Attention: Your Conversational Data is What They Need

Author(s):
Merane, Jakob

Publication date:
2024-07

Permanent link:
https://doi.org/10.3929/ethz-b-000689289

Attention: Your Conversational Data is What They Need

Jakob Merane [1]

Abstract

As Generative AI chatbots become widely used and increasingly multimodal, an unprecedented amount of conversational data is being generated. The use of this data by model developers for further training is a critical and timely issue. This paper investigates, through a legal lens, how leading providers—OpenAI, Google, Mistral AI, and Anthropic—use conversational data, and it identifies potential privacy-related legal shortcomings. We also situate the findings within the larger framework of law and technology.

1. Introduction

The launch of ChatGPT, the first consumer-facing large language model (LLM), made the much-anticipated AI revolution tangible for the general public. With its impressive ability to provide tailored responses, ChatGPT quickly captivated users (Wu et al., 2023). Many credited its usability to the integration of supervised fine-tuning and reinforcement learning from human feedback applied to the foundation model (Ouyang et al., 2022). With approximately 100 million weekly active users (OpenAI, 2024a), OpenAI now possesses a vast amount of conversational data, which is valuable for further fine-tuning. However, this vast trove of data has raised concerns about data privacy.

Recent research has demonstrated that LLMs can memorize and leak data from both pre-training datasets (Nasr et al., 2023) and fine-tuning datasets (Borkar, 2023). This has led some companies to fear that proprietary information could be leaked, prompting them to restrict the use of these models (Tilley and Kruppa, 2023). While much scholarly attention has focused on how leakage implicates intellectual property (Sag, 2023; Samuelson, 2023), less attention has been paid to the privacy implications of using conversational data for fine tuning. Recently, this issue has sparked heated debate as highlighted by the backlash from Slack users (Belanger, 2024) and Meta's sudden pause in its plan to use user data to train AI models (Meta, 2024).

In this work, we outline how Generative AI chatbots use conversational data. By *conversational data*, we mean all user inputs (prompts and other content) and the LLM-generated output. Our analysis in Section 2 reveals that most chatbots use this data by default. Section 3 presents a legal analysis and includes a benchmark that highlights potential legal shortcomings. Section 4 situates our findings within the larger framework of law and technology.

2. Background

In this section, we briefly outline the launch of four Generative AI chatbots—*ChatGPT*, *Gemini*, *LeChat*, and *Claude*—in Europe.[1] Given that ChatGPT was the first and most widely used (Shafer, 2024), we focus particularly on OpenAI. We analyze the approaches these companies take in using conversational data for further model training. Specifically, we investigate whether they employ an *opt-in* or *opt-out* approach for data usage. *Opt-in* means users must give their consent beforehand, whereas in an *opt-out* approach, companies use the data unless the user actively opts out. We focus exclusively on consumer-facing chat interfaces, excluding data processed through APIs.

2.1. OpenAI

OpenAI's approach to user data evolved over time. Upon its launch on 30 November 2022, ChatGPT allowed the use of conversations for further model training, without providing an opt-out option. However, OpenAI stated they would remove personally identifiable information from the data. This was also true with previous instruct beta models (Ouyang et al., 2022), where OpenAI stated human annotators would remove personal data.

In December 2022, OpenAI swiftly introduced an email-based opt-out option for API users. However, progress for ChatGPT users was slower. Significant changes only began after the Italian Data Protection Agency (Garante)

[1]ETH Zurich, Center for Law and Economics, Zurich, Switzerland. Correspondence to: Jakob Merane <jakob.merane@gess.ethz.ch>.

Accepted to the 2^{nd} *Workshop on Generative AI and Law*, co-located with the *International Conference on Machine Learning*, Vienna, Austria. 2024.

[1]Information sourced from current company websites, archived websites on the Internet Archive, and review of terms of service and privacy policies.

opened an investigation. In March 2023, the Italian Garante imposed a temporary processing limitation, arguing, inter alia, that Italian users should have a right to opt out from their data being used (Garante, 2023; Chiara, 2023). On April 25, 2023, OpenAI announced a new setting allowing users to deactivate training (OpenAI, 2023). Before this update, OpenAI reverted to using a Google Form for opt-outs, which was suprising for a company valued at over 30 billion USD at the time (Glasner, 2023). On April 28, 2023, ChatGPT became accessible to Italian users again, although the investigation is still ongoing (Garante, 2024).

Despite introducing the opt-out option, criticism persisted. One issue was that opting out of model training prevented retaining chat history. In early 2024, OpenAI introduced temporary chats, allowing users to disable data usage for training while saving history outside a session-based window (OpenAI, 2024b). By disentangling model training and history, OpenAI may have addressed some criticism. Another criticism emerged from January to April 2024, when users on various platforms, including the OpenAI Community Forum, reported a bug that could only be resolved by disabling the opt-out setting (OpenAI Community, 2024). OpenAI's delayed response led to speculation that the friction was intentional to coerce users into sharing data.

2.2. Google

Google entered the market later, launching its Generative AI chatbot BARD in the US in March 2023, which was rebranded as Gemini in 2024. Initially, Google planned a European launch for June 2023 but postponed it to mid-July 2023 after discussions with the Irish Data Protection Commission (Data Protection Commission). One concern was the use of conversational data for further model training without offering an opt-out option.

However, with the delayed launch in Europe, Google introduced a new Privacy Hub allowing users to opt out of model training. Furthermore, they explain that human reviewers might process conversations if users do not opt out. As safeguards, they separate conversations from user identities and advise users not to share anything in the chat they would not want seen by a reviewer. Additionally, only a random portion of all conversations are reviewed.

2.3. Mistral AI

The French AI start-up Mistral AI, known for its open-source models, launched LeChat on February 26, 2024 (Mistral AI Team, 2024). LeChat allows users to opt out only if they subscribe to the paid chat service, which is currently not yet launched; otherwise, there is no opt-out option. Users are asked to pseudonymize prompts to exclude personal data from training data. It is unclear what privacy safeguards Mistral AI has in place for LeChat.

2.4. Anthropic

Anthropic launched its Generative AI chatbot Claude in Europe on May 14, 2024. Anthropic, which aims to develop safe and ethical AI systems, does not use conversational data from users by default for further model training. It only does so if a user opts in, although currently, there is no opt-in option. This approach clearly sets Anthropic apart from other Generative AI chatbot providers.

2.5. Summary

In Figure 1, we present the timeline of chatbot product launches by the Generative AI chatbot companies in Europe. The patterns and colors used in the chart indicate the status of the opt-out option: **white** signifies no opt-out option available, **vertical lines** indicate an opt-out option that is available but not fully functional, and **grey** represents a fully functional opt-out option, or for Anthropic, an opt-in.

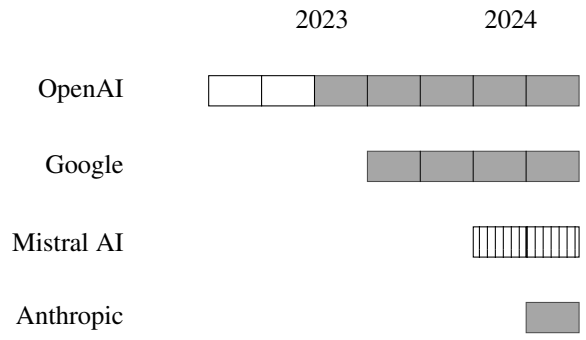

Figure 1. Timeline of product launches by Generative AI companies, grouped by quarters from November 2022 to July 2024.

3. European Legal Framework

The legal landscape governing generative AI chatbots is still in its infancy and continuously evolving. We begin by examining the relevant provisions of the EU's General Data Protection Regulation (GDPR), a comprehensive European privacy law. Next, we explore the forthcoming AI Act. Finally, we present a legal benchmarking framework, based on GDPR criteria, to assess company compliance specifically in using conversational data for further model training.

3.1. The General Data Protection Regulation

Applicability. Many users include personal data in their prompts. Therefore, using user conversations likely involves processing personal data, placing it within the material scope of the GDPR's (Dieker, 2024). Companies cannot evade compliance merely by including a clause in their terms requesting users to avoid inputting personal data. The responsibility of complying remains with the company, re-

gardless of such disclaimers.

Mistral AI, based in Paris, clearly falls under GDPR jurisdiction. The GDPR may have extraterritorial application when US-based companies serve European users.[2] Google and Anthropic both fall under the jurisdiction of the Irish Data Protection Commission as lead regulator, as their main EU establishments are in Ireland. OpenAI, however, did not initially establish a European presence prior to launching its product. As a result, it did not benefit from the one-stop-shop mechanism until February 2024 (EDPB, 2024).

Lawfulness. A key provision of the GDPR mandates that the processing of personal data requires a legal basis.[3] The legal basis that most closely aligns with the idea of privacy self-management would be obtaining the user's consent (Schaub et al., 2015). However, in the context of Generative AI chatbots, almost all companies currently rely on legitimate interest. According to recent case law from the European Court of Justice, product improvement may indeed be considered a necessary legitimate interest.[4]

Legitimate interest. Data Protection Authorities are expected to challenge the necessity of such processing for model improvement purposes (EDPB, 2024). While using real user data might enhance alignment, less intrusive alternatives, such as processing data from users who consent, could suffice (Sartor, 2023). Since companies claim that only a small portion of data is used, making processing conditional on consent might still yield sufficient data. However, suggesting reliance solely on internal alignment trainers or external contractors might be overly restrictive as it might not accurately reflect real user data.

Furthermore, it is required that the legitimate interests outweigh those of its users. This balancing depends on three main factors. First, the reasonable expectations of the affected users (Veil, 2018). Since interacting with Generative AI chatbots is quite a new phenomenon, many users might not anticipate how their input is used beyond the conversation. For instance, users likely do not realize the implications when they provide a thumbs up on a response. Second, legitimate interest might not be a sufficient legal basis, when the user data involves certain sensitive conversations, such as when a user discusses personal health issues.[5] Third, the impact of the processing and whether the company implements specific safeguards to protect personal data, such as anonymization, pseudonymization, and data minimization.

Opt-out. In addition, the legal basis of legitimate interest requires that users be informed and given the option to opt out, with non-compliance subject to fines the GDPR.[6] Users must be informed at the time of the first communication (Forgó, 2024). Making it available as part of the privacy policy, might meet the legal minimum requirements if it is clearly highlighted (Carmichael et al., 2023). However, it is problematic to impose costs or hide or complicate the opt-out option to discourage users.[7]

3.2. The AI Act

Transparency. While the GDPR provides a framework for privacy, the AI Act introduces specific provisions for AI systems. The new provisions for foundation models under the EU's AI Act are expected to take effect on 2 August 2025. In Chapter 5, the EU has introduced new rules for General Purpose AI models (GPAIs) that will apply alongside the GDPR rules mentioned earlier.[8] It is likely that most of the models behind the discussed Generative AI chatbots will be classified as a *Systemic risk* GPAI because the floating-point operations for training exceed 10^{25} (Epoch AI, 2024) or due to other criteria such as a high number of users.[9] Despite this, the AI Act does not require opting in for the use of data for model improvement. However, it imposes documentation duties on GPAIs vis-à-vis the Office for AI and providers of AI systems regarding the training of their models (Friedl and Gasiola, 2024).[10] These requirements may increase overall transparency.

3.3. Legal Benchmarking

In the Generative AI community, it is common to compare the performance of an LLM on leaderboards. However, this practice has increasingly been questioned due to benchmark overfitting (Alzahrani et al., 2024) and benchmark leakage (Zhang et al., 2024). Beyond performance metrics, it is also important to consider privacy aspects. To assess how well companies are adhering to these legal requirements, we have developed a benchmarking framework based on key principles derived from the GDPR.

In Table 1, we compare the adherence to the legal principles discussed earlier. We focus on the opt-out option (*opt-out*) and whether it is explicitly communicated in the first use of the chatbot (*first use*). Merely including the option in the privacy policy (*fine print*) is not ideal as research has

[2] Article 3(1) and Article 3(2) GDPR. For improved readability, references to laws will be found in footnotes.

[3] Article 6(1) GDPR.

[4] Meta Platforms Inc. and Others v. Bundeskartellamt, Case C-252/21, ECLI:EU:C:2023:537, para. 122 (2023) ('it cannot be ruled out from the outset that the controller's interest in improving the product or service with a view to making it more efficient and thus more attractive can constitute a legitimate interest capable of justifying the processing of personal data and that such processing may be necessary in order to pursue that interest').

[5] Article 9(1) GDPR.

[6] Article 21 GDPR; Article 85(5)(b) GDPR.

[7] Article 12 GDPR.

[8] Recital 10 of the AI Act.

[9] Article 51(2) and 51(1)(b) AI Act and its Annex XIII.

[10] Article 53 and 55 AI Act.

shown that consumers rarely read or understand these documents (Bakos et al., 2014). Users should also be able to opt out without unnecessary friction in the settings (*settings*). Additionally, users should have access to clear documentation about privacy safeguards (*privacy safeguards*) and the retention limitation (*retention*).

Table 1. Legal comparison of Generative AI chatbots based on privacy criteria. ●: likely compliant, ◑: partial compliance, ○: potential non-compliance. Note: Claude is excluded as it does not use conversational data by default. (July 2024)

	CHATGPT	GEMINI	LECHAT
LEGAL BASIS	◑	◑	◑
OPT-OUT	●	●	◑
First use	◑	●	○
Settings	●	◑	○
Fine print	●	●	●
PRIVACY SAFEGUARDS	●	◑	○
RETENTION	○	●	○

Surprisingly, Mistral AI, often praised as a European champion, does not seem to champion privacy. Currently, regulatory oversight is heavily focused on OpenAI and Google. However, if EU laws are not merely industrial policy tools, what the US perception is sometimes claimed, and if Mistral AI does not change its policies for LeChat, they too will likely become a target for regulatory authorities.

4. Discussion

The two hats. The challenge of Generative AI chatbot providers may stem from two conflicting roles: that of model developer and a chatbot provider. A model developer has the incentive to collect as much user data as possible to optimize the training and improve the model. As a chatbot provider, one might be incentivized to increase user satisfaction through user- and privacy-friendly chat interfaces. Therefore, it is not surprising that HuggingFace offers HuggingChat (HuggingFace), which provides users with a chat interface to use different open-source models. In this case, the data is not used for further training or forwarded to the model makers. The interest situation is different here, as the chat provider is not the model developer. The dual role as a chat provider and model developer presents a challenge.

Consumer vulnerabilities. It has become an industry standard that model developers do not use the data of companies that send data through APIs for business applications. However, professional business users can handle much more complexity than regular consumers, who are more vulnerable due to power imbalances and knowledge gaps (Helberger et al., 2022). Although there is a business rationale for providing better conditions for companies, having weaker default privacy setting for consumers is not reasonable from a vulnerability standpoint. This is especially true as Generative AI chatbots become a household phenomenon, but might differ during an early phase when only early tech adopters use them.

Law and technology. In the law and tech literature, the so-called *pacing* problem describes how the law often lags behind technological advancements (Marchant, 2011). Indeed, early ChatGPT users could not opt out from further training. By the time the Italian Garante intervened, GPT-4 had already been launched, likely having been trained on new user data. Another factor contributing to the *pacing* problem is the phenomenon known as the *Collingridge dilemma* (Collingridge, 1982), which highlights the difficulty regulators face in finding the ideal time to intervene with new technologies—not too early and not too late. Finding this sweet spot is challenging and often results in either regulatory laissez-faire or regulatory overreach.

However, the *pacing* problem is only one part of the story, there is also a *racing* problem. ChatGPT's rapid global launch illustrates Silicon Valley's philosophy of 'move fast and break things' (Taplin, 2017). As the prominent example of Uber shows (The Guardian, 2022), some companies blatantly exploit the *pacing* problem by *racing* to release their products, even amidst legal uncertainties.

Due to its reactive nature, legal frameworks can thus reward those who are bold and launch things prematurely. With Generative AI, data collection becomes increasingly challenging as regulatory oversight intensified and resistance grows (Metz, 2024). However, a first mover can particularly benefit in the early stages while it flies under the radar. A first mover can also leverage network effects, quickly amass a large user base and create a cycle of improvement and user growth. If fine-tuning is the secret sauce behind the success of Generative AI chatbots, the run for conversation data is intense. It is questionable whether Data Protection Authorities can effectively impose sanctions that impact the market advantage gained by first movers.

Innovation. Excessive legal intervention can hinder or even prevent innovation. Not without reason have existing AI companies been very secretive about the training dataset for pre-training, although this seems to be slowly changing (Liu et al., 2023). This strategy may be understandable, as the massive data collection that enables the technology could be threatened by legal concerns. However, collecting conversational data for fine-tuning is a different legal issue that does not prevent the technology as a whole. There are various ways conversational data can be used: with user consent or by default, with a clear opt-out option.

When Meta had to pause its plans due to regulatory pressure, it called it a 'step backwards for European innovation' and announced it would not launch Meta AI in Europe (Meta, 2024). This reaction mirrors earlier debates on other Internet phenomena, such as the regulation of spam around the turn of the millennium. EU law requires consent before one can send marketing emails, while the US follows an opt-out system. Before the EU introduced this, there were concerns that it could 'kill e-commerce in Europe' (Magee, 2003), but these fears did not materialize.

Meta's response to regulatory challenges is rigid. However, alternative approaches to default data collection and disincentivized opt-out procedures exist. For example, companies could incentivize voluntary opt-in. Users who contribute to fine-tuning could be rewarded with perks such as early access to improved models. Additionally, session-based sharing offers a nuanced approach, allowing users to share data selectively based on context—such as consenting to share a chat about a recipe but not a personal medical analysis. Many companies, including the recent case with X (Fingas, 2024), choose to disincentivize voluntary opt-out instead of encouraging voluntary opt-in.

Friction. Disincentivizing users from exercising their opt-out rights through deceptive design is not a new phenomenon online (Acquisti et al., 2015). This practice, often referred to as *dark patterns*, is well-studied in the context of cookie consent (Luguri and Strahilevitz, 2021). A newer study by (Kyi et al., 2023) also found a lack of transparency and deceptive design patterns in user interfaces when companies rely on legitimate interest. For Generative AI chatbots, friction could be introduced by complicating the opt-out process or by degrading performance for non-technical reasons after opting out. Empirical research has shown early on that defaults significantly influence behavior (Johnson et al., 2002), suggesting that these tactics might be intentional. The legal qualification of this issue will be determined on a case-by-case basis. However, in a regulatory investigation, a company would need to provide truthful information.

5. Conclusion

LLMs are becoming widely used and multimodal (Stanford University, 2024). With this widespread adoption comes the unprecedented generation of conversational data. Although the pre-training and copyright law receive much scholarly attention, the issue of using conversational data to improve models raises issues regarding privacy. In this study, we survey current EU legal requirements governing privacy, and present a benchmark that highlights potential legal shortcomings. We further situate our findings within the larger framework of law and technology.

Acknowledgements

Special thanks to Selina Ehrenzeller, Florian Geering, Aniket Kesari, Dominik Stammbach, Luca Strässle, Vandit Sharma, Jingwei Ni and the anonymous reviewers for their feedback and beneficial discussions. The author also acknowledges the use of GPT-4o for assistance with proofreading. All URLs referenced in this paper were accessed prior to July 31, 2024.

References

Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Privacy and Human Behavior in the Age of Information. *Science*, 347(6221):509–514, 2015.

Norah Alzahrani, Hisham Abdullah Alyahya, Yazeed Alnumay, Sultan Alrashed, Shaykhah Alsubaie, Yusef Almushaykeh, Faisal Mirza, Nouf Alotaibi, Nora Altwairesh, Areeb Alowisheq, M Saiful Bari, and Haidar Khan. When Benchmarks Are Targets: Revealing the Sensitivity of Large Language Model Leaderboards. 2024. URL https://arxiv.org/abs/2402.01781.

Yannis Bakos, Florencia Marotta-Wurgler, and David R. Trossen. Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts. *The Journal of Legal Studies*, 43(1):1–35, 2014.

Ashley Belanger. Slack Defends Default Opt-in for AI Training on Chats Amid User Outrage. *Ars Technica*, 2024. URL https://arstechnica.com/tech-policy/2024/05/slack-defends-default-opt-in-for-ai-training-on-chats-amid-user-outrage/.

Jaydeep Borkar. What Can We Learn from Data Leakage and Unlearning for Law? 2023. URL https://arxiv.org/abs/2307.10476.

Laura Carmichael, Emma Cradock, and Sophie Stalla-Bourdillon. Article 6. In Indra Spiecker gen. Döhmann and Vagelis Papakonstantinou and Gerrit Hornung and Paul De Hert, editor, *General Data Protection Regulation: Article-by-Article Commentary*, page 524. 2023.

Pier Giorgio Chiara. Italian DPA v. OpenAI's ChatGPT: The Reasons Behind the Investigations and the Temporary Limitation to Processing. *Journal of Law and Technology*, 31, 2023.

David Collingridge. The Social Control of Technology. 1982.

Data Protection Commission. Annual Report 2023. URL https://www.dataprotection.ie/sites/default/files/uploads/2024-05/DPC%20EN_AR%202023_Final%20.pdf.

Amon Dieker. Datenschutzrechtliche Zulässigkeit der Trainingsdatensammlung. *Zeitschrift für Datenschutz*, pages 132–137, 2024.

EDPB. Report of the Work Undertaken by the ChatGPT Taskforce, 2024. URL https://www.edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-chatgpt-taskforce_en.

Epoch AI. Epoch AI Database, 2024. URL https://epochai.org/data/epochdb/visualization?startDlEra=1950-12-28.

Jon Fingas. X is Training Grok AI on Your Data—Here's How to Stop It. *Ars Technica*, 2024. URL https://arstechnica.com/ai/2024/07/x-is-training-grok-ai-on-your-data-heres-how-to-stop-it/.

Nikolaus Forgó. Article 21. In Heinrich Amadeus Wolff and Stefan Brink and Antje v. Ungern-Sternberg, editor, *BeckOK Datenschutzrecht: DS-GVO, DA, DGA, BDSG. Datenschutz und Datennutzung*, page 318. 48th edition, 2024.

Paul Friedl and Gustavo Gil Gasiola. Examining the EU's Artificial Intelligence Act. *Verfassungsblog*, 2024. URL https://verfassungsblog.de/examining-the-eus-artificial-intelligence-act/.

Garante. Temporary Limitation of Processing of Italian Users' Data Against OpenAI, 2023. URL https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english.

Garante. ChatGPT: Italian DPA Notifies Breaches of Privacy Law to OpenAI, 2024. URL https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9978020#english.

Joanna Glasner. Startup Valuation Fluctuations. *Crunchbase*, 2023. URL https://news.crunchbase.com/venture/startup-valuation-fluctuations-ai-openai-msft-eoy-2023/.

Natali Helberger, Martin Sax, Joanna Strycharz, and Hans-W. Micklitz. Choice Architectures in the Digital Economy: Towards a New Understanding of Digital Vulnerability. *Journal of Consumer Policy*, 45:175–200, 2022.

HuggingFace. HuggingChat. URL https://huggingface.co/chat.

Eric J. Johnson, Steven Bellman, and Gerald L. Lohse. Defaults, Framing and Privacy: Why Opting In-Opting Out. *Marketing Letters*, 13:5–15, 2002.

Lin Kyi, Sushil Ammanaghatta Shivakumar, Cristiana Teixeira Santos, Franziska Roesner, Frederike Zufall, and Asia J. Biega. Investigating Deceptive Design in GDPR's Legitimate Interest. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, 2023. URL https://doi.org/10.1145/3544548.3580637.

Zhengzhong Liu, Aurick Qiao, Willie Neiswanger, Hongyi Wang, Bowen Tan, Tianhua Tao, Junbo Li, Yuqi Wang, Suqi Sun, Omkar Pangarkar, Richard Fan, Yi Gu, Victor Miller, Yonghao Zhuang, Guowei He, Haonan Li, Fajri Koto, Liping Tang, Nikhil Ranjan, Zhiqiang Shen, Xuguang Ren, Roberto Iriondo, Cun Mu, Zhiting Hu, Mark Schulze, Preslav Nakov, Tim Baldwin, and Eric Xing. LLM360: Towards Fully Transparent Open-Source LLMs. 2023. URL https://arxiv.org/abs/2312.06550.

Jamie Luguri and Lior Jacob Strahilevitz. Shining a Light on Dark Patterns. *Journal of Legal Analysis*, 13(1):43–109, 2021.

John Magee. The Law Regulating Unsolicited Commercial E-Mail: An International Perspective. *Santa Clara Computer & High Technology Law Journal*, 19:333, 2003.

Gary E. Marchant. *The Growing Gap Between Emerging Technologies and the Law*, pages 19–33. 2011.

Meta. Building AI Technology for Europeans in a Transparent and Responsible Way, 2024. URL https://about.fb.com/news/2024/06/building-ai-technology-for-europeans-in-a-transparent-and-responsible-way/.

Cade Metz. AI Data Restrictions: How New Regulations Are Shaping the Future of Technology. *The New York Times*, 2024. URL https://www.nytimes.com/2024/07/19/technology/ai-data-restrictions.html?smid=nytcore-android-share.

Mistral AI Team. Le Chat, 2024. URL https://mistral.ai/news/le-chat-mistral/.

Milad Nasr, Nicholas Carlini, Jonathan Hayase, Matthew Jagielski, A. Feder Cooper, Daphne Ippolito, Christopher A. Choquette-Choo, Eric Wallace, Florian Tramèr, and Katherine Lee. Scalable Extraction of Training Data from (Production) Language Models. 2023. URL https://arxiv.org/abs/2311.17035.

OpenAI. New Ways to Manage Your Data in ChatGPT. 2023. URL https://openai.com/index/new-ways-to-manage-your-data-in-chatgpt/.

OpenAI. Start Using ChatGPT Instantly, 2024a. URL https://openai.com/index/start-using-chatgpt-instantly/.

OpenAI. Memory and New Controls for ChatGPT, 2024b. URL https://openai.com/index/memory-and-new-controls-for-chatgpt/.

OpenAI Community. Chat History Off - Conversation Key Not Found Error, 2024. URL https://community.openai.com/t/chat-history-off-conversation-key-not-found-error/594342.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F Christiano, Jan Leike, and Ryan Lowe. Training Language Models to Follow Instructions with Human Feedback. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 27730–27744, 2022.

Matthew Sag. Copyright Safety for Generative AI. *Houston Law Review*, 61:295, 2023.

Pamela Samuelson. Generative AI Meets Copyright. *Science*, 381(6654):158–161, 2023.

Giovanni Sartor. Article 6. In Indra Spiecker gen. Döhmann and Vagelis Papakonstantinou and Gerrit Hornung and Paul De Hert, editor, *General Data Protection Regulation: Article-by-Article Commentary*, page 318. 2023.

Florian Schaub, Rebecca Balebako, Adam L. Durity, and Lorrie Faith Cranor. A Design Space for Effective Privacy Notices. In *Eleventh Symposium on Usable Privacy and Security (SOUPS 2015)*, pages 1–17, 2015.

Seth Shafer. Consumer Insights: One-quarter of US Adults Have Used Generative AI Tools. *S&P Global Market Intelligence*, 2024.

Stanford University. The AI Index Report, 2024. URL https://aiindex.stanford.edu/report/.

Jonathan Taplin. *Move Fast and Break Things: How Facebook, Google, and Amazon Have Cornered Culture and What It Means for All of Us*. 2017.

The Guardian. Uber Files, 2022. URL https://www.theguardian.com/news/series/uber-files.

Aaron Tilley and Miles Kruppa. Apple Restricts Use of ChatGPT, Joining Other Companies Wary of Leaks. *The Wall Street Journal*, 2023. URL https://www.wsj.com/articles/apple-restricts-use-of-chatgpt-joining-other-companies-wary-of-leaks-d44d7d34?

Winfried Veil. Einwilligung oder berechtigtes Interesse?: Datenverarbeitung zwischen Skylla und Charybdis. *Neue Juristische Wochenschrift*, pages 3337–3344, 2018.

Tianyu Wu, Shizhu He, Jingping Liu, Siqi Sun, Kang Liu, Qing-Long Han, and Yang Tang. A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development. *IEEE/CAA Journal of Automatica Sinica*, 10(5):1122–1136, 2023.

Hugh Zhang, Jeff Da, Dean Lee, Vaughn Robinson, Catherine Wu, Will Song, Tiffany Zhao, Pranav Raja, Dylan Slack, Qin Lyu, Sean Hendryx, Russell Kaplan, Michele Lunati, and Summer Yue. A Careful Examination of Large Language Model Performance on Grade School Arithmetic, 2024. URL https://arxiv.org/abs/2405.00332.

A. Privacy Policies and Terms of Service

Excerpts from the four Generative AI chatbot companies are provided. Key points are highlighted in bold for emphasis.

Privacy Policy Excerpt: OpenAI (June 23, 2023)

[[3] How we use Personal Data We may use Personal Data for the following purposes: • To provide and maintain our Services; • **To improve and develop our Services and new features and conduct research**; • To communicate with you, including to send you information or marketing about our Services and events; • To prevent fraud, criminal activity, or misuses of our Services, and to protect the security of our systems and Services; and • To comply with legal obligations and to protect the rights, privacy, safety, or property of our users, us, our affiliates, or any third party. **Aggregated or De-Identified Information. We aggregate or de-identify Personal Data** so that it can no longer be used to identify you and use this information to analyze the effectiveness of our Services, to improve and add features to our Services, to conduct research and for other similar purposes. In addition, from time to time, we may share or publish aggregated information like general user statistics with third parties. We collect this information through the Services, through cookies, and through other means described in this Privacy Policy. We will maintain and use de-identified information in anonymous or de-identified form and **we will not attempt to re-identify** the information, unless required by law. As noted above, **we use Content you provide us to improve our Services, for example to train the models that power our Services**. Read our instructions(opens in a new window) on **how you can opt out of our use of your Content to train our models**.

[6] Your rights You have the following statutory rights in relation to your Personal Data: • Access your Personal Data and information relating to how it is processed. • Delete your Personal Data from our records. • Rectify or update your Personal Data. • Transfer your Personal Data to a third party (right to data portability). • Restrict how we process your Personal Data. • Withdraw your consent—where we rely on consent as the legal basis for processing at any time. • Lodge a complaint with your local data protection authority (see below). You have the following rights to object: • Object to our processing of your Personal Data for direct marketing at any time. • **Object to how we process your Personal Data when our processing is based on our legitimate interests**. You can exercise some of these rights through your OpenAI account. If you are unable to exercise your rights through your account, please submit your request through privacy.openai.com(opens in a new window) or send it to dsar@openai.com.

[8] Legal bases for processing When we process your Personal Data for the purposes described above, we rely on the following legal bases: To improve and develop our Services and new features and conduct research • Account Information • User Content • Communication Information • Other Information You Provide • Data We Receive From Other Sources • Log Data • Usage Data • Device Information • Cookies and Similar Technologies • See here for more specific information on the data used to train our models. Where necessary for our **legitimate interests** and those of third parties and broader society, including in **developing, improving, or promoting our Services, such as when we train and improve our models**. See here(opens in a new window) for more information.]

Terms of Service Excerpt: OpenAI (March 14, 2023)

[[2] Content Your content. You may provide input to the Services ("Input"), and receive output from the Services based on the Input ("Output"). Input and Output are collectively "Content". You are responsible for Content, including ensuring that it does not violate any applicable law or these Terms. You represent and warrant that you have all rights, licences, and permissions needed to provide Input to our Services. Ownership of content. As between you and OpenAI, and to the extent permitted by applicable law, you (a) retain your ownership rights in Input and (b) own the Output. We hereby assign to you all our right, title, and interest, if any, in and to Output. Similarity of content. Due to the nature of our Services and artificial intelligence generally, Output may not be unique and other users may receive similar output from our Services. Our assignment above does not extend to other users' output or any Third Party Output. Our use of content. We can use your Content worldwide to provide, maintain, develop, and improve our Services, comply with applicable law, enforce our terms and policies and keep our Services safe. **Opt out**. If you do not want us to use your Content to train our models, you have **the option to opt out by updating your account settings**. Further information can be found in this Help Center article(opens in a new window). Please note that in some cases this may limit the ability of our Services to better address your specific use case.]

Privacy Policy Excerpt: Google (March 28, 2024)

[We use data to build better services We use the information we collect from all our services for the following purposes: Maintain & improve our services We also use your information to ensure our services are working as intended , such as tracking outages or troubleshooting issues that you report to us. And we use your information to **make improvements to our services** — for example, understanding which search terms are most frequently misspelled helps us improve spell-check features used across our services.

Legal bases of processing We process your information for the purposes described in this policy, based on the following legal grounds: When we're pursuing legitimate interests We process your information for our **legitimate interests** and those of third parties while applying appropriate safeguards that protect your privacy. This means that we process your information for things like: **Providing, maintaining, and improving our services** to meet the needs of our users **Developing new products and features** that are useful for our users **Understanding how people use our services** to ensure and improve the performance of our services Customizing our services to provide you with a better user experience (and, if relevant, tailoring the experience to be age-appropriate) Marketing to inform users about our services Providing advertising, which allows us to offer many of our services without a fee(and when ads are personalized, we ask for your consent) Detecting, preventing, or otherwise addressing fraud, abuse, security, or technical issues with our services Protecting against harm to the rights, property or safety of Google, our users, or the public as required or permitted by law, including disclosing information to government authorities Performing research that improves our services for our users and benefits the public Fulfilling obligations to our partners like developers and rights holders Enforcing legal claims, including investigation of potential violations of applicable Terms of Service]

Terms of Service Excerpt: Google (May 29, 2024)

[**Google Gemini Apps Privacy Notice** What data is collected and how it's used Google collects your Gemini Apps conversations, related product usage information, info about your location, and your feedback. Google uses this data, consistent with our Privacy Policy , to provide,improve, and develop Google products and services and machine-learning technologies, including Google's enterprise products such as Google Cloud. **Gemini Apps Activity is on by default if you are 18 or older**. Users under 18 can choose to turn it on. How human reviewers improve Google AI To help with quality and improve our products (such as generative machine-learning models that power Gemini Apps), **human reviewers read, annotate, and process your Gemini Apps conversations**.We take steps to **protect your privacy** as part of this process. This includes **disconnecting your conversations** with Gemini Apps from your Google Account before reviewers see or annotate them. **Please don't enter confidential information in your conversations or any data you wouldn't want a reviewer to see or Google to use** to improve our products, services, and machine-learning technologies.

Configuring your settings Visit your Google Account to access settings and tools that let you safeguard your data and protect your privacy. **To stop future conversations from being reviewed or used** to improve Google machine-learning technologies, **turn off Gemini Apps Activity** . You can review your prompts or delete your conversations from your Gemini Apps Activity at myactivity.google.com/product/gemini . **Conversations that have been reviewed or annotated by human reviewers (and related data like your language, device type, location info, or feedback) are not deleted when you delete your Gemini Apps activity** because they are kept separately and are not connected to your Google Account. Instead,they **are retained for up to three years**. Even when Gemini

Apps Activity is off, your conversations will be saved with your account for up to 72hours. This lets Google provide the service and process any feedback. This activity won't appear in your Gemini Apps Activity. Learn more .

How can I object to the processing of my data or ask for inaccurate data in Gemini Apps'responses to be corrected? LLM experiences (Gemini Apps included) can hallucinate and present inaccurate information as factual. Under certain privacy laws, including the General Data Protection Regulation in the EU, you may have the right to:1. **Object to the processing of your personal data**, or 2.Ask for inaccurate personal data in Gemini Apps' responses to be corrected. To exercise these rights, you can create a request in our Help Center.

What are Google's legal bases of processing Gemini Apps data under European Union (EU) or United Kingdom (UK) data protection law? If European Union (EU) or United Kingdom (UK) data protection law applies to the processing of your information, please read the below carefully. When you use Gemini Apps, Google processes your information for the purposes, and on the legal grounds, described below. When we refer to "your Gemini Apps information" below we mean all of the following: (i) your Gemini Apps conversations, related product usage information (which includes your location information), and any supplemental information processed when Gemini is your mobile assistant;and (ii) your feedback. Google's legal bases are: Performance of a contract. We process your Gemini Apps information, and any of your other information you give Gemini Apps permission to process when you integrate Gemini Apps with another service, so that we can provide and maintain the Gemini Apps service you have requested under the Google Terms of Service . For example, we process your Gemini Apps information to respond to your queries, and to provide various Gemini Apps features and functionalities such as code generation. **Google and third parties' legitimate interests with appropriate safeguards** to protect your privacy. We process information from publicly accessible sources and your Gemini Apps information so that we can provide, maintain, improve, and develop Google products, services, and machine learning technologies. Processing this information for this purpose is necessary for the legitimate interests of Google and our users in: **Providing, maintaining, and improving our services** to meet the needs of our users (such as **using conversations to fine-tune models** and improve Gemini Apps' responses for safety and accuracy). Developing new products and features that are useful for our users (such as learning how to route requests to new large language models best suited to answer a particular question or training new models to handle these requests). **Understanding how people use our services** to ensure and improve the performance of our services (such as generating metrics to understand how users are using Gemini Apps to better tailor user experience). Customizing our services to provide users with a better experience (such as using your location information and your past conversations so Gemini Apps provide a more relevant answer). We also process your Gemini Apps information so that we can maintain the functionality,safety and reliability of Gemini Apps, including by detecting, preventing, and responding to fraud, abuse, security risks, and technical issues that could affect Google, our users, or the public. Processing this information for this purpose is necessary for the legitimate interests of Google, our users, and the public in: Detecting, preventing, or otherwise addressing fraud, abuse, security, or technical issues with our services (such as fixing bugs and troubleshooting failures). Protecting against harm to the rights, property, or safety of Google, our users, or the public as required or permitted by law (such as updating safety classifiers and model filters). Performing research that improves our services for our users and benefits the public. Enforcing legal claims, including investigation of potential violations of applicable terms of service (such as reviewing suspicious activity and interactions flagged as problematic). Processing this information for this purpose is also necessary for the legitimate interests of Google and our commercial partners in fulfilling obligations to our partners like developers and rights holders (such as honoring removal requests from intellectual property rightsholders). To respond to your request, for example to summarize an email from a named contact, we process personal data about others that you've given Gemini Apps (such as when you integrate Gemini Apps with another service like Google Workspace). Legal obligations. We'll also process your Gemini Apps information to meet any applicable law,regulation, legal process, or enforceable governmental request (such as if we get a legal request for information from a governmental authority). Your consent. We rely on your consent when you use certain features such as Voice Match . As Gemini Apps develop, we may ask for your consent to process your information for specific purposes. Where we rely on your consent to process information, you will have the right to withdraw your consent at any time. What data is collected When you interact with Gemini Apps, Google collects your: **Conversations** Location Feedback Usage information About your Gemini Apps Activity control When Gemini Apps Activity is on, Google stores your Gemini Apps activity (such as your prompts, responses, and feedback) in your Google Account. Even when Gemini Apps Activity is off, your conversations will be saved with your account for up to 72 hours to allow us to provide the service and process any feedback. This activity will not show up in your Gemini Apps Activity.

How **human review helps improve models Google uses conversations (as well as feedback and related data) from Gemini Apps users to improve Google products (such as the generative machine-learning models that power GeminiApps), so we can make them safer, more helpful, and work better for all users. Human review is a necessary step of the model improvement process.** Through their review, rating, and rewrites,humans help enable quality improvements of generative machine-learning models like the ones that power Gemini Apps. How we protect your privacy in this process We take a number of **precautions to protect your privacy during this human review process: Conversations (as well as feedback and related data like your language, device type, or location info) that reviewers see and annotate are not associated with any user accounts. We pick a random sample for such human review, and only a portion of all Gemini Apps conversations are reviewed**. How the process works **Our trained reviewers look at conversations to assess if Gemini Apps' responses are low-quality, inaccurate, or harmful. From there, trained evaluators suggest higher-quality responses. These are then used to create a better dataset for generative machine-learning models to learn from so our models can produce improved responses in the future**. How long is reviewed data retained Gemini Apps conversations that have been reviewed by human reviewers (as well as feedback and related data like your language, device type, or location info) are not deleted when you delete your Gemini Apps activity because they are kept separately and are not connected to your Google Account. Instead, they are **retained for up to 3 years**. **How you can control what's shared with reviewers If you turn off Gemini Apps Activity , future conversations won't be sent for human review or used to improve our generative machine-learning models**. **Don't enter anything you wouldn't want a human reviewer to see or Google to use**. For example, don't enter info you consider confidential or data you don't want to be used to improve Google products, services, and machine-learning technologies. Learn more about how to turn off, manage, and delete your Gemini Apps activity at any time. Also,you can control Google's storage of Gemini Apps activity in My Activity at any time.]

Privacy Policy Excerpt: Mistral AI (June 26, 2024)

[[1] Definitions The capitalized words in this document will have the meaning given below: • "Data Controller": means the person who makes decisions on Your Personal Data. For instance, the Data Controller decides which Personal Data to collect, where to store such data, for how long, etc. • "Data Processor": means the person who uses Your Personal Data on behalf of the Data Controller and under the Data Controller's instructions. For instance, our hosting services provider acts as Data Processor when it stores Your Personal Data on Our behalf and under Our instructions. • **"Incognito Mode": means the feature that allows You to not display your Prompts and Outputs History when using Our Chat Services**.

[3] What kind of Personal Data do We collect ? 3.1. Personal Data You provide to Us • Identity, account and contact data. We collect Personal Data about Your identity when You sign-up to Our Services (first name, last name, email address, identifiers, etc.), when You subscribe to receive information from Us (newsletters, etc.) and/or when You contact Us. • Payment and billing information. We collect Your payment details (payment method, applicable fees, etc.) and your billing information (billing address, billing contact, etc.) when You subscribe to Our fee-based Services. • Prompts and Outputs. We only use Your Prompts and/or Outputs: o To monitor abuse, meaning any breach by You of the Terms of Use or the applicable Terms of Service, or o When You report an illicit Output. In such a case, we may use Your Prompt and/or Your Outputs to improve our services. o To display your Prompts and Outputs history on Your Account when You use Our Chat Services, unless You activate Our Incognito Mode. o **To improve Our Models. We may use your Prompts or Outputs or information (e.g. language or topic field statistics) related to them to improve Our Models only when You Use Our Chat Services and You don't opt-out of the Mistral AI Training Data, when such option is available for the Paid Chat Services You use the free version of Codestral and You don't opt-out of the Mistral AI Training Data. You can opt-out of the Mistral AI Training Data at any time by making a request directly on Our Support Chatbot available on Our Platform. Please note that in that case, Your opt-out will only be effective for future Prompts and Outputs. • User Input Data (for Customers acting as Consumers). We only use Your User Input Data to allow You to Fine-Tune Models via Our Fine-Tuning API.** • Feedback. Where applicable, We may collect and use the Feedback You provide to Us to improve our Service.

[4] . Why do we use Your Personal Data? We use Your Personal Data for the following purposes: Provide Our Services • Create and administer Your account on the Platform. • Manage the security of the Services, for Customers acting as Consumers. • Generate Outputs based on Your Prompts. • Allow You to Fine-Tune

Our Models using Our Fine-Tuning API, for Customers acting as Consumers. • Communicate with You for purposes other than marketing. • Answer to your assistance requests. • When You use Our Chat Services as a Consumer, provide technical support (fixing the bugs You notify us). • Display Your Prompts and Outputs to You. • **Train Our Models (Large Language Models) to answer questions, generate text according to context/Prompts (e-mails, letters, reports, computer code, etc), translating, summarizing and correcting text, classifying text, analyzing feelings, etc.,** (i) Possibly, if such Personal Data is publicly available, and even if we apply good practices to filter such data, and (ii) **if You use Our Chat Services and You don't opt-out of the Mistral AI Training Data when such option is available for the Paid Chat Services (iii) if You use the free version of Codestral.** • Make aggregated statistics about the use of the Services. Legal basis: Performance of the contract. **Our legitimate interest in (1) providing quality Services and continuously improving our Services and (2) developing Our Models for the purpose of providing them to You.**

[5] How long do We store Your Personal Data ? We may keep Your Personal Data for as long as necessary to achieve the purposes mentioned in Section 4 of this Privacy Policy. We may retain your Personal Data for longer periods when We are required by applicable law to do so or when it is necessary to exercise Our rights in legal proceedings. For illustrative purposes, please find below the applicable data retention periods: Personal Data We use to Provide the Services: • Identity and contract/subscription data : for the duration of your registration on the Platform and for 5 years from the end - of your registration for evidentiary purposes. • Account data : for the duration of your registration on the Platform and for 1 year from the end of your registration for - evidentiary purposes. • Security data: the security logs are stored for 1 rolling year. • User Input Data (for Customers Acting as Consumers): for the duration of your registration on the Platform and/or until You delete such data. • Technical support/assistance requests: for the duration of the Processing the request and for 5 years from the processing of - Your request for evidentiary purposes.]

Terms of Service Excerpt: Mistral AI (June 26, 2024)

[Your Prompts and Outputs. Your Prompts and Outputs history will be stored by Us in order to be displayed to You as Your Le Chat history. By entering into this Agreement, **You authorize Mistral AI to use Your Prompts and Outputs to improve its Models and the Chat Services**. You acknowledge that: • If You wish to exclude Personal Data (in a clear, identifiable form) from the Mistral AI Training Data, where high-level information about your Prompts and Outputs may be used, **You must pseudonymize Your Prompts, • Your must not include sensitive Personal Data in Your Prompts, • You must not include sensitive business data such as, for instance, trade-secrets, know-how, etc, in Your Prompts.** If You decide to **subscribe to the Paid Chat Services, You will have the option to: • Activate the incognito mode ("Incognito Mode"). If You activate the Incognito Mode, Your Prompts and Outputs will not be displayed in Your Le Chat history; and • Opt-out of the Mistral AI Training Data at any time by activating the relevant option on Your Account. In such a case, Mistral AI will not use Your Prompts and Outputs to get high level information (e.g. language or field of the prompts) to improve future training of its Models.** Please note that Your opt-out may only be effective for future Prompts and Outputs, as technical limitations inherent to the Chat Services may prevent Mistral AI from deleting all information previously extracted from Your Prompts and Outputs from Mistral AI Training Data, even if you have opted out. Mistral AI may also use Your Prompts and Outputs to monitor abuse as set out in Section (Your User Data) of the Terms of Use.

How we use Your User Data. We only use Your User Data: • To provide the Services (e.g. to generate outputs based on Your Prompts, to display Your Conversation history, etc.), • To monitor abuse (meaning, to monitor any breach by You of the Agreement). To this end, we retain Your Prompts and Outputs for a period of thirty (30) days, • **To improve Our Models. We may use your Prompts or Outputs or information (e.g. language or topic field statistics) related to them to improve Our Models only when (a) you Use Our Chat Services and (b) You don't opt-out of the Mistral AI Training Data, when such option is for the Paid Chat Services**.]

Privacy Policy Excerpt: Anthropic (June 5, 2024)

[Personal data you provide to us directly Identity and Contact Data: Anthropic collects identifiers, such as your name, email address, and phone number when you sign up for an Anthropic account, or to receive information on our Services. We may also collect or generate indirect identifiers (e.g., "USER12345"). Payment Information: We may collect your payment information if you choose to purchase access to Anthropic's products and services. Inputs and Outputs: Our AI services allow you to prompt the Services in a variety of media including but not limited to the format of text,files and documents, photos and images, and other materials along with the metadata and other information contained therein ("Prompts" or "Inputs") , which generate responses("Outputs" or "Completions") based on your Inputs. If you include personal data in your Inputs, we will collect that information and this information may be reproduced in your Outputs.

Feedback on your use of our Services: We appreciate feedback, including ideas and suggestions for improvement or rating an Output in response to an Input (" Feedback "). **If you rate an Output in response to an Input—for example, by using the thumbs up/thumbs down icon—we will store the related conversation as part of your Feedback**. Applicable Data Protection Laws We will only use your personal data in accordance with applicable laws. We rely on the following grounds where permitted under and in accordance with data protection laws, such as in the European Union (our " Legal Bases "): Where we need it to perform a contract with you.For example, we process Identity and Contact Data, Inputs, Outputs and Payment Information in order to provide Services to you. In circumstances where we do not have a contract with you, such as where you are an end user of our Commercial Services, **we instead rely on our legitimate interests**. Where it is necessary for our legitimate interests (or those of a third party) and your interests and rights do not override our interests.Our legitimate interests include: **providing, maintaining and improving our products and services**; research and development, including developing new products and features; marketing our products and services; detecting, preventing and enforcing violations of our terms including misuse of services, fraud, abuse, and other trust and safety protocols; and protecting our rights and the rights of others. Where you have given us your consent . You have the right to withdraw your consent at any time. Where we need to comply with our legal obligations . **We will not use your Inputs or Outputs to train our models, unless: (1) your conversations are flagged for Trust Safety review** (in which case we may use or analyze them to improve our ability to detect and enforce our Usage Policy , including training models for use by our Trust and Safety team,consistent with Anthropic's safety mission), or (2) **you've explicitly reported the materials to us (for example via our feedback mechanisms)**, or (3) **you've otherwise explicitly opted in** to the use of your Inputs and Outputs for training purposes.

6. Data Retention and Data Lifecycle Aggregated or De-Identified Information We may process personal data in an aggregated or de-identified form to analyze the effectiveness of our Services, conduct research, study user behavior, and train our AI models as permitted under applicable laws. For instance: When you submit Feedback and provide us permission, we **disassociate Inputs and Outputs from your user ID** to use them for training and improving our models. If our systems flag Inputs or Outputs for potentially violating our Usage Policy , we disassociate the content from your user ID to train our trust and safety classification models and internal generative models. However, we may re-identify the materials to enforce our Usage Policy with the responsible user if necessary. To improve user experience, we may analyze and aggregate general user behavior and usage data.This information does not identify individual users.]

Terms of Service Excerpt: Anthropic (May 13, 2024)

[4 Our use of Materials. We may use Materials to provide, maintain, and improve the Services and to develop other products and services. We will not train our models on any Materials that are not publicly available, except in two circumstances:1. **If you provide Feedback to us (through the Services or otherwise) regarding any Materials**,we may use that Feedback in accordance with Section 5 (Feedback).2. If your Materials are flagged for trust and safety review, we may use or analyze those Materials to improve our ability to detect and enforce Acceptable Use Policy violations, including training models for use by our trust and safety team, consistent with Anthropic's safety mission.

5. Feedback We appreciate feedback, including ideas and suggestions for improvement or rating an Output in response to a Prompt (" Feedback "). If **you rate an Output in response to a Prompt—for example, by using the thumbs up/thumbs down icon—we will store the related conversation as part of your Feedback**. You have no obligation to give us Feedback, but if you do, you agree that we may use the Feedback however we choose without any obligation or other payment to you.]