft / Meta
	AnswerAI	T, P, E	Google / Microsoft
Email not sent	Genspark	T	Google, Meta, Microsoft, X (Twitter), Naver
	Grok	T	Google

case of task-specific chatbots, the sensitive conversational context that gets leaked carries medical and legal details, which users should be aware that they are sharing with an advertising network.

Moreover, this content does not stay anonymous. On Table 2, we present the context-leaking generic and task-specific platforms together with the email leakage that accompanies the conversation. For the task-specific chatbots, all eight platforms also transmit the user's email to a third party, and for seven of them a single company receives both these signals, allowing for a one-to-one linkage of the conversation to the user's real-world identity. In the case of Edubrain, Intercom receives the conversation but not the email of the user, while Google receives the email and conversation context through different requests with no shared identifier. As for the general-purpose, Google is the dominant recipient, positioned to re-identify users on at least two of these seven platforms. In the starkest case, the legal assistant AI Lawyer sends the conversation title and the user's email in the *same* request to Meta and ByteDance. As such, what users treat as a personal health or legal consultation becomes an identified record for advertisers.

4.2.3. Leakage Across Task-Specific Chatbot Categories. In Figure 10, we present the per-category average number of requests that each platform makes to trackers, along with the subset of those requests that carry the conversation title, prompt, or the user's email. We note that these category-level observations reflect the subset of platforms we were able to analyze within each category, rather than each category's ecosystem as a whole. Among the platforms we study, education platforms are both the heaviest trackers, at roughly 367 tracker requests per platform, and the most leak-prone, with about 24% of those requests carrying user content. Legal platforms send a comparable volume (317 per platform) but leak a much smaller share (9%), while health platforms track the least (124 per platform) and leak the least (3%). Within each category the leakage is highly concentrated rather than uniform, e.g., SciSpace is responsible for roughly 80% of education's leaking requests and AI Lawyer for nearly all of legal's, further indicating that these differences are driven by individual platforms rather than inherent to each category.

4.3. Interaction Decomposition

Across all three categories, a substantial share of each platform's traffic reaches trackers, even though the fraction

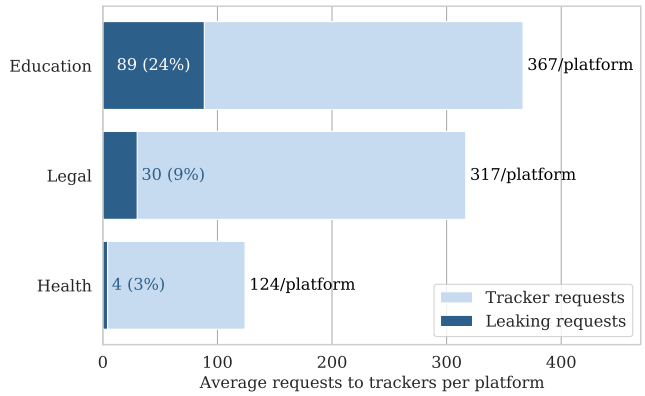


Figure 10. Per-category tracking and leakage: average tracker requests per platform (light), and the share that leak the title, prompt, or email (dark).

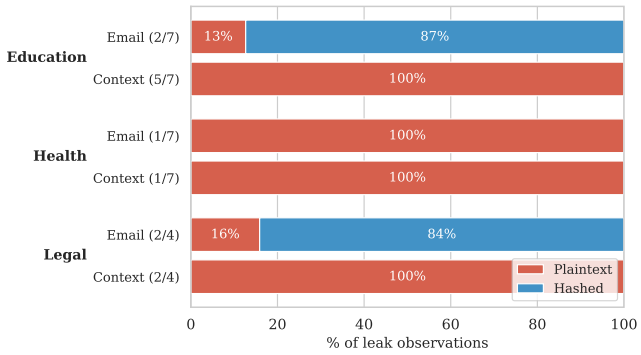


Figure 11. Email vs. content leak encoding by category. The *Email* and *Content* rows show how each category encodes the leaked email and the leaked title/prompt (plaintext vs. hashed); the counts give the number of platforms per category (of its total) that leak.

that carries leaked context differs, with the education platforms in our set leaking the most and the health platforms the least. It is already concerning that these single-purpose tools leak sensitive conversations to trackers at all, and more concerning that they do so at this scale. Moreover, we observe that a smaller leakage footprint does not necessarily translate into better protections for the users. In Figure 11, we present how each category encodes the email and the conversation content when they leak. Across all three categories, the conversation content, whether the title or the prompt, is transmitted in plaintext. In contrast, the email presents a less uniform approach, with the education and legal platforms predominantly hashing it and sending a SHA-256 digest of the email address. However, the health platform that leaks the email transmits it in completely unencrypted form. Although the health chatbots in our set represent the smallest leakage footprint, sending the email in raw form represents a more direct privacy risk. Unlike hashed identifiers, a raw email address requires no additional processing to link a user to their interactions, making it straightforward to re-identify the user. As such, even a platform with a small leakage footprint can pose a significant identity-privacy risk, handing third parties the data needed to deanonymize the user in plaintext.

To determine which user actions trigger contextual leakage, we isolate five interaction scenarios per platform, each executed in a separate VV8 session with independent execution traces. For this experiment, we focused only on the platforms where we found contextual leakage to third parties. From our analysis we exclude two platforms (ChatGPT and Grok) which have updated their web interface and removed the existing trackers before our experiment took place. Additionally, we also exclude two character-based AI chatbots where the title of the conversation is directly leaked upon navigation and corresponds to the character the user interacts with. For these two platforms we did not find any third party tracker collecting the full conversation context (prompt). We start by analyzing all outbound network operations for the presence of two markers: the raw user prompt and the platform-generated conversation title. By comparing the third-party destinations observed in each scenario against the prompt submission baseline, we identify which user action is the causal trigger for each leakage flow.

Table 3 summarizes which platforms exhibit contextual leakage to at least one third party under each scenario. Six platforms already leak conversational context upon *submission* of the prompt, before any navigation takes place. Specifically, on Zenochat and Writesonic, Microsoft Clarity captures the raw prompt text from the input text field in the DOM even before the platform has generated the title of the conversation. On You.com the prompt is embedded directly on the page URL so every tracker that populates their payload with the `location.href` (googleads.g.doubleclick.net, www.google.com, www.googleadservices.com) receives the whole prompt upon submission found in the `tiba, url, ref` query parameters. On Kimi, Genspark and DeepSeek, the platform sets the conversation title fast enough that it is already present in `document.title` when the first analytics page view fires during the initial session.

Because conversation titles are dynamically generated upon response, some tracking scripts have fired the corresponding event before the title is rendered in the DOM. This is evident as we see that *all* nine platforms expose the conversation title to third parties during navigation to an existent conversation. This is because navigating within a single-page application, which handles different chats as sub-pages, fires a synthetic page view event in embedded analytics and advertising tags, which read and transmit the sub-page title, which also holds the title of the conversation.

The remaining scenarios produce minimal additional leakage. *Scroll* and *Click* re-fire the same analytics tags already observed on Submit and Navigate, without introducing new third-party destinations. *Delete* triggers leakage on only one platform (Writesonic), where deleting a conversation causes a page view event that carries the title to Google Analytics. These results indicate that the primary leakage vectors are the initial prompt submission and the first subsequent navigation, with other interactions contributing little beyond what these two scenarios already expose.

TABLE 3. PROMPT OR TITLE LEAKAGE TO THIRD PARTIES BY INTERACTION.

Platform	Submit	Navigate	Delete	Scroll	Click
Zenochat	✓	✓	✗	✓	✓
You.com	✓	✓	✗	✓	✓
Quillbot	✗	✓	✗	✗	✗
Manus	✗	✓	✗	✗	✗
Kimi	✓	✓	✗	✓	✓
DeepSeek	✓	✓	✗	✗	✗
Genspark	✓	✓	✗	✓	✓
Writesonic	✓	✓	✓	N/A	N/A
ChatOn	✗	✓	✗	N/A	N/A

While these isolated scenarios allow us to infer *when* contextual leakage occurs, we cannot explain *how* each tracker obtains the conversation context. However, VV8's execution traces record every JavaScript API call a script performs allowing us to link each outbound request to the specific read operation that preceded it. Analyzing these traces, starting from the outbound requests containing the taint values (prompts and titles) we identify the execution script, resolve its origin and enumerate the read operations performed on the DOM before transmission. In our analysis, we correlate reads to transport via script identifier: if the same script context initiates a request towards a third party endpoint and also called `document.title` as a property getter, we consider that a read and transmission chain.

As illustrated by Table 4, the most prevalent mechanism is a property read of `document.title` observed on eight out of the nine platforms. Analytics services such as Google Analytics and Volcengine, as well as session replay tools like Clarity, utilize this standard property to access the specific value and populate the corresponding parameter (e.g., `dt=` or `tl=`) values for outbound requests.

On the other hand, Meta Pixel employs an entirely different approach. On all five platforms where it is present (Zenochat, Genspark, Manus, Quillbot, and Writesonic), its script (`fbevents.js`) never accesses the `document.title` property. Instead, it calls `document.querySelector("title")` to select the `<title>` HTML element from the DOM and then reads its `textContent` to obtain the page title. Interestingly, TikTok Pixel on Writesonic uses both mechanisms: it reads `document.title` as a property and also queries the DOM via `querySelector`.

Microsoft Clarity, present on Zenochat, Writesonic, and Quillbot, captures context through a broader mechanism. Beyond reading `document.title`, Clarity walks the DOM tree and reads `textContent` and `value` properties across page elements as part of its session replay functionality. Thus, Clarity captures not only the conversation title but also visible conversation text rendered in the page, including the *raw prompt* even before a title has been generated.

On Manus, we observe a capture mechanism unique among the platforms we study. Amplitude's autocapture feature records every user click, logging the `textContent` and HTML attributes of the clicked element and its ancestors. When a user clicks a conversation entry in the sidebar, Amplitude captures the `title` attribute of the `<span>` element along with its visible text. Similarly, Sentry's performance monitoring records the CSS selector of UI interactions, which includes the element's `title` attribute.

These mechanisms, deployed by third-party tracking frameworks, were designed for "traditional" websites. Page titles are static descriptors and rarely carry sensitive information. On the other hand, on the web interfaces of AI chatbots, the design pattern of treating each conversation as a navigable sub-page with a dynamically generated, context-derived title that inadvertently transforms these mechanisms to capture and propagate conversational context. Finally, as also observed by the outbound requests, You.com represents a distinct leakage behavior. The platform embeds the user's prompt directly in the initial page URL so any tracker that reads the `location.href` receives the prompt without needing to access the page title at all. Google Ads operating on You.com obtains the prompt through this mechanism, in the values of the `tiba, url` and `ref` parameters.

4.4. Privacy Policies

Finally, we look further into the Privacy and Cookie Policies [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101], [102] of the platforms where we identified contextual leakage. Our findings suggest a more subtle and systemic issue: AI chatbot platforms integrate third-party analytics, advertising, and session replay frameworks under the same assumptions that govern their deployment on traditional websites. In multiple instances we find that the privacy notices state that such frameworks are deployed to collect information about the browsing activity of the user within the platform covering page-views, click events, or session duration.

Notably, several platforms explicitly state in their privacy notices that user data is not sold or shared for targeted advertising purposes, yet we observe third-party advertising and analytics pixels actively operating on their interfaces and collecting conversational context. This discrepancy suggests that platforms may not fully account for the data collection practices of the third-party frameworks they integrate, beyond what is explicitly configured. Third-party tracking pixels and similar analytics frameworks are designed to also capture structural page components in order to measure engagement and attribute conversions. On a conventional e-commerce website for example, such signals carry little sensitive information: a page titled "Men's Running Shoes" or a URL containing `/checkout` reveals commercial intent.

AI chatbot platforms, however, fundamentally alter this assumption. Unlike static web pages, chatbot interfaces dynamically construct some of their structural components from user input. The prompt itself reveals the user's intent, topic of interest, or personal circumstances directly, and conversation titles are auto-generated semantic compressions of the user's prompt that are set as the *title* of the current chat session. Each chat instance is treated as a separate sub-page having a unique title which however encapsulates conversation context. As described in Table 4, the same

TABLE 4. VV8 CONTEXTUAL CAPTURE MECHANISMS USED BY TRACKERS ON AI CHAT PLATFORMS.

Capture Method	Operation Description	Trackers	Platforms
`document.title` (getter)	Property access on `document.title`	Google Analytics, Clarity, Naver, X/Twitter, Amplitude, Volcengine, TikTok Pixel	Zenochat, Writesonic, ChatOn, Genspark, Kimi, Manus, Quillbot, DeepSeek
`document.querySelector("title")`	DOM query for `<title>` element, then reads `textContent`	Meta Pixel, TikTok Pixel	Zenochat, Genspark, Manus, Quillbot, Writesonic
`document.getElementsByTagName ("title")`	DOM query for `<title>` element collection	Google Analytics	Zenochat
`element.textContent` / `element.value`	Walks DOM tree, reads all visible text and form values	Clarity	Zenochat, Writesonic, Quillbot
`element.getAttribute` + `element.textContent`	Reads HTML attributes and text of clicked/interacted DOM elements	Amplitude, Sentry	Manus
`location.href`	Reads page URL which contains prompt directly	Google Ads	You.com

capture mechanisms that trackers are implemented to capture page components describing navigation patterns and commercial interactions now are retrieving conversational context rather than generic page metadata. This structural shift raises significant privacy concerns as we find that privacy policies disclose the operation of these pixels covering pageviews and interaction metrics, with no acknowledgment that on a chatbot, these signals carry a categorically different informational payload.

Key Takeaways

- *Pervasive third-party tracking.* AI chatbot platforms embed conventional third-party tracking and analytics frameworks, which are the same as those deployed across traditional websites.
- *Contextual leakage is systemic.* We audit 42 platforms and find that 21 of them leak conversation titles and nine leak verbatim user prompts to third-party trackers, indicating that contextual leakage is a systemic issue.
- *Leaked conversational context is re-identifiable.* Trackers receiving conversational context frequently also obtain the user's email, allowing even sensitive disclosures to be tied to a named individual.
- *Conventional analytics is a privacy risk on AI chatbots.* Tracking tools on AI chatbot platforms require careful scrutiny: deploying traditional-web analytics on AI chatbots should be treated as a privacy risk, since such tools capture conversational context that users reasonably expect to remain private.

5. Limitations

Experimental Repetitions. Each platform was measured under identical conditions and interaction scenarios, varying only the registration email, between May 6 and May 20, 2026. Our results describe the state of these platforms during that period. Tracker deployments change as platforms add or remove vendors, so the set of platforms leaking conversational context, and the third parties receiving it, may differ outside our measurement window.

Location and Consent. We measured from a single US-East vantage point with tracking consent accepted. This configuration isolates our central question, namely whether contextual leakage occurs and through what mechanism. Consent state and region determine which trackers a platform loads, so we do not study whether these platforms honor consent choices or how their deployments vary across jurisdictions.

Platform Coverage. We study web-based chatbot platforms that offer free account registration, which excludes paid tiers and the growing number of AI chatbots distributed primarily as mobile or desktop applications. Extending this analysis to application-based clients would provide additional insights regarding the state of this ecosystem.

6. Related Work

Model and Application Leakage. A separate line of work has studied the privacy issues and leakage vectors originating from the LLM itself or from applications built on top of it. At the model level, studies have focused on personally identifiable information (PII) memorization in pretrained models [103], [104], multi-step jailbreaks extracting personal information from ChatGPT [105] or inappropriate information disclosure [106]. At the application level, Priyanshuet al. [107] documented prompt regurgitation in chatbots, while Yanet al. [108] studied how file exfiltration content could be achieved through prompt injection in third-party GPT-based applications. Moreover, Shakeel et al. [109] demonstrated that PII retained across sessions by AI chatbots can be exploited to disclose a user's personal information in unrelated conversations or to adversarial actors. These studies address privacy risks inherent to the LLM or its application layer but do not consider the tracking infrastructure embedded in the chatbots' web interfaces, where third-party scripts can silently extract conversation context and share it with third-party actors.

Third-party Tracking on the Web. Privacy in the online ecosystem has been extensively studied with works focused on identifying and measuring the prevalence of third-party actors on the web, both through client-side [5], [6], [7], [8], [9], [110], [30] and server-side [111], [112] tracking mechanisms. Studies have also examined the presence and impact of such trackers on "sensitive" websites [11], [12], [13], [113], [15], highlighting the increased concern when tracking occurs on platforms that handle sensitive user information. For example, Moti et al. [10] measured tracking and advertising components on children-related websites, while Huo et al. [14] explored the disclosure of protected health information to third parties from within web-based healthcare systems.

Complementing these measurement efforts, numerous studies have characterized the PII that trackers collect [27], [28], [29], [30], [31], [32], [34]. Kubicek et al. [33] automated the submission of registration forms to study PII collection by analyzing outbound traffic towards third-party destinations. Kieserman et al. [35] measured such PII leakage by injecting artificial forms into websites and inspecting the network traffic generated upon their submission. More recently, studies have turned to analyzing the structural components of tracking implementations themselves, uncovering *contextual* leakage as an additional privacy risk. Bekos et al. [25] demonstrated that Meta Pixel can collect contextual information describing user interactions with a website's components, while Ghani et al. [26] further examined the structural configuration of Meta Pixel deployments that enable such collection practices.

Third-party Tracking on AI Chatbots. Although AI chatbots have become an integral part of online activity, the presence of third-party trackers on their web interfaces has received limited attention. Existing work has primarily focused on the ChatGPT ecosystem, studying how third-party developers build plugins and apps that interface with the platform and the data collection practices these integrations introduce [22], [23]. Vekaria et al. [114] audited nine GenAI browser extensions for tracking, profiling, and personalization, revealing that extensions can exfiltrate sensitive page content, including form inputs. Concurrently to our work, Jazlan et al. [19] measured content and identity exposure across 20 general-purpose AI chatbots by capturing network traffic during a single-prompt session. Similarly, Girish et al. [20] examined third-party tracking across four chatbot platforms through HAR traffic analysis, documenting the transmission of conversation URLs and user identifiers to advertising and analytics services.

Our work fundamentally differs from prior studies in that we treat tracking as an *interaction-driven* phenomenon. Tracking pixels and analytics frameworks are fundamentally event-driven, designed to trigger on user interactions rather than solely on page load. A single-prompt network capture is therefore insufficient to understand when and how contextual leakage occurs on these platforms. The class-action complaints recently filed against both OpenAI [16], [17] and Perplexity [18] further underscore this point, alleging that embedded trackers transmitted conversational context to third parties upon user *interaction* with the chatbot. Our study addresses these gaps by capturing both contextual and identity leakage towards third-party actors across structured multi-interaction sessions, and then decomposing these sessions into isolated interaction scenarios to attribute each data flow to its causal trigger. Our analysis through VV8 further shows that trackers capture this context through unmodified, generic page-metadata primitives, a structural finding that network-only measurements cannot establish.

7. Discussion

Third-party tracking scripts have been designed to operate on traditional websites, where the structural components they capture, such as page titles and URLs, describe generic user interactions like product views or general content engagement. However, when deployed on AI chatbot platforms under the same default configuration, these same mechanisms retrieve structural components that encode conversational context: the user's prompt and the dynamically generated conversation title derived from it. This distinction is fundamental, as the metadata of a chat session do not include only generic page information but also a direct reflection of what the user has disclosed. Deploying such frameworks on AI chatbot platforms therefore requires deliberate consideration of the privacy implications of the page's structural components, and these platforms should not be treated as traditional websites when configuring third-party tracking integrations.

Our findings reveal a systematic mismatch between the intent of conventional tracking frameworks and the context in which they are deployed on AI chatbot platforms. Standard client-side privacy tools remain relevant here, as blocking these scripts would prevent the observed leakage. A more principled long-term direction, however, is the development of AI-specific analytics frameworks that decouple usage measurement from conversational content, allowing operators to collect meaningful platform metrics without exposing user intent to third parties. Such an outcome could be approached from two directions. Platform developers could minimize the embedding of conversational context in dynamic page components, for example by generating less specific conversation titles that omit sensitive details rather than directly reflecting the user's query. Third-party vendors, on the other hand, could develop dedicated analytics solutions for conversational interfaces, or extend existing frameworks with stricter privacy constraints that prevent the collection of content-bearing page attributes in chatbot contexts.

Furthermore, when conducting privacy audits of such platforms, developers should evaluate the full range of user interactions rather than limiting their assessment to page load or prompt submission events. Tracking pixels are event-driven by nature and may fire on interactions such as navigation, button clicks, or chat deletion, each of which can independently trigger contextual leakage that would otherwise go undetected.

During the course of our study, ChatGPT and Grok had third party trackers operating on their platforms that were collecting conversational context. Both platforms have since removed these trackers, which we verified prior to submission, suggesting that the issue is actionable and that platforms can remediate it once identified. For all remaining platforms where contextual leakage was observed, we have followed responsible disclosure practices and notified the respective teams. During our responsible disclosure process, four platforms (Manus, Kimi, QuillBot and SciSpace) received our notification and corresponding data and forwarded them to the appropriate teams, two platforms (Genspark and Chaton) acknowledged receipt of the disclosure, and one platform (TheLawGPT) directly remediated the issue by updating the configuration of the corresponding tracker responsible for the leakage.

8. Conclusion

In this study, we present a systematic measurement study of contextual leakage in AI chatbot platforms, auditing 42 platforms spanning general-purpose and task-specific chatbots in sensitive domains. We find that contextual leakage is a systemic phenomenon, with 21 platforms transmitting conversation titles and 9 transmitting verbatim user prompts to third-party trackers. Beyond content leakage, the majority of context-leaking platforms also expose the user's email to the same trackers, enabling direct linkage of conversational context to a named individual. Through VV8-based interaction decomposition, we attribute each leakage flow to its causal trigger and reconstruct the JavaScript-level mechanisms through which trackers extract conversational context from the page.

Our findings highlight a fundamental mismatch between how AI chatbot platforms are instrumented and the nature of the content they handle. Unlike traditional websites, chatbot interfaces dynamically encode user intent into structural page components, causing standard tracking integrations to inadvertently capture even *sensitive* user disclosures. Therefore, these platforms should be carefully evaluated for third-party tracking deployments in this context and privacy audits must cover the full range of user interactions, as leakage may manifest under different interaction events.

9. Acknowledgments

We would like to thank the anonymous reviewers for their constructive comments and feedback. This work was supported by the National Science Foundation (NSF) under grant CNS-2211575, with additional support by Amazon through the Amazon Research Awards program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the funding agencies and sponsoring organizations that supported this work.

References

[1] Backlinko, "ChatGPT / OpenAI Statistics: How Many People Use ChatGPT?" https://backlinko.com/chatgpt-stats.

[2] Instant Press, "AI & ChatGPT Statistics for 2026," https://www.instantpress.co/ai-statistics.

[3] Elfsight, "Chatbot Statistics 2025–2026: Market Size, Usage & Trends," https://elfsight.com/blog/chatbot-statistics-and-trends/.

[4] Brookings, "How are Americans using AI? Evidence from a nationwide survey," https://www.brookings.edu/articles/how-are-americans-using-ai-evidence-from-a-nationwide-survey/.

[5] P. Bekos, P. Papadopoulos, E. P. Markatos, and N. Kourtellis, "The hitchhiker's guide to Facebook web tracking with invisible pixels and click IDs," in *Proceedings of the 32nd ACM Web Conference (WWW)*, 2023.

[6] I. Fouad, N. Bielova, A. Legout, and N. Sarafijanovic-Djukic, "Missed by filter lists: Detecting unknown third-party trackers with invisible pixels," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2020.

[7] Q. Chen, P. Ilia, M. Polychronakis, and A. Kapravelos, "Cookie swap party: Abusing first-party cookies for web tracking," in *Proceedings of the 30th ACM Web Conference (WWW)*, 2021.

[8] S. Munir, S. Siby, U. Iqbal, S. Englehardt, Z. Shafiq, and C. Troncoso, "CookieGraph: Understanding and detecting first-party tracking cookies," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2023.

[9] S. Dambra, I. Sanchez-Rola, L. Bilge, and D. Balzarotti, "When Sally met trackers: Web tracking from the users' perspective," in *Proceedings of the 31st USENIX Security Symposium*, 2022.

[10] Z. Moti, A. Senol, H. Bostani, F. Z. Borgesius, V. Moonsamy, A. Mathur, and G. Acar, "Targeted and troublesome: Tracking and advertising on children's websites," in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2024.

[11] X. Yu, N. Samarasinghe, M. Mannan, and A. Youssef, "Got sick and tracked: Privacy analysis of hospital websites," in *Proceedings of the IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 2022.

[12] J. Chanenson, B. Sloane, N. Rajan, A. Morril, J. Chee, D. Y. Huang, and M. Chetty, "Uncovering privacy and security challenges in K-12 schools," in *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, 2023.

[13] N. Samarasinghe, P. Kapoor, M. Mannan, and A. Youssef, "No salvation from trackers: Privacy analysis of religious websites and mobile apps," in *Proceedings of the Data Privacy Management, Cryptocurrencies and Blockchain Technology: ESORICS International Workshops*, 2023.

[14] M. Huo, M. Bland, and K. Levchenko, "All eyes on me: Inside third party trackers' exfiltration of PHI from healthcare providers' online systems," in *Proceedings of the 21st Workshop on Privacy in the Electronic Society*, 2022.

[15] S. Rauti, E. Vuorinen, R. Carlsson, and P. Puhtila, "Data leaks to third-party services on medical websites," in *Proceedings of the 16th International Conference on Security of Information and Networks (SIN)*, 2023.

[16] "Lim v. OpenAI global, LLC," May 2026.

[17] "Couture v. OpenAI global, LLC," May 2026.

[18] "Doe v. Perplexity AI, inc., Meta platforms, inc., and Google LLC," Mar. 2026.

[19] M. Jazlan, E. Wang, Y. Vekaria, and Z. Shafiq, "Tracking conversations: Measuring content and identity exposure on ai chatbots," 2026.

[20] A. Girish, G. Oliveira, G. Suarez-Tangil, J. G. Herrero, M. Sanchez, N. Vallina-Rodriguez, and T. Jackevicius, "LeakyLM: AI Assistants Are Leaking Your Conversations," https://leakylm.github.io/, 2026.

[21] J. Jueckstock and A. Kapravelos, "VisibleV8: In-browser monitoring of javascript in the wild," in *Proceedings of the Internet Measurement Conference (IMC)*, 2019.

[22] C. Yan, R. Ren, M. H. Meng, L. Wan, T. Y. Ooi, and G. Bai, "Exploring ChatGPT app ecosystem: Distribution, deployment and security," in *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2024.

[23] Y. Wu, E. Jaff, K. Yang, N. Zhang, and U. Iqbal, "An in-depth investigation of data collection in LLM app ecosystems," in *Proceedings of the ACM Internet Measurement Conference (IMC)*, 2025.

[24] KOI Research, "8 Million Users' AI Conversations Sold for Profit by "Privacy" Extensions," https://www.koi.ai/blog/urban-vpn-browser-extension-ai-conversations-data-collection.

[25] P. Bekos, P. Papadopoulos, N. Kourtellis, and M. Polychronakis, "PIIxel Leaks: Passive identification of personally identifiable information leakage through Meta Pixel," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2025.

[26] A. Ghani, Y. Vekaria, and Z. Shafiq, "PixelConfig: Longitudinal measurement and reverse-engineering of Meta Pixel configurations," in *Proceedings of the ACM Internet Measurement Conference (IMC)*, 2026.

[27] H. Dao and K. Fukuda, "Alternative to third-party cookies: Investigating persistent PII leakage-based web tracking," in *Proceedings of the 17th International Conference on Emerging Networking EXperiments and Technologies (CoNEXT)*, 2021.

[28] J. Brookman, P. Rouge, A. Alva, and C. Yeung, "Cross-device tracking: Measurement and disclosures," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2017.

[29] S. Englehardt, J. Han, and A. Narayanan, "I never signed up for this! Privacy implications of email tracking," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2018.

[30] G. Acar, S. Englehardt, and A. Narayanan, "No boundaries: Data exfiltration by third parties embedded on web pages," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2020.

[31] O. Starov, P. Gill, and N. Nikiforakis, "Are you sure you want to contact us? Quantifying the leakage of PII via website contact forms," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2016.

[32] A. Senol, G. Acar, M. Humbert, and F. Z. Borgesius, "Leaky forms: A study of email and password exfiltration before form submission," in *Proceedings of the 31st USENIX Security Symposium*, 2022.

[33] K. Kubicek, J. Merane, A. Bouhoula, and D. Basin, "Automating website registration for studying GDPR compliance," in *Proceedings of the ACM Web Conference (WWW)*, 2024.

[34] M. Chatzimpyrros, K. Solomos, and S. Ioannidis, "You shall not register! Detecting privacy leaks across registration forms," in *Computer Security*. Springer-Verlag, 2019.

[35] J. B. Kieserman, A. Andreou, C. Geeng, T. Lauinger, and D. McCoy, "Tracker installations are not created equal: Understanding tracker configuration of form data collection," in *Proceedings of the Privacy Enhancing Technologies Symposium (PETS)*, 2025.

[36] "Top ai chatbots and tools websites ranking," https://www.similarweb.com/top-websites/ai-chatbots-and-tools/, 2026.

[37] Microsoft, "Microsoft Copilot," https://copilot.microsoft.com/, 2025.

[38] DeepSeek, "DeepSeek," https://chat.deepseek.com/, 2025.

[39] Anthropic, "Claude AI," https://claude.ai/, 2025.

[40] Mistral AI, "Le Chat Mistral," https://chat.mistral.ai/chat.

[41] TextCortex AI, "Zenochat," https://textcortex.com/.

[42] Meta, "Meta AI," https://www.meta.ai/, 2025.

[43] Google, "character.ai," https://character.ai/.

[44] Learneo, "QuillBot," https://quillbot.com/.

[45] Manus, "manus," https://manus.im/.

[46] Hugging Face, "Hugging Chat," https://huggingface.co/chat/.

[47] Alibaba Cloud, "Qwen," https://qwen.ai/home.

[48] xAI, "Grok," https://grok.com/, 2025.

[49] Writesonic, "Writesonic," https://app.writesonic.com/.

[50] SuSea, Inc., "you.com," https://you.com/chat.

[51] Moonshot AI, "Kimi AI," https://www.kimi.com/.

[52] Perplexity AI, "Perplexity," https://www.perplexity.ai/, 2025.

[53] OpenAI, "ChatGPT," https://chatgpt.com/, 2025.

[54] Quora, "Poe," https://poe.com/, 2025.

[55] Duck Duck Go, Inc., "Duck.ai," https://duck.ai/.

[56] SUNRISEAI PTE. LTD, "SEAART AI," https://www.seaart.ai/.

[57] Inflection AI, "Pi," https://pi.ai.

[58] Mainfunc Inc., "Genspark AI Workspace," https://www.genspark.ai/.

[59] AIBY, "ChatOn," https://chaton.ai/.

[60] Cloud Whale Interactive Technology LLC, "PolyBuzz," https://www.polybuzz.ai/.

[61] Google, "Gemini," https://gemini.google.com/app, 2025.

[62] Khan Academy, "Khanmigo," https://www.khanmigo.ai/.

[63] EduBrain, "EduBrain," https://edubrain.ai/.

[64] CK-12 Foundation, "Flexi," https://www.ck12.org/flexi/.

[65] SciSpace, "SciSpace," https://scispace.com/.

[66] Elicit AI, "Elicit," https://elicit.com/.

[67] AI Search Inc., "iAsk AI," https://iask.ai/.

[68] Answer.AI, "Answer.AI," https://answerai.pro/.

[69] Free AI (not a) Therapist, "Free AI (not a) Therapist," https://www.freeainotatherapist.com/.

[70] FeelBetterBot, "FeelBetterBot," https://feelbetterbot.com/.

[71] Wellzy Inc., "Wellzy," https://wellzy.io/.

[72] AI Search Inc., "iAsk AI - Companion Mode," https://iask.ai/companion/.

[73] Earkick AI, "Earkick," https://chat.earkick.com/.

[74] Abby Intelligence, Inc., "Abby," https://abby.gg/.

[75] AITherapistFree, "AITherapistFree," https://www.aitherapistfree.com/.

[76] LawConnect, "LawConnect," https://lawconnect.com/.

[77] NyayGuru AI, "NyayGuru," https://nyayguru.com/.

[78] AI Lawyer, "AI Lawyer," https://app.ailawyer.pro/.

[79] TheLawGPT, "TheLawGPT," https://www.thelawgpt.com/.

[80] "Puppeteer," https://pptr.dev/, 2018.

[81] Disconnect, "Tracker Protection Lists," https://github.com/disconnectme/disconnect-tracking-protection.

[82] DuckDuckGo, "DuckDuckGo Tracker Radar," https://github.com/duckduckgo/tracker-radar.

[83] Character.AI, "Character.AI Privacy Policy," https://policies.character.ai/privacy, 2025.

[84] OpenAI, "Privacy policy," https://openai.com/policies/row-privacy-policy/, 2025.

[85] Chaton, "Privacy Policy," https://chaton.ai/legal/privacy/en/, 2025.

[86] Deepseek, "DeepSeek Privacy Policy," https://cdn.deepseek.com/policies/en-US/deepseek-privacy-policy.html, 2025.

[87] Genspark, "Privacy Policy," https://www.genspark.ai/privacy, 2025.

[88] xAI, "xAI Privacy Policy," https://x.ai/legal/privacy-policy, 2025.

[89] Moonshot AI PTE. LTD, "Privacy Policy," https://www.kimi.com/user/agreement/userPrivacy?version=v2, 2025.

[90] Butterfly Effect Pte. Ltd, "Privacy," https://manus.im/privacy, 2025.

[91] Polybuzz, "Privacy Policy," https://www.polybuzz.ai/privacy-policy, 2025.

[92] Quillbot, "Quillbot's Privacy Policy," https://quillbot.com/privacy, 2025.

[93] Writesonic, "Privacy Policy," https://writesonic.com/legal/privacy-policy, 2025.

[94] You.com, "Privacy Policy," https://you.com/privacy, 2025.

[95] TextCortex, "TextCortex Privacy Policy," https://textcortex.com/privacy-policy, 2025.

[96] AI Lawyer, "Privacy Policy," https://ailawyer.pro/privacy-policy, 2025.

[97] Answer.AI, "Answer.AI Privacy Policy," https://answerai.pro/privacy-policy.html, 2025.

[98] EduBrain, "Privacy Policy," https://edubrain.ai/privacy-policy/, 2025.

[99] Elicit AI, "Privacy Policy," https://elicit.com/operations/privacy, 2025.

[100] SciSpace, "Privacy," https://scispace.com/privacy, 2025.

[101] TheLawGPT, "Privacy Policy," https://www.thelawgpt.com/privacy-policy, 2025.

[102] AI Search Inc., "Privacy Policy," https://iask.ai/privacy, 2025.

[103] S. Kim, S. Yun, H. Lee, M. Gubri, S. Yoon, and S. J. Oh, "ProPILE: Probing privacy leakage in large language models," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

[104] S. Zanella-Béguelin, L. Wutschitz, S. Tople, V. Rühle, A. Paverd, O. Ohrimenko, B. Köpf, and M. Brockschmidt, "Analyzing information leakage of updates to natural language models," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2020.

[105] H. Li, D. Guo, W. Fan, M. Xu, J. Huang, F. Meng, and Y. Song, "Multi-step jailbreaking privacy attacks on ChatGPT," in *Findings of the Association for Computational Linguistics (EMNLP)*, 2023.

[106] N. Mireshghallah, H. Kim, X. Zhou, Y. Tsvetkov, M. Sap, R. Shokri, and Y. Choi, "Can LLMs keep a secret? Testing privacy implications of language models via contextual integrity theory," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2024.

[107] A. Priyanshu, S. Vijay, A. Kumar, R. Naidu, and F. Mireshghallah, "Are chatbots ready for privacy-sensitive applications? An investigation into input regurgitation and prompt-induced sanitization," 2023.

[108] C. Yan, B. Guan, Y. Li, M. H. Meng, L. Wan, and G. Bai, "Understanding and detecting file knowledge leakage in GPT app ecosystem," in *Proceedings of the ACM Web Conference (WWW)*, 2025.

[109] A. Shakeel and S. Thaseen, "Investigating leakage of sensitive information in LLM-powered chatbots," in *Proceedings of the International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, 2025.

[110] J. R. Mayer and J. C. Mitchell, "Third-party web tracking: Policy and technology," in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2012.

[111] A. El fraihi, N. Amieur, W. Rudametkin, and O. Goga, "Client-side and server-side tracking on Meta: Effectiveness and accuracy," in *Proceedings on Privacy Enhancing Technologies Symposium (PETS)*, 2024.

[112] I. Fouad, C. Santos, and P. Laperdrix, "The devil is in the details: Detection, measurement and lawfulness of server-side tracking on the web," in *Proceedings on Privacy Enhancing Technologies Symposium (PETS)*, 2024.

[113] A. B. Friedman, R. M. Merchant, A. Maley, K. Farhat, K. Smith, J. Felkins, R. E. Gonzales, L. Bauer, and M. S. McCoy, "Widespread third-party tracking on hospital websites poses privacy risks for patients and legal liability for hospitals," *Health Affairs (Millwood)*, 2023.

[114] Y. Vekaria, A. L. Canino, J. Levitsky, A. Ciechonski, P. Callejo, A. M. Mandalari, and Z. Shafiq, "Big help or big brother? Auditing tracking, profiling, and personalization in generative AI assistants," in *Proceedings of the 34th USENIX Security Symposium*, 2025.

Appendix A. Extended Findings

Figure 12 presents the email–provider asymmetry of whole-string (title or prompt) leak volume, measured using the per-platform $\log_2$ ratio of leaks under a Gmail vs. non-Gmail account, with one session per configuration. Positive means more leaks under Gmail. Kimi, ZenoChat, ChatGPT, and Manus leak $\sim$12–24$\times$ more under Gmail, with only You.com, Writesonic and Genspark leaning the other way.

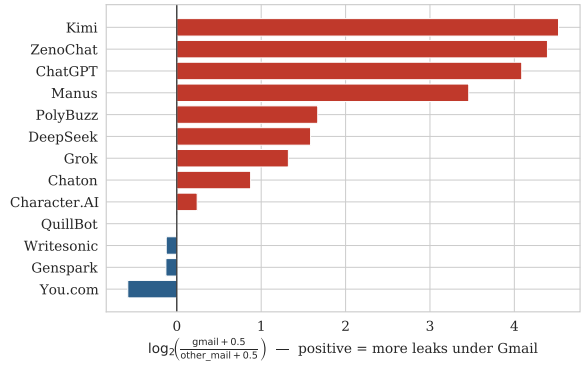


Figure 12.

Figure 13 provides a per-domain breakdown of the data aggregated by company in Figure 8. Google's dominance is distributed across multiple distinct endpoints, with *google.com*, *doubleclick.net*, and *googleadservices.com* each independently accumulating substantial context-bearing requests, underscoring the breadth of its tracking infrastructure across the chatbot ecosystem.

Similarly to our findings for the generic AI chatbots, we find that the ecosystem of third-party companies operating on these platforms is densely populated by three companies: Google, Meta and Microsoft, as shown in Figure 14. These companies account for 70% of all tracking-related requests. More importantly, we find that all chatbot platforms heavily favour the adoption of Google services, as it is the company found in all the network traces.

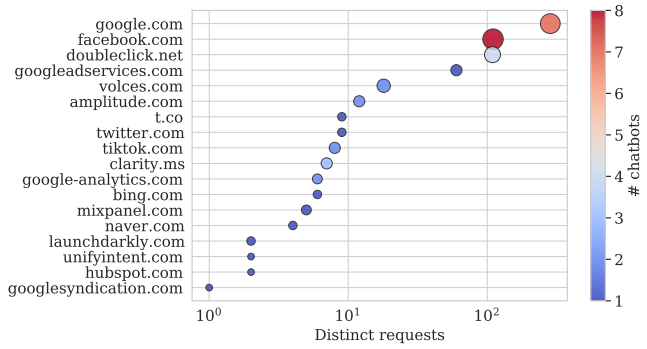


Figure 13. Distinct context-bearing requests to tracker domains.

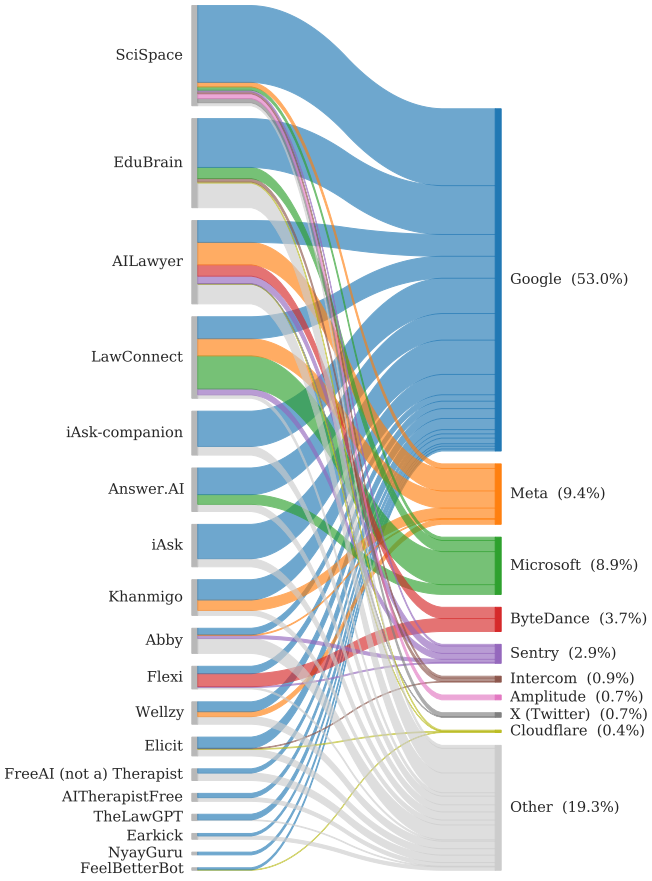


Figure 14. Flow of third-party requests from task-specific chatbots to tracker owner companies.

As illustrated by Figure 15, we find eight task-specific chatbots sharing conventional context, in the form of *titles* with a total of seven tracker-owning companies. Interestingly enough, similarly to You.com, we find a heavy concentration of third party requests that include conversation titles towards Google from SciSpace. In general, we find heavier conventional leakage patterns in academic and educational chatbots. Finally, as illustrated by Figure 16, we also find five instances of full *prompt* leakage towards four companies

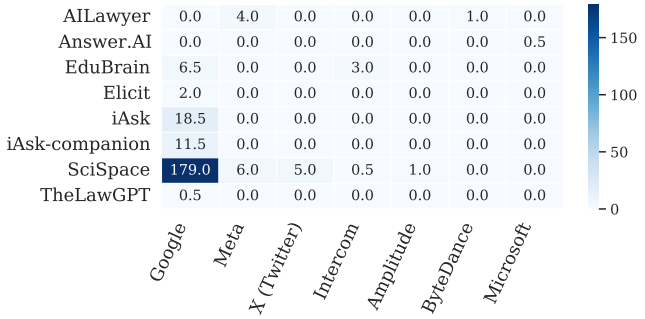


Figure 15. Whole-string title leaks per task-specific chatbot × owner company, averaged per prompt.

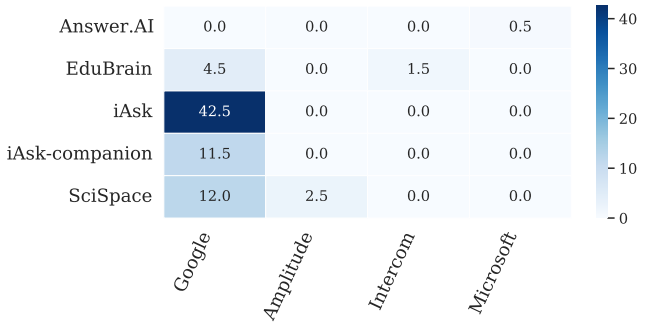


Figure 16. Whole-string prompt leaks per task-specific chatbot × owner company, averaged per prompt.

operating third-party analytics and tracking scripts on these platforms.

Appendix B. Open Science

We provide an artifact accompanying this paper at https://zenodo.org/records/22831655. It includes (i) the network traffic logs collected across all 42 platforms and (ii) the VisibleV8 JavaScript execution traces, each accompanied by the analysis scripts that reproduce our findings.

Appendix C. AI Usage

In this paper, we conduct a passive measurement study on the privacy implications of interacting with online AI chatbot platforms. As such, AI platforms are the object of the study. Specifically, we audit 42 general-purpose and task-specific chatbots for third-party leakage of conversational context by passively measuring the requests that these platforms send during ordinary use. In terms of using generative AI as a tool in the development of this work, we employed Anthropic's Claude Opus 4.7 to generate the prompts that were used in the interactions with the AI chatbot platforms, and to develop and refine our data analysis pipeline. All generative AI outputs were manually verified

and meticulously reviewed. Generative AI was used for editorial purposes in this manuscript, and all outputs were inspected by the authors to ensure accuracy and originality. Specifically, we used Claude to tighten and spellcheck the manuscript, as well as to fix grammar and syntax errors. As above, all editorial generative AI suggestions were inspected by the authors and manually verified.

Appendix D. Ethics Considerations

Our study involves passive measurement of publicly accessible AI chatbot web interfaces. No real user data was collected or processed at any point, as our experiments were conducted using two dedicated accounts created exclusively for this study, with no involvement of real users or their data. Interactions with live platforms were conducted in a controlled, low-volume manner using browser instrumentation (Puppeteer and VisibleV8) across manual sessions. We generated a minimal set of prompts per platform and a minimal number of interactions per session, deliberately avoiding traffic patterns that could disrupt normal platform operations. We do not access, modify, exfiltrate, or interfere with other users' data or platform functionality. Our methodology is strictly observational, recording only outbound network traffic and JavaScript execution traces originating from our own controlled sessions.

Appendix E. Meta-Review

The following meta-review was prepared by the program committee for the 2027 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

E.1. Summary

This paper presents a measurement study of privacy leakage in the web interfaces of AI chatbot platforms, spanning both general-purpose and domain-specific systems. Using a combination of Puppeteer-based network interception and VisibleV8 JavaScript instrumentation, the authors analyze how conversational data and user identifiers are transmitted to third-party trackers, attribute each leak to specific user interactions, and identify the underlying JavaScript mechanisms responsible. Their analysis reveals that leakage of conversation titles and verbatim user prompts are present and that a substantial amount of chatbot platforms leak users' registration email addresses, enabling the association of sensitive conversational content with user identities. The study further shows that many of these leaks arise unintentionally through generic metadata collection, such as the extraction of dynamically generated page titles that contain conversation summaries. The paper highlights widespread privacy risks in current AI chatbot deployments and provides new insights into the interactions and implementation mechanisms that lead to contextual leakage.

E.2. Scientific Contributions

- Identifies an Impactful Vulnerability
- Provides a Valuable Step Forward in an Established Field

E.3. Reasons for Acceptance

1) The paper identifies an important privacy vulnerability in the web interfaces of AI chatbots specifically studying health, legal, and education platforms and providing valuable insight into the privacy implications of sensitive-domain applications.
2) This paper provides a valuable step forward in an established field. Web privacy is an ongoing concern due to the widespread collection and tracking of user data. The paper presents novel characterization of contextual leakage and explains when and how it occurs in the context of AI chatbots. By attributing leaks to specific user interactions and identifying the DOM-read primitives used by individual trackers, the paper provides insights that network-only analyses cannot.

E.4. Noteworthy Concerns

Consistency and reproducibility of the measured behavior: Reviewers raised concerns the paper relies on a limited number of measurements and sessions.