Privacy in the Age of Frontier AI

A Comparative Analysis of Data Practices, Regulatory Enforcement, and Transparency Failures Across Major AI Platforms

Moiz Ahmed
Wzdom Research, Wzdom LLC
Q2 2026 · Published May 2026
moiz@wzdom.ai · wzdom.ai/research

Independent research. No advertising accepted. No payment accepted from any company reviewed in this paper.

Citation and Use Policy

Abstract

This paper examines how twelve AI platforms — OpenAI (ChatGPT), Google (Gemini), Anthropic (Claude), Meta (Meta AI), Microsoft (Copilot), xAI (Grok), DeepSeek, Perplexity AI, Mistral (Le Chat), Ollama, Apple Intelligence, and Amazon Bedrock — collect, retain, and use personal data from their users. Drawing on primary policy documents, published system cards, government transparency reports, regulatory enforcement records, and peer-reviewed academic research, we find that: (1) the majority of major consumer AI providers use conversation data to train AI models by default, with opt out rather than opt in as the dominant model; (2) industrywide transparency is declining, not improving — average Foundation Model Transparency Index scores dropped from 58/100 in 2024 to 40/100 in 2025; (3) data retention periods and training use policies are inconsistently disclosed, frequently incomplete, and in some cases actively misleading; and (4) regulatory enforcement remains fragmented and largely ineffective — the only confirmed GDPR fine against a frontier AI chatbot was annulled by an Italian court in March 2026. We identify Ollama's fully local architecture and Apple's Private Cloud Compute as structural outliers demonstrating that privacy preserving AI at consumer scale is technically achievable. At the other extreme, DeepSeek scores zero across all dimensions — the only platform where the risk is jurisdictional rather than merely a policy failure.

Keywords: *AI privacy, data retention, LLM training data, GDPR enforcement, Foundation Model Transparency Index, opt in consent, membership inference, EU AI Act, ChatGPT, Gemini, Claude, DeepSeek, Perplexity, Mistral, Ollama, local LLMs, privacy scoring*

1. Introduction

Every day, hundreds of millions of people share their thoughts, questions, health concerns, financial situations, and personal relationships with AI chatbots. They are doing so under the reasonable assumption that this information is handled responsibly. The evidence suggests that assumption is largely wrong.

The same systems people confide in are, by default, using those conversations to train the next generation of AI models. Conversations are retained — in some cases indefinitely. A subset are reviewed by human contractors. Data is shared with third party vendors. Government agencies submit legal requests, and companies comply.

None of this is secret. It is, however, buried. The average AI privacy policy runs to thousands of words, changes without notice, and is written to satisfy legal reviewers rather than to inform users. A 2025 survey found that 82% of Americans view AI related data loss as a serious personal threat (Relyance AI, 2025).

Scope: We cover twelve platforms: OpenAI (ChatGPT), Google DeepMind (Gemini), Anthropic (Claude), Meta (Meta AI), Microsoft (Copilot), xAI (Grok), DeepSeek, Perplexity AI, Mistral (Le Chat), Ollama, Apple (Apple Intelligence), and Amazon (Bedrock). Ollama is included as a privacy baseline — a tool for which the questions this paper asks do not apply, and which demonstrates what structural privacy looks like. Where we found no verifiable information, we say so explicitly and treat the gap as a finding in itself.

2. The Technical Privacy Threat Surface

2.1 Memorization

Large language models do not simply learn patterns from training data — they memorize it. Carlini et al. (2021) demonstrated that GPT-2 could be prompted to reproduce verbatim PII: names, phone numbers, email addresses, and unique identifiers. A follow-on study (Carlini et al., 2023) established that memorization scales with model size, data repetition, and context length. A conversation shared with an AI chatbot that is subsequently used for training may later be reproduced in response to a different user's query. This has been demonstrated experimentally on production models.

2.2 Membership Inference

Membership inference attacks (MIA) can determine, with statistical confidence, whether specific data was included in a model's training set (Li et al., 2024; Huang et al., 2025). An entity-level MIA framework (EL-MIA, 2024) demonstrated that specific names, phone numbers, and sensitive attributes can be confirmed as present in a training corpus without directly extracting them.

2.3 The De-identification Problem

All major AI providers claim to apply de-identification before using consumer data for training. However, de-identification is not a solved problem — datasets can be re-identified when cross-referenced with other sources. The Canadian Privacy Commissioner's 2026 investigation into OpenAI found that health information and children's data had been scraped and used for training despite de-identification claims (OPC, 2026).

Core finding: users face a nontrivial risk that sensitive information shared with AI chatbots will be incorporated into model weights, potentially reproducible by future users, and not fully removable after account deletion.

3. Company-by-Company Data Practices

3.1 Summary Comparison

Company	Training Default	Consumer Retention	Human Review	System Card
OpenAI (ChatGPT)	**Opt out**	Indefinite active; 30 day post delete purge	Yes	Yes — GPT-4.1 missing
Google (Gemini)	**Opt out**	18 months; human reviewed: 3 years	Yes (third party)	Delayed on 2.5 Pro
Anthropic (Claude)	**Opt out** (was opt in until Sept 2025)	30 days (non-training); 5 years (opted in)	Limited	Yes — all models
Meta (Meta AI)	**Opt out**	**Not disclosed**	Yes (vendors)	**None** (consumer)
Microsoft (Copilot)	**Opt out**	18 months	Yes	No — narrative reports
xAI (Grok)	**Opt out** + trains on all public X posts	30 days post deletion	Unknown	**None**
Apple Intelligence†	**N/A — no training**	On device: not sent; PCC: not stored	No	Technical papers only
Amazon Bedrock†	**Opt in required**	Not retained for training	No	No
DeepSeek	**No opt out**	Stored in China; duration undisclosed	Unknown	**None**
Perplexity AI	**No training on user queries**	90 days	No	Minimal
Mistral (Le Chat)	**No training via API**	Session (API); 30 days (chat)	No	Research papers only
Ollama / Local LLMs	**N/A — fully local**	None — never transmitted	No	Open source

3.2 OpenAI (ChatGPT)

For free and Plus consumer accounts, conversation data is used for model training by default. A critical and widely misunderstood detail: the 30 day figure frequently cited refers to the post

deletion purge window — active conversations are retained **indefinitely** unless deleted by the user. GPT-4.1 (April 2025) was released without a model card. Canada's OPC found six categories of violations in May 2026, including scraping of health information and children's data (OPC PIPEDA 2026-002). Italy's EUR 15 million GDPR fine was annulled by a Rome court in March 2026.

3.3 Google (Gemini)

Consumer Gemini defaults to 18 month auto deletion, but human reviewed conversations are retained for **up to three years** even after the user has deleted their activity. Human reviewers include third party service providers. Gemini 2.5 Pro was released without a model card in March 2025; the card was published weeks later.

3.4 Anthropic (Claude)

Anthropic was the only frontier AI developer with a genuine opt in model as of May 2025 (King et al., 2025). This changed September 28, 2025, when Anthropic shifted to opt out for consumer accounts. Current default retention: 30 days (non-training). Anthropic applies automated PII filtering before any training use, does not sell user data, and has published government transparency reports for H1 and H2 2024. No regulatory actions or fines as of May 2026. Signed GPAI Code of Practice (August 2025).

3.5 Meta (Meta AI)

Meta AI conversations are used for personalization and — as of December 2025 — for ad targeting, by default. Users in the EU, UK, and South Korea are excluded due to data protection law. **Meta has published no data retention periods for consumer Meta AI.** Meta explicitly refused to sign the EU AI Act GPAI Code of Practice — the only major frontier AI developer to do so — exposing it to full assessment by the European AI Office without the safe harbor accorded to signatories.

3.6 Microsoft (Copilot)

Consumer Copilot stores conversations for 18 months by default and uses data for training. The opt out path is buried: mobile app → profile → Account → Privacy → Training on conversation activity. Uploaded files are not used for training regardless of settings — a meaningful carve-out not present in all competitor policies. Enterprise Microsoft 365 Copilot does not use customer data for training. Signed GPAI Code (August 2025).

3.7 xAI (Grok)

Grok is the most privacy opaque major AI platform: no system card, no model card, no government transparency report, no GDPR legitimate interest assessment, no independent safety evaluation. Grok trains on all conversations **and all public X/Twitter posts** — including from users who have never used Grok. The Irish DPC secured a permanent suspension of EU tweet collection in August 2024 and opened a new investigation in February 2026 covering AI

training and harmful image generation. Australia, Canada, India, Indonesia, Malaysia, the UK ICO, and Ofcom have all opened or signaled cases.

3.8 Apple (Apple Intelligence)

Apple Intelligence runs the majority of features entirely on device. For tasks requiring more compute, Private Cloud Compute (PCC) uses data only to fulfill the immediate request — data is never stored server side. PCC source code is available to independent security researchers. Apple does not use user interactions for training. No regulatory actions as of May 2026.

3.9 Amazon Bedrock†

Amazon Bedrock is Amazon Web Services' managed API service for accessing and deploying foundation models from third party providers — including Anthropic Claude, Meta Llama, Mistral, and Amazon's own Nova and Titan series. Bedrock is a B2B enterprise product, not a consumer chatbot. Its customers are developers and enterprises building AI applications, not end users sharing personal conversations directly with Amazon.

Amazon's default stance is the most privacy protective of any cloud hosted AI service in this review: customer data submitted to Bedrock is not used to train or improve foundation models by default. Opt in is required for any training use — the inverse of every other cloud frontier AI provider except Apple. Conversation data is not retained for model training. No human review of customer inputs for training purposes is documented. No ad targeting. Amazon signed the EU AI Act GPAI Code of Practice in August 2025. No AI specific regulatory actions or fines as of May 2026. *Primary limitation: Amazon has published no system cards for its own Titan and Nova models, and there is no Bedrock specific government data request transparency report. The practical privacy protections are strong; the formal transparency layer lags behind the architecture.*

3.10 DeepSeek

DeepSeek is a Chinese AI company whose R1 and V3 models achieved competitive benchmark performance against OpenAI's frontier models at a fraction of the training cost. Its consumer chatbot at deepseek.com attracted tens of millions of users globally within weeks of launch in January 2025. The privacy implications of that rapid adoption are severe.

DeepSeek's privacy policy states that user data — including conversation history, device identifiers, keystroke patterns, and clipboard content — is stored on servers in the People's Republic of China. Under China's Personal Information Protection Law (PIPL, 2021) and National Intelligence Law (2017), Chinese authorities can compel data disclosure with no requirement for judicial oversight and no avenue for foreign users to contest access. There is no opt out from training use and no disclosed retention period.

Regulatory response was immediate: Italy's Garante banned DeepSeek in January 2025, citing failure to demonstrate a lawful basis for processing EU user data in China — the fastest AI enforcement action in any jurisdiction to date. Ireland's DPC, Australia's privacy regulator, and South Korea's PIPC all launched parallel investigations. The US Navy, NASA, and several state

governments banned DeepSeek from official devices. As of May 2026, DeepSeek remains inaccessible to Italian users and under active investigation in at least five jurisdictions. Transparency is the lowest of any platform in this review: no system card, no model card, no government data request report, no GDPR assessment, and no independent safety evaluation.

3.11 Perplexity AI

Perplexity is a search-native AI assistant that retrieves real time web results and synthesizes answers with citations. Its privacy profile differs structurally from the frontier chatbots reviewed above: Perplexity does not train its own foundation models on user queries, removing the primary privacy risk driver — conversation data used for model training — that defines the worst practices in this review.

Conversation history is retained for 90 days by default. Users can clear history at any time. Perplexity does not use conversation data for ad targeting and explicitly states it does not sell user data. Data is shared with search API providers (Bing, Google, and others) to execute queries — a structural sharing requirement rather than a policy choice. Transparency limitations: no formal system card, no government transparency report, and no GPAI Code of Practice signature. A September 2024 Wired investigation found that Perplexity's crawler was ignoring robots.txt exclusions and reproducing proprietary content verbatim — a copyright concern distinct from user privacy. No regulatory actions as of May 2026.

3.12 Mistral AI (Le Chat)

Mistral AI is a French AI company with two distinguishing characteristics: its products operate under EU jurisdiction by default, and it releases open weight models users can run entirely locally. Le Chat is Mistral's consumer facing chat interface, separate from the API and open weight releases.

Mistral's API documentation states explicitly that customer data submitted via API is not used for training foundation models — one of the cleaner defaults in this review. Le Chat stores conversations for session continuity; retention periods are not prominently disclosed. No human review of conversations is documented. No ad targeting. No third party commercial data sharing beyond infrastructure subprocessors. Mistral signed the EU AI Act GPAI Code of Practice in August 2025 — the only French-headquartered major AI lab to do so. No regulatory enforcement actions as of May 2026. Transparency limitation: the Stanford FMTI 2025 recorded Mistral's transparency score dropping from 55/100 in 2024 to 18/100 in 2025 — the largest single year drop of any company in the index.

3.13 Ollama / Local LLMs

Ollama is an open source tool that enables users to download and run large language models — including Meta's Llama series, Mistral, Google's Gemma, Microsoft's Phi, and dozens of others — entirely on local hardware. No data of any kind leaves the user's device. There is no server, no account, no telemetry, no API call, and no training data collection. The model weights are the product; the user's conversations are not.

Ollama is included in this review because it represents the privacy baseline: what AI privacy looks like when it is structurally impossible for data to be misused, rather than merely promised not to be. Every privacy concern documented in this paper — training data consent, retention, human review, third party sharing, regulatory exposure, government data requests — is eliminated by architecture, not by policy. Practical limitations: local models require sufficient CPU/GPU and RAM (typically 8–16 GB for a capable model), are slower than cloud hosted alternatives, and require some technical comfort to set up. Ollama scores 100/100 — not because it is perfect software, but because the privacy questions this paper asks do not apply to it.

3.14 Privacy Scoring — Wzdom Research Rankings

Each company is scored across six dimensions totaling 100 points. Every score traces to a primary source in this paper. Scores reflect consumer facing products; enterprise only products are noted with †.

Scoring Rubric

Dimension	Weight	Criteria Summary
A. Training Consent	25 pts	25 = no training; 20 = opt in required; 12 = opt out with PII filtering; 8 = opt out with exceptions; 0 = trains on nonconsenting users' data
B. Data Retention	20 pts	20 = no server retention; 16 = <30 days; 12 = 30 days; 8 = 1–6 months; 4 = 6–18 months; 1 = >18 months or indefinite; 0 = not disclosed
C. Transparency	20 pts	4 sub-criteria at 5 pts each: system card for all models; government data request report; GPAI Code signed; no documented failures
D. Regulatory Record	15 pts	15 = clean; 12 = investigated, no violations; 8 = violations remediated; 4 = active investigations + confirmed violations; 0 = multiple active investigations
E. Data Use	10 pts	10 = no ad targeting, no commercial sharing; 7 = no ads, limited contractor sharing; 0 = ad targeting by default
F. User Control	10 pts	10 = on device/enterprise opt in by design; 8 = prominent opt out; 5 = settings-buried opt out; 0 = no meaningful control

Company Scores by Dimension

All 12 platforms analysed. Apple Intelligence and Amazon Bedrock are device/enterprise products noted with † and are not directly comparable to consumer chatbots.

Company	A. Training /25	B. Retention /20	C. Transparency /20	D. Regulatory /15	E. Data Use /10	F. Control /10	TOTAL
Ollama / Local LLMs	25	20	20	15	10	10	**100**
Apple Intelligence†	25	20	16	15	10	10	**96**

Amazon Bedrock†	20	16	10	12	10	8	**76**
Anthropic (Claude)	12	12	20	15	10	7	**76**
Mistral (Le Chat)	22	14	8	13	8	7	**72**
Perplexity AI	22	14	10	12	8	6	**72**
Microsoft (Copilot)	8	4	14	13	7	5	**51**
Google (Gemini)	10	1	18	12	7	3	**51**
OpenAI (ChatGPT)	8	1	18	4	7	5	**43**
Meta (Meta AI)	5	0	0	8	0	2	**15**
xAI (Grok)	2	6	0	0	0	0	**8**
DeepSeek	0	0	0	0	0	0	**0**

Rankings and Grades

Rank	Company	Score	Grade	Key Finding
1	Ollama / Local LLMs	100/100	A	Fully local; no server, no account, no telemetry, no training. Privacy is architectural — not a policy.
2	Apple Intelligence†	96/100	A	No server side retention; no training on user interactions; PCC independently verifiable
3	Amazon Bedrock†	76/100	B	Opt in required; data not retained for training; limited transparency documentation
3	Anthropic (Claude)	76/100	B	Perfect transparency (20/20); clean regulatory record; 30 day default retention; sole prior opt in consumer AI
5	Mistral (Le Chat)	72/100	B	EU jurisdiction; API does not train on customer data; GPAI Code signatory; transparency declining (FMTI: 55→18)
5	Perplexity AI	72/100	B	Does not train on user queries; 90 day retention; no ad targeting; limited formal documentation
7	Microsoft (Copilot)	51/100	D	Signed GPAI Code; responsible AI reports published; opt out buried; 18 month retention
7	Google (Gemini)	51/100	D	Good documentation; human reviewed conversations retained 3 years even after user deletion
9	OpenAI (ChatGPT)	43/100	D	Strong documentation; active conversations retained indefinitely; six category violations (Canada OPC, May 2026)
10	Meta (Meta AI)	15/100	F	No retention policy disclosed; no consumer system card; ad targeting by default; refused GPAI Code

11	xAI (Grok)	8/100	F	No system card; no transparency report; trains on all public X posts without consent; most active regulatory exposure
12	DeepSeek	0/100	F	No opt out; China jurisdiction servers; no retention disclosure; banned in Italy; under investigation in 5+ jurisdictions

† Apple Intelligence is an on device product; Amazon Bedrock is an enterprise B2B API. Neither is a general purpose consumer chatbot. Scores reflect their respective architectures.

4. Regulatory Landscape

4.1 GDPR Enforcement — Limited and Contested

Italy's Garante banned ChatGPT in March 2023; service was restored in April 2023 after remediation. The Garante imposed a EUR 15 million fine in December 2024 — the first GDPR fine against a frontier AI chatbot. A Rome court annulled the fine on March 19, 2026. **As of May 2026, no GDPR fine against a frontier AI chatbot is currently in force.** Italy also banned DeepSeek in January 2025 for storing EU user data in China without legal basis. The EDPB's Opinion 28/2024 (December 2024) provided the first Europe-wide framework for AI training data and GDPR.

4.2 US Enforcement — Sectoral and Indirect

The US has no federal AI specific privacy law. The largest AI related settlement to date: Amazon paid $30.8 million in May/June 2023 — $25 million for Alexa's illegal COPPA violations and $5.8 million for Ring employees' unauthorized access to private camera footage (FTC, 2023). The FTC launched “Operation AI Comply” in September 2024, producing more AI enforcement actions in 2024 than in the prior three years combined. Biden's EO 14110 (AI safety mandates) was revoked January 20, 2025; Trump's EO 14179 removes restrictions and promotes US AI leadership.

4.3 Canada — Most Recent Significant Enforcement

On May 6, 2026, Canada's OPC with three provincial regulators published findings from a two-year investigation into OpenAI: six categories of violations including scraping of health information and children's data without consent. BC and Alberta refused to resolve, stating retroactive consent cannot remedy the initial violation (OPC PIPEDA 2026-002).

4.4 EU AI Act — Forward-Looking Framework

The EU AI Act entered into force August 2024. The GPAI Code of Practice was endorsed August 2025 with 26 signatories including Amazon, Anthropic, Google, IBM, Microsoft, and OpenAI. Meta refused to sign. **GPAI enforcement begins August 2, 2026** with penalties up to EUR 15 million or 3% of global revenue.

5. The Transparency Deficit

Stanford FMTI average scores: October 2023 = 37/100; May 2024 = 58/100; December 2025 = **40/100**. Meta dropped from 60 to 31; Mistral from 55 to 18. Data practices remain the most opaque category industrywide. Specific documented gaps: xAI has no system card, no model card, no transparency report, and no GDPR assessment — the largest transparency gap of any frontier AI company. Meta has no consumer system card and no retention policy. OpenAI released GPT-4.1 without a model card. Google published the Gemini 2.5 Pro model card weeks late.

6. Industry Standards: Claimed vs. Verified

Standard	Status	Confirmed Companies	Notes
ISO/IEC 42001	Published Dec 2023	IBM (certified Nov 2025), Microsoft (third party audits)	No public certification from OpenAI, Google, Anthropic, Meta, xAI, Apple, or Amazon
NIST AI RMF 1.0 + GenAI Profile	Jan 2023 / July 2024	None confirmed	Voluntary, self-declared; only 24% of GenAI projects include recommended security measures
EU AI Act GPAI Code	Endorsed Aug 2025	Amazon, Anthropic, Google, IBM, Microsoft, OpenAI, Mistral, Cohere (26 total)	Meta explicitly refused; signatories receive presumption of conformity
Frontier Model Forum	Founded July 2023	Google, Microsoft, OpenAI, Anthropic	Self-regulatory; no binding certification

7. The Apple Outlier and What It Proves

Apple's approach demonstrates that privacy preserving AI at consumer scale is technically achievable. Key decisions: on device processing where possible; PCC for server side tasks with stateless processing (data used then removed, never stored); PCC source code available to independent researchers; no training on user interaction data. The result: Apple Intelligence has zero regulatory exposure on AI training data. Apple has demonstrated that the privacy cost tradeoff in AI is not inevitable — it is a design choice.

8. Conclusions and Recommendations

8.1 Summary Findings

- **Finding 1:** Every major cloud based frontier AI chatbot uses consumer conversation data to train models by default. No major cloud frontier AI provider uses a genuine opt in model as of May 2026.
- **Finding 2:** Retention periods are inconsistent and incompletely disclosed. Google retains human reviewed conversations for up to three years after user deletion. Meta discloses no retention period.
- **Finding 3:** Most transparent: Anthropic. Most opaque: xAI. Most privacy protective

architecture: Apple. Greatest regulatory risk: Meta and xAI.

- **Finding 4:** Regulatory enforcement has not produced a deterrent result. The only GDPR fine against a frontier AI chatbot was annulled.
- **Finding 5:** Industry transparency is declining. Stanford FMTI scores dropped 18 points in one year.

8.2 Recommendations

For policymakers: Mandate opt in for training data use; require standardized AI privacy disclosures; establish binding retention limits; require AI specific government request reporting; treat model cards as legal compliance documents.

For AI developers: Publish training data transparency reports; make opt out mechanisms prominent; publish legitimate interest assessments when relying on Article 6(1)(f) GDPR.

For users: Assume free tier conversations train the model unless opted out. Use Temporary Chat modes for sensitive queries. Exercise GDPR/CCPA rights. For maximum privacy: Apple Intelligence (on device) or Anthropic Claude (best consumer cloud option).

References

Carlini, N., Tramèr, F., Wallace, E., et al. (2021). Extracting Training Data from Large Language Models. *30th USENIX Security Symposium*, pp. 2633–2650. https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting

Carlini, N., et al. (2023). Quantifying Memorization Across Neural Language Models. *ICLR 2023*. arXiv:2202.07646.

King, C., Klyman, J., Capstick, A., Saade, R., & Hsieh, G. (2025). User Privacy and Large Language Models. *AIES 2025*. arXiv:2509.05382.

Li, Q., et al. (2024). LLM-PBE: Assessing Data Privacy in Large Language Models. *VLDB 2024*, vol. 17.

Bommasani, R., et al. (2024). The Foundation Model Transparency Index v1.1. Stanford CRFM. https://crfm.stanford.edu/fmti/May-2024/index.html

Wan, A., et al. (2025). The 2025 Foundation Model Transparency Index. Stanford CRFM. https://crfm.stanford.edu/fmti/paper.pdf

Stanford HAI. (2025). AI Index Report 2025. https://hai.stanford.edu/ai-index/2025-ai-index-report/responsible-ai

AI Now Institute. (2025). Artificial Power: 2025 Landscape Report. https://ainowinstitute.org/publications/research/ai-now-2025-landscape-report

Apple Security Research. (2024). Private Cloud Compute. https://security.apple.com/blog/private-cloud-compute/

Office of the Privacy Commissioner of Canada. (2026). PIPEDA 2026-002. https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2026/pipeda-2026-002/

Relyance AI. (2025). Consumer AI Trust Survey 2025. https://www.relyance.ai/consumer-ai-trust-survey-2025

EDPB. (2024). Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models. https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-art-64/opinion-282024-certain-data-protection-aspects-related_en

FTC. (2023). FTC and DOJ charge Amazon with violating children's privacy law. https://www.ftc.gov/news-events/news/press-releases/2023/05/ftc-doj-charge-amazon-violating-childrens-privacy-law-keeping-kids-alexa-voice-data-years-deleting

Huang, F., et al. (2025). Membership Inference Attacks on Large-Scale Models: A Survey. arXiv:2503.19338.

NIST. (2024). AI Risk Management Framework: Generative AI Profile (NIST-AI-600-1). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

---