Tracking GPTs Third Party Service: Automation, Analysis, and Insights

Chuan Yan
University of Queensland
Brisbane, QLD, Australia

Liuhuo Wan
University of Queensland
Brisbane, QLD, Australia

Bowei Guan
University of Queensland
Brisbane, QLD, Australia

Fengqi Yu
University of Queensland
Brisbane, QLD, Australia

Guangdong Bai
National University of Singapore
University of Queensland
QLD, Australia

Jin Song Dong
National University of Singapore

ABSTRACT

ChatGPT has quickly advanced from simple natural language processing to tackling more sophisticated and specialized tasks. Drawing inspiration from the success of mobile app ecosystems, OpenAI allows developers to create applications that interact with third-party services, known as GPTs. GPTs can choose to leverage third-party services to integrate with specialized APIs for domain-specific applications. However, the way these disclose privacy setting information limits accessibility and analysis, making it challenging to systematically evaluate the data privacy implications of third-party integrate to GPTs. In order to support academic research on the integration of third-party services in GPTs, we introduce GPTs-ThirdSpy, an automated framework designed to extract GPTs' privacy settings. GPTs-ThirdSpy provides academic researchers with real-time, reliable metadata on third-party services used by GPTs, enabling in-depth analysis of their integration, compliance, and potential security risks. By systematically collecting and structuring this data, GPTs-ThirdSpy facilitates large-scale research on the transparency and regulatory challenges associated with the GPT app ecosystem.

CCS CONCEPTS

• **Software and its engineering → Application specific development environments**.

KEYWORDS

Large Language Model, Testing, Privacy

ACM Reference Format:

Chuan Yan, Liuhuo Wan, Bowei Guan, Fengqi Yu, Guangdong Bai, and Jin Song Dong. 2025. Tracking GPTs Third Party Service: Automation, Analysis, and Insights. In *Companion Proceedings of the 33rd ACM Symposium on the Foundations of Software Engineering (FSE '25), June 23–27, 2025, Trondheim, Norway*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

ChatGPT, introduced by OpenAI [14] in 2023, is a leading large language model (LLM) that showcases cutting-edge advancements in AI-powered natural language processing (NLP) technology. It has been widely adopted across various domains, utilizing GPT models to streamline tasks and improve efficiency [12, 17, 19, 24, 27, 28]. As LLMs continue to evolve, user demand for personalized and customizable AI solutions has been steadily increasing. In response this trend, OpenAI introduce the GPT Store [16], featuring GPT applications (GPTs) that not only enable access to real-time data, perform computations, and integrate with third-party services but also allow users to create and customize GPTs without any coding. This innovation marks a significant step toward a more open and diverse GPT ecosystem.

However, while interactions with third-party services enhance the functionality of GPTs, they also introduce significant security risks. Previous studies analyze and confirm these security concerns from multiple perspective. For example, attackers can exploit leaked GPTs data to clone fraudulent GPTs for phishing or scams [10, 22]. Certain third-party services are vulnerable to unauthorized access, leading to potential privacy breaches [25]. And uploaded knowledge files could be inadvertently exposed, increasing the likelihood of data leaks [23]. To conduct these security analyses, it is essential to first obtain the third-party service information provided by GPTs developers. Most existing research relies on foundational GPTs data collected by third-party platforms. However, third-party platforms suffer from poor data timeliness, making it difficult to track updates to GPTs. Researchers often face challenges in obtaining up-to-date information, which limits the accuracy and reliability of analyses. For example, in its latest version, AI PDF Drive [13] specifies the domain names and corresponding privacy policies of three third-party services (`account.myaidrive.com`, `aipdf.myaidrive.com`,`pdf-creator.myaidrive.com`). Nevertheless, GPTs App [7] only records two domains of them, preventing researchers from obtaining a comprehensive view of its most recent privacy compliance updates.

Our work. To reduce experimental costs while improving data timeliness, we develop GPTs-ThirdSpy, a framework for automatically extracting GPTs third-party privacy setting data from the GPT Store. A major challenge is that OpenAI has implemented strict anti-scraping mechanisms [15] and dynamic content loading in the GPT Store, making traditional automation tools such as Selenium [9] and Puppeteer [5] ineffective for interacting with its page structure.

To overcome this, we build on previous research [25] by leveraging AppleScript [2] to simulate user interactions, bypassing the anti-scraping defenses. Additionally, we introduce an innovative approach using cliclick [4], enabling precise position-based clicking to extract the complete GPTs privacy setting data. GPTs-ThirdSpy detects 109 GPTs relying on third-party services within a sample of 500 GPTs and successfully extracts their associated privacy setting data. In future work, we will build a more comprehensive dataset that covers a broader range of GPTs. This enables future research by supporting privacy compliance audits, supporting privacy compliance audits, penetration testing, and other security analyses to comprehensively examine the usage patterns and potential risks of third-party services in GPTs.

Contribution. The main contributions of this work are as follows.

- **Metadata-driven categorization of GPTs.** We categorize GPTs into three distinct types based on their metadata characteristics, providing a structured framework to better understand their functional differences and usage patterns.
- **A systematic security assessment tool.** We propose GPTs-ThirdSpy, a framework designed to automatically detect and extract GPTs third-party service data from the GPT Store. Unlike existing methods that rely on third-party platforms, our framework captures real-time and accurate data, providing a more reliable foundation for future research.
- **Analyzing Third-Party Service Usage in GPT Store and Future Directions.** Our results reveal that the *status quo* of utilizing third-party services in GPTs. This not only provides valuable data for privacy compliance assessments but also supports future research in security evaluations, malicious GPTs detection, and other areas that rely on up-to-date data.

Availability. The source code of GPTs-ThirdSpy and relevant artifacts are available on Github [1].

2 EXPERIMENTAL STUDY

To explore the deployment process and regulations of third-party services on the GPT Store, we specifically develop GPTs as a test case for validation. Table 1 presents the seven types of metadata defined by OpenAI for GPTs. Among them, *name* and *instructions* are mandatory, meaning that every GPTs must include these two metadata types (our work does not consider empty-shell GPTs that only include *name*). The remaining five metadata types are optional, allowing GPTs to incorporate them as needed to enhance functionally and personalization. We use g represents a GPTs, $\mathcal{G}$ represents GPTs set, thus we have $\mathcal{G} = \{g | g : \{name, instructions, description', conversation', knowledge', capabilities', actions'\}\}$, where $'$ denotes optional metadata.

Based on the functions and properties of different GPTs metadata, we define three types of GPTs.

Prompt-based GPTs ($\mathcal{G}_p$). *Prompt-based indicates that the GPTs relies solely on prompts for description and generation, without incorporating any background knowledge or external resources.*

$$g \in \mathcal{G}_p \Leftrightarrow \{knowledge\} \notin g \wedge \{actions\} \notin g \tag{1}$$

Table 1: Metadata of GPTs

Metadata	Explanation
Name†	The non-unique, potentially repetitive identifier for GPTs.
Description‡	A brief summary of what the GPTs does.
Instructions†	Define the behavior and limitations.
Conversation‡	Predefined prompts to guide user interactions.
Knowledge‡	Uploaded files that enhance the GPTs' responses.
Capabilities‡	Additional features the GPTs can use.
Actions‡	Custom third-party API integrations for external interactions.

† Mandatory: Metadata that developers are required to provide on a mandatory basis.
‡ Discretionary: Metadata that developers may choose to provide.

Knowledge-based GPTs ($\mathcal{G}_k$). *Knowledge-based GPTs allow developers to upload local knowledge, such as text documents. By integrating external information, these GPTs enhance their accuracy, retain domain-specific expertise, and generate more context-aware responses.*

$$g \in \mathcal{G}_k \Leftrightarrow \{knowledge\} \in g \wedge \{actions\} \notin g \tag{2}$$

Action-based GPTs ($\mathcal{G}_a$). *Action-based GPTs can interact with external systems by making API calls, enabling dynamic and real-time responses. Unlike prompt-based or knowledge-based, these GPTs can fetch live data, execute commands, and integrate with third-party services.*

$$g \in \mathcal{G}_a \Leftrightarrow \{actions\} \in g \tag{3}$$

Among the three types of GPTs, Prompt-based GPTs are the most fundamental, pioneering a zero-code approach to creating custom applications, allowing users to build their own GPTs simply by writing prompts. A step further is Knowledge-based GPTs, which enable developers to upload local knowledge, equipping GPTs with domain-specific expertise for more accurate responses and personalized services. The only type that interacts with third-party services is Action-based GPTs, which leverage APIs to access external data and perform complex computations. However, this reliance on external services also introduces greater security and privacy risks, including data breaches and unauthorized API access [8, 22, 25]. Therefore, research on this category of GPTs is essential to address these challenges and develop more secure and compliant management frameworks.

3 GPTS-THIRDSPY

Leveraging prior knowledge from the experimental study in Section 2, we design GPTs-ThirdSpy, which can automatically crawl third-party service data from GPTs. Figure 1 illustrates the complete workflow of GPTs-ThirdSpy.

Workflow. Each GPTs has a GizmoID, which serves as its unique 9-character alphanumeric identifier and provides access to its interaction page. Based on this, we first construct the access link based on the GizmoID of each GPTs ❶ (e.g., GizmoID: 1abcD2EFG, Url: https://chatgpt.com/g/g-1abcD2EFG). After successfully accessing

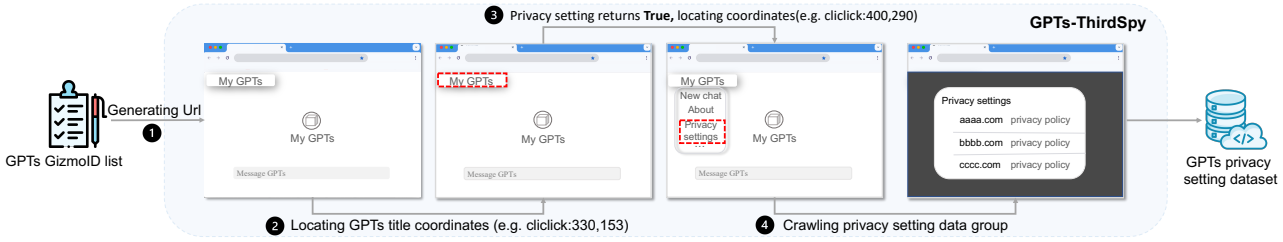

Figure 1: The workflow of GPTs-ThirdSpy

the GPTs interaction page, the next step it to retrieve its detailed information, which will only be shown after clicking the dropdown button. Therefore, we need to locate and click this button. Since the GPTs interaction page has strict anti-automation mechanisms, methods relying on element IDs for interaction, such as JavaScript or Selenium, often fail to execute or trigger CAPTCHA verification, making automated testing ineffective. To overcome this challenge, we use AppleScript in combination with Cliclick. Cliclick is a commandline tool for macOS that simulates mouse clicks. By precomputing the absolute screen coordinates of the GPTs dropdown button, we can bypass anti-automation restrictions and successfully expand the GPTs details panel ❷.

According to OpenAI's guidelines, if a GPTs utilizes external APIs (i.e. Actions), it must provide a corresponding privacy policy. As a result, a "Privacy settings" button appears in the GPTs's dropdown menu. Leveraging this characteristic, we precompute the absolute screen coordinate of the "Privacy settings" button and use Cliclick for precise clicking, ensuring seamless access to the relevant privacy data ❸. When the "Privacy settings" button is clicked, a window pops up displaying the domain names of all third-party services used by the GPTs, along with their corresponding privacy policy links. To comprehensively collect this data, we iterate through each domain entry, extract and parse its associated privacy policy link, and access these links to crawl the underlying privacy policy text. Finally, we organize the collected data into the GPTs privacy setting dataset for further analysis and research ❹.

4 DATASET

Data source and experiment setup. We use the 500 most popular GPTs (ranked by conversation count) provided by GPTsHunter as our data source. Since GPTs-ThirdSpy is built with AppleScript, it relies on the macOS environment. Therefore, we deploy it on two Mac devices: a 16GB M1 Pro and a 16GB M2 Pro.

4.1 Domain Analysis

Figure 2 illustrates the distribution of third-party service usage among the 500 most popular GPTs. Among them, 409 GPTs do not rely on any third-party services. These GPTs operate solely based on their own knowledge or GPT model, without calling external API or transmitting additional data, making them more self-contained in terms of privacy protection. 79 GPTs integrate a single third-party service, primarily to extend their functionality, such as accessing external databases, providing real-time information, or enhancing

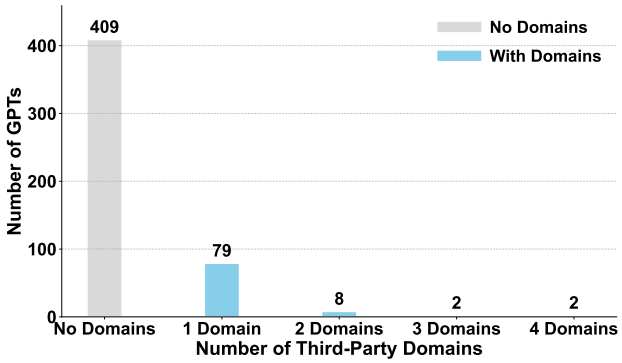

Figure 2: GPTs domain count distribution

computational capabilities. This represents the most common pattern of third-party service usage among GPTs. Additionally, 12 GPTs utilize two or more third-party services. Manual verification reveals that these GPTs often require more advanced functionalities, such as cross-platform integrations, real-time data analysis, or leveraging multiple AI APIs for task optimization. However, this also introduces greater security and privacy risks, as data may be processed across multiple external systems, increasing the complexity of access control and the potential for information leakage. We leave the analysis and assessment of these security and privacy risks for future work.

Figure 3 shows the distribution of these third-party service domains. Third-party services for AI application development are the most frequently used, with services such as `gpts.WebPilot.ai` [20], `gpt-tools.co` [6] and `b12.io` [3]. Following closely behind are third-party services that leverage AI for specialized tasks. For example, `swan-api.jobright.ai` [11] uses AI to assist users in job searching, improving efficiency and optimizing the matching process.

4.2 Privacy Policy

We systematically access and store the text content of privacy policy links in our dataset for further analysis. Although OpenAI mandates that GPTs using third-party services provide corresponding privacy policy links, it does not enforce content review, leading to potential issues with validity and compliance. Among the 109 domains associated with third-party services, 92 privacy policy links are accessible, while others exhibit various issues. 9

Figure 3: Domain Usage Distribution of GPTs

Table 2: Privacy policy accessibility distribution

Accessible	Inaccessible			
-	Broken link	Official website	Timeout	Server error
92	9	5	2	2

links are invalid, including 1 GPTs that directly uses a placeholder link (i.e. `https://app.example.com/privacy_policy`). 5 GPTs provide only the homepage of the third-party service instead of a dedicated privacy policy, which could prevent users from obtaining clear information on data handling and privacy protection. Additionally, 2 links result in connection timeouts, while another 2 return server errors. These findings indicate that despite OpenAI's requirement for privacy policy links, many GPTs still fail to provide valid or compliant policies, underscoring the need for stricter review and regulatory mechanisms to ensure users have access to meaningful full and enforceable privacy data.

5 DISCUSSION

GPTs-ThirdSpy addresses the issue of outdated GPTs metadata, enabling developers to access the latest third-party service data directly from the official GPT Store in real-time, ensuring the accuracy and reliability of research and applications. Based on these features, GPTs-ThirdSpy can be further extended in several key areas.

Privacy compliance. Academic researchers and industry regulators can use this tool to monitor the enforcement of GPTs' privacy policy in real-time, identifying non-compliant third-party data-sharing practices and promoting a more transparent and secure AI ecosystem, building upon prior work [21, 26]. Additionally, GPTs-ThirdSpy can be extended to automatically detect inconsistencies in privacy policies, analyzing whether GPTs provide contradictory privacy statements across different periods or platforms.

Third-Party service monitoring. GPTs-ThirdSpy also can be used to track changes in third-party services that GPTs rely on, allowing it to detect when a GPTs adds or replaces an external API. In such cases, the system can issue alerts, prompting regulatory agencies or enterprises to reassess data access permissions. This functionality is particularly valuable for identifying sudden shifts in data-sharing practices, ensuring that new integrations comply with privacy and security standards, and preventing unauthorized access to sensitive user information. Moreover, by maintaining a historical record of these changes, GPTs-ThirdSpy enables trend analysis, helping researchers and policymakers understand long-term patterns in GPTs' reliance on third-party services and anticipate potential regulatory challenges.

Penetration testing. GPTs-ThirdSpy can be further extended to conduct penetration testing on third-party service domains relied upon by GPTs, allowing for security assessments and the identification of potential vulnerabilities. Since third-party APIs used by GPTs could serve as entry points for data breaches or supply chain attacks [18], the tool can automatically identify and track these domains while integrating penetration testing techniques for comprehensive security evaluation. GPTs-ThirdSpy also monitors third-party services for risks such as API abuse, authentication flaws, and potential privilege escalation vulnerabilities, enabling enterprises, researchers, and regulatory agencies to assess the security of the GPT app ecosystem more effectively and proactively mitigate emerging threats.

6 CONCLUSION

In this work, we first categorize GPTs into three distinct types based on their metadata characteristics. This categorization not only enhances understanding of GPTs in terms of functionality, dependencies, and privacy compliance but also establishes a systematic foundation for future research. Additionally, we develop GPTs-ThirdSpy, an automated framework for extracting third-party service data from GPTs. This framework enables real-time and precise retrieval of GPTs' interactions with external services, providing researchers with a reliable analytical foundation.

ACKNOWLEDGMENTS

This research has been partially supported by Australian Research Council Discovery Projects (DP230101196, DP240103068) and the Ministry of Education, Singapore under its Academic Research Fund Tier 3 (MOET32020-0003).

REFERENCES

[1] 2025. *Tracking GPTs Third Party Service: Automation, Analysis, and Insights (GPTs-ThirdSpy Source Code).* https://github.com/UQ-Trust-Lab/GPTs-ThirdSpy
[2] Apple. 2024. *Introduction to AppleScript Language Guide.* https://developer.apple.com/library/archive/documentation/AppleScript/Conceptual/AppleScriptLangGuide/introduction/ASLR_intro.html
[3] b12. 2024. *b12 Official Website.* https://www.b12.io/
[4] Carsten Blüm. 2024. *cliclick: macOS CLI tool for emulating mouse and keyboard events.* https://github.com/BlueM/cliclick
[5] Google. 2024. *Puppeteer website.* https://pptr.dev/
[6] gptbuilder. 2024. *gptbuilder Official Website.* https://gptbuilder.tools/
[7] GPTsApp.io. 2024. *AI PDF Drive: Chat, Create, Organize.* https://gptsapp.io/gpts/ai-pdf-ai/xf37cpcez
[8] Xinyi Hou, Yanjie Zhao, and Haoyu Wang. 2024. On the (in) security of llm app stores. *arXiv preprint arXiv:2407.08422* (2024).
[9] Jason Huggins. 2024. *Selenium website.* https://www.selenium.dev/
[10] Umar Iqbal, Tadayoshi Kohno, and Franziska Roesner. 2024. Llm platform security: applying a systematic evaluation framework to openai's chatgpt plugins. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Vol. 7. 611–623.
[11] jobright. 2024. *jobright Official Website.* https://jobright.ai/
[12] Ningke Li, Yuekang Li, Yi Liu, Ling Shi, Kailong Wang, and Haoyu Wang. 2024. Drowzee: Metamorphic testing for fact-conflicting hallucination detection in large language models. *Proceedings of the ACM on Programming Languages* 8, OOPSLA2 (2024), 1843–1872.
[13] myaidrive. 2024. *AI PDF Drive: Chat, Create, Organize.* https://chatgpt.com/g/g-V2KlUZSj0-ai-pdf-drive-chat-create-organize

[14] OpenAI. 2023. *OpenAI official website.* https://openai.com/
[15] OpenAI. 2024. *ChatGPT: Verify that you are human.* https://community.openai.com/t/verify-that-you-are-human-stop-it/857988
[16] OpenAI. 2024. *Introducing the GPT Store.* https://openai.com/index/introducing-the-gpt-store/
[17] Liuhuo Wan, Kailong Wang, Kulani Mahadewa, Haoyu Wang, and Guangdong Bai. 2024. Don't Bite Off More than You Can Chew: Investigating Excessive Permission Requests in Trigger-Action Integrations. In *Proceedings of the ACM Web Conference 2024*. 3106–3116.
[18] Liuhuo Wan, Kailong Wang, Haoyu Wang, and Guangdong Bai. 2024. Is it safe to share your files? an empirical security analysis of google workspace. In *Proceedings of the ACM Web Conference 2024*. 1892–1901.
[19] Zihan Wang, Zhongkui Ma, Xinguo Feng, Ruoxi Sun, Hu Wang, Minhui Xue, and Guangdong Bai. 2024. Corelocker: Neuron-level usage control. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2497–2514.
[20] WebPilot. 2024. *WebPilot Official Website.* https://www.webpilot.ai/post-gpts/
[21] Fuman Xie, Yanjun Zhang, Chuan Yan, Suwan Li, Lei Bu, Kai Chen, Zi Huang, and Guangdong Bai. 2022. Scrutinizing privacy policy compliance of virtual personal assistant apps. In *Proceedings of the 37th IEEE/ACM international conference on automated software engineering*. 1–13.
[22] Yinglin Xie, Xinyi Hou, Yanjie Zhao, Kai Chen, and Haoyu Wang. 2024. LLM App Squatting and Cloning. *arXiv preprint arXiv:2411.07518* (2024).
[23] Chuan Yan, Bowei Guan, Yazhi Li, Mark Huasong Meng, Liuhuo Wan, and Guangdong Bai. 2025. Understanding and Detecting File Knowledge Leakage in GPT App Ecosystem. In *THE WEB CONFERENCE 2025*.
[24] Chuan Yan, Mark Huasong Meng, Fuman Xie, and Guangdong Bai. 2024. Investigating Documented Privacy Changes in Android OS. *Proceedings of the ACM on Software Engineering* 1, FSE (2024), 2701–2724.
[25] Chuan Yan, Ruomai Ren, Mark Huasong Meng, Liuhuo Wan, Tian Yang Ooi, and Guangdong Bai. 2024. Exploring chatgpt app ecosystem: Distribution, deployment and security. In *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering*. 1370–1382.
[26] Chuan Yan, Fuman Xie, Mark Huasong Meng, Yanjun Zhang, and Guangdong Bai. 2024. On the quality of privacy policy documents of virtual personal assistant applications. *Proceedings on Privacy Enhancing Technologies* (2024).
[27] Zhiqiang Yuan, Mingwei Liu, Shiji Ding, Kaixin Wang, Yixuan Chen, Xin Peng, and Yiling Lou. 2024. Evaluating and improving chatgpt for unit test generation. *Proceedings of the ACM on Software Engineering* 1, FSE (2024), 1703–1726.
[28] Xiaoyu Zhang, Juan Zhai, Shiqing Ma, Qingshuang Bao, Weipeng Jiang, Chao Shen, and Yang Liu. 2025. Unveiling Provider Bias in Large Language Models for Code Generation. *arXiv preprint arXiv:2501.07849* (2025).