THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

Exploring ChatGPT App Ecosystem: Distribution, Deployment and Security

Chuan Yan†, Ruomai Ren†, Mark Huasong Meng§, Liuhuo Wan†, Tian Yang Ooi†, Guangdong Bai†

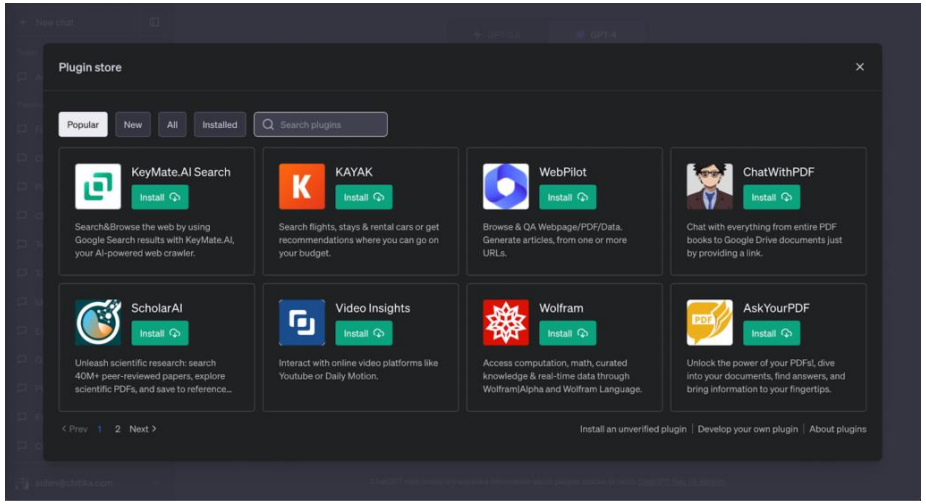
† The University of Queensland

§ Technical University of Munich Germany

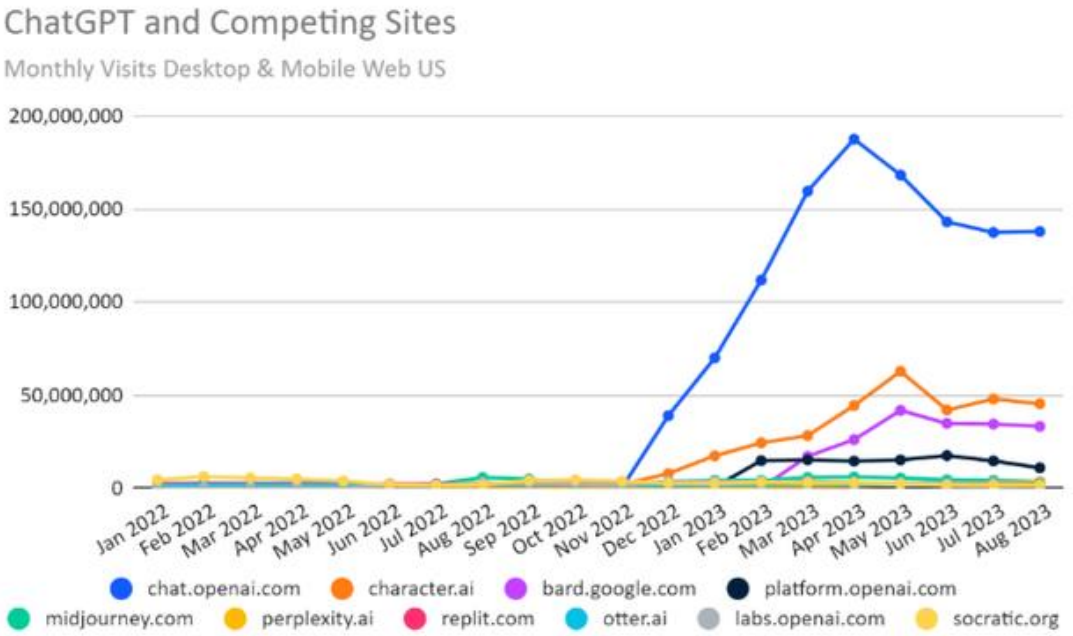
39th IEEE/ACM International Conference on Automated Software Engineering

Background

ChatGPT is a flagship LLM product of OpenAI launched in 2023. It represents the state-of-the-art advancement in AI-driven natural language processing technology.

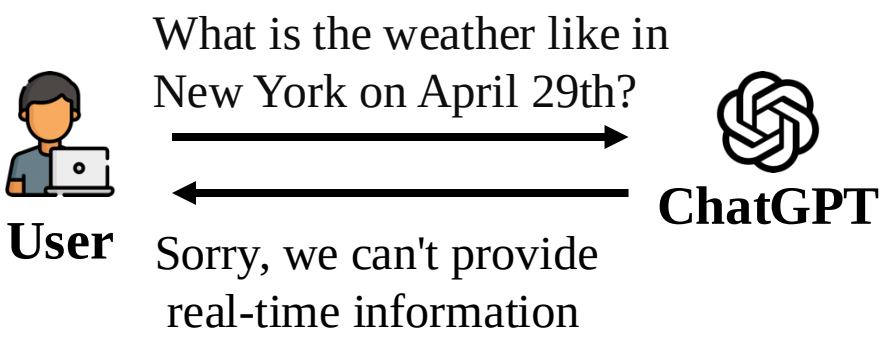


ChatGPT plugin store screenshot



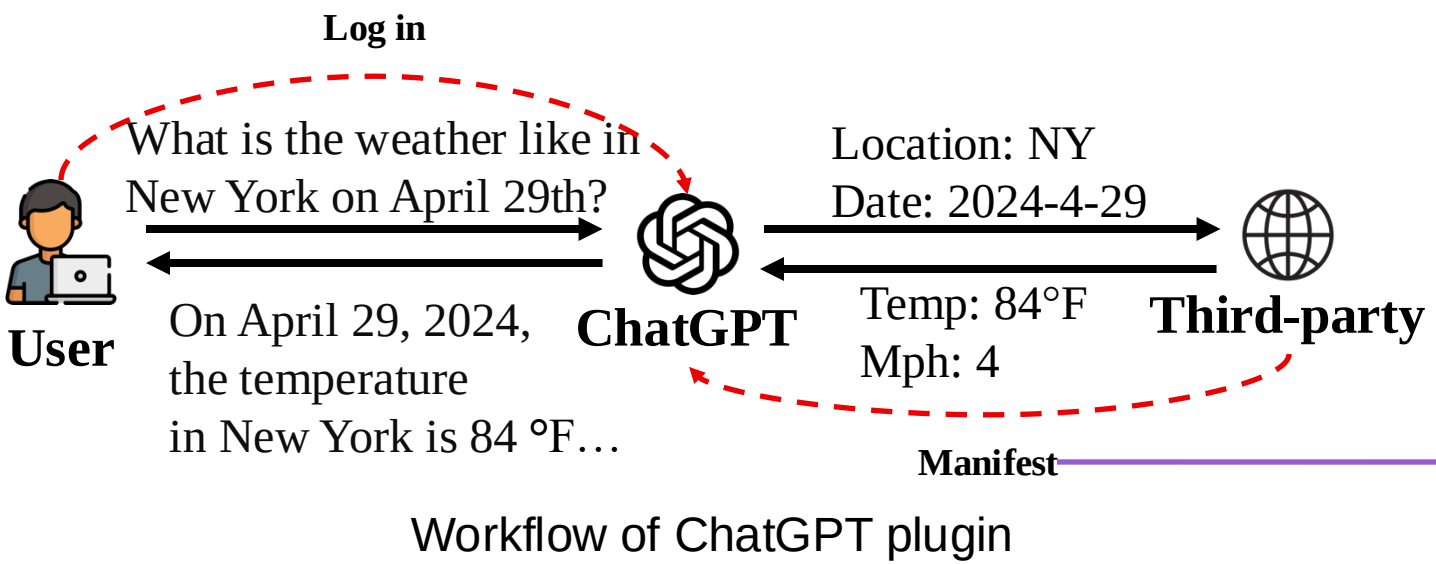
Changes in ChatGPT visits since the beginning of 2023

Background


What is the weather like in New York on April 29th?
User
ChatGPT
Sorry, we can't provide real-time information

Background

These plugins allow ChatGPT to connect the **current conversation** to **external data** sources and **services**, such as mobile apps with internet access.



Workflow of ChatGPT plugin

Plugin manifest webpage

https://weatherMananger.com/.well-known/ai-plugin.json

{"name_for_human": Weather Manager,
"name_for_model": Weather Manager,
"description_for_human": Your Weather manager, the weather conditions of your location anytime,
"description_for_model":…,
"auth":{"type": "none"},
"api": {"type": "openai", "url":…},
"logo_url": …,
"contact_email": …,
"legal_info_url": …}

Our Works

Characteristics of existing apps, app development, deployment and distribution mechanisms, security and privacy implications

Our study examines the **distribution and deployment models** in the integration of LLMs and third-party apps, and assesses their **security and privacy implications.**

RQ1: *what are the characteristics of the plugins available in the store*

RQ2: *what are the deployment model and runtime execution model that integrate third-party apps and LLMs*

RQ3: *what are the security and privacy issues associated with the integration*

RQ1: Characteristics of the plugins

1038 plugins

App Store Google Play amazon appstore

Plugin Category Definitions

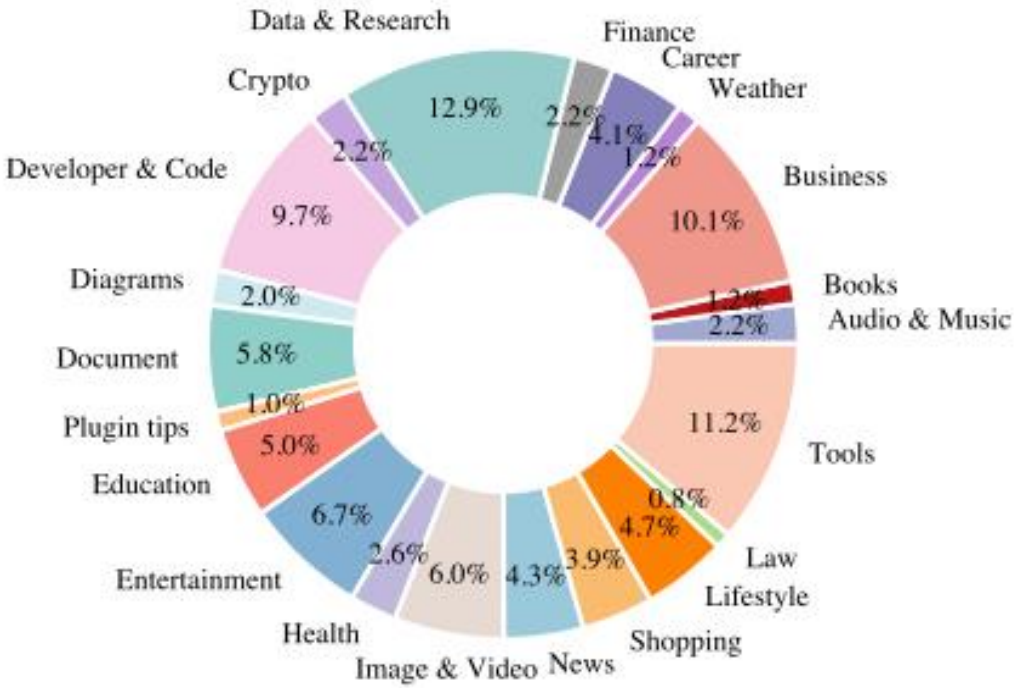
21 plugin categories

BART-large-mnli

Classification

Accuracy	Macro-F1 score
83.3%	0.91

Benchmarking performance

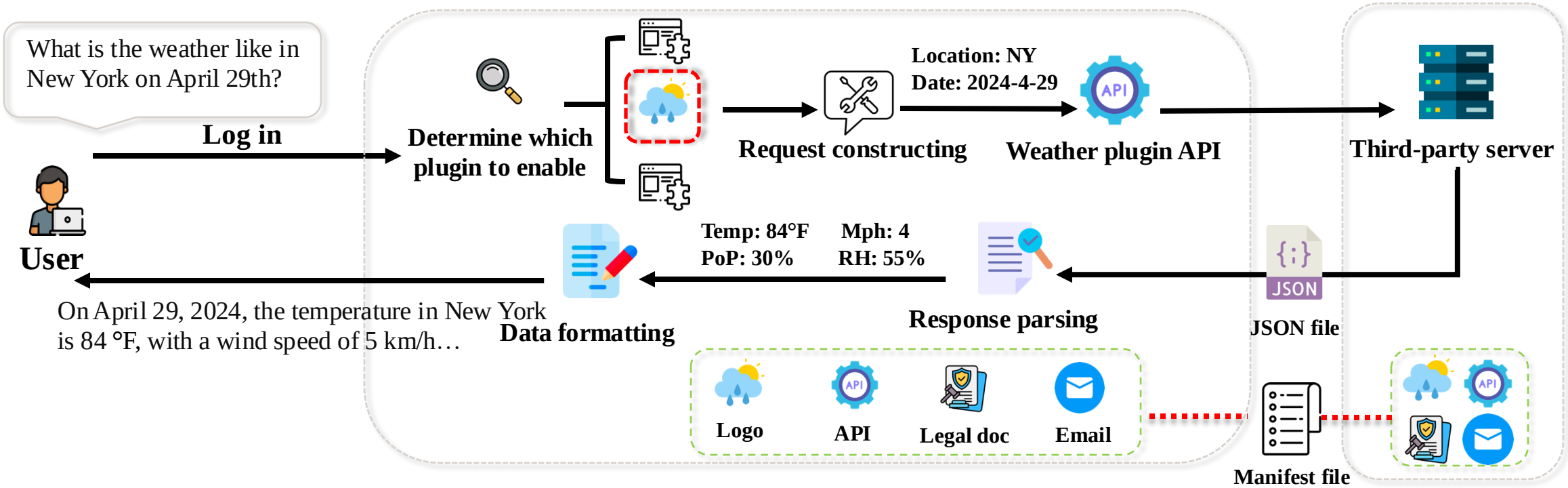


Distribution of the number of plugins for the 21 categories

The distribution of country-specific plugins

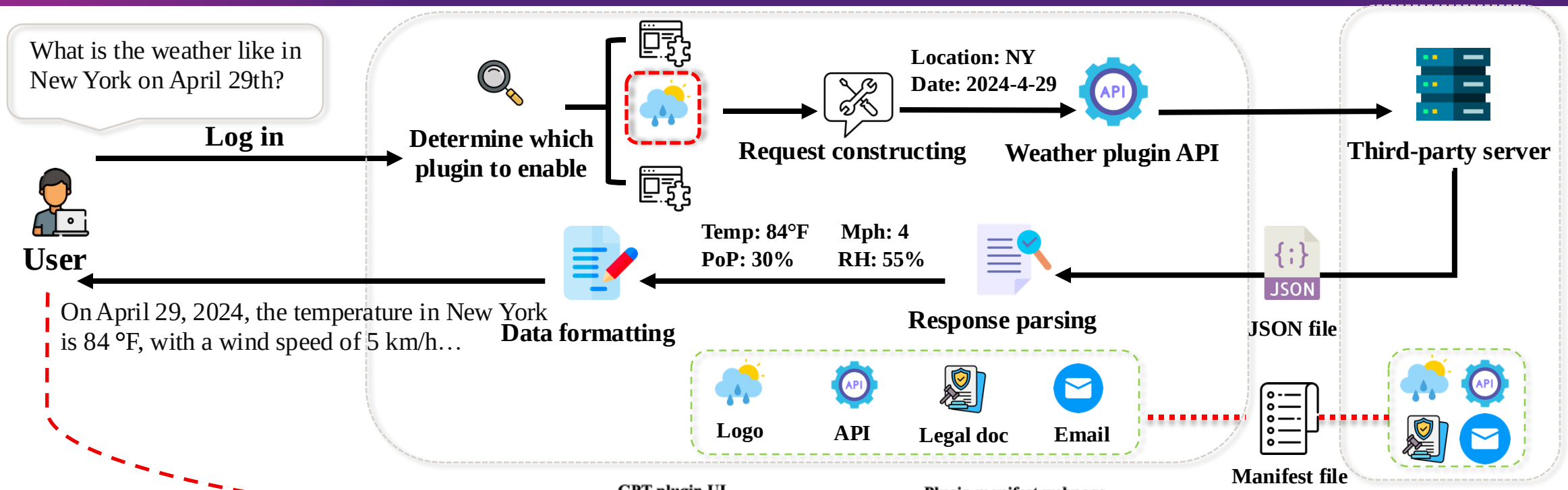
Country & Region	Plugin Number	Country & Region	Plugin Number
Japan	27	USA	24
UK	8	Australia	6
Korea	5	Canada	4
Germany	4	Singapore	4
China	3	Switzerland	3
Austria	1	Brazil	1
India	1	Ireland	1
Israel	1	Italy	1
Taiwan	1	Portugal	1
Turkey	1	Netherlands	1

Understanding app deployment and runtime execution



The workflow of security assessment model based on the plugin operating mechanism

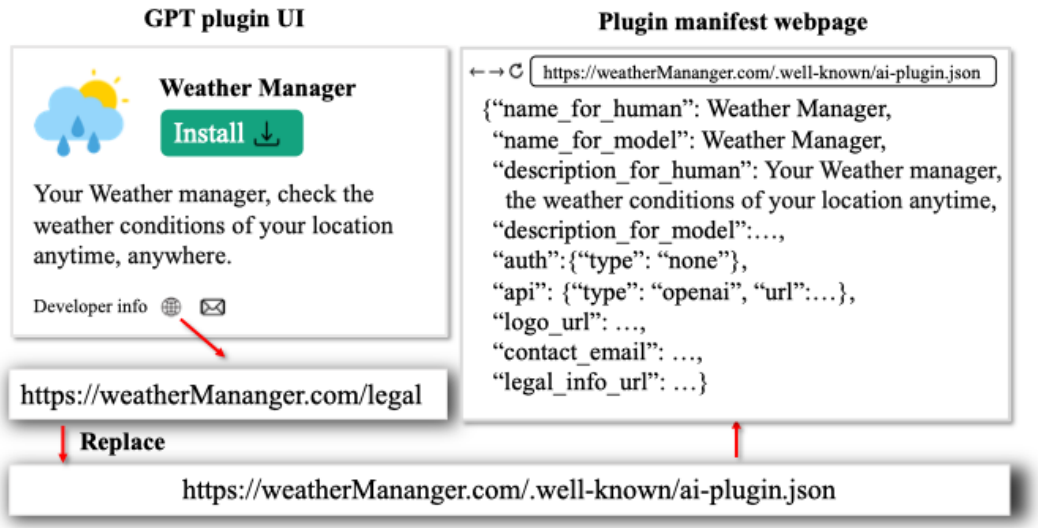
RQ3: Security and privacy issues associated with the integration



File leakage detection layer

Exposure 1: Non-Empty Manifest

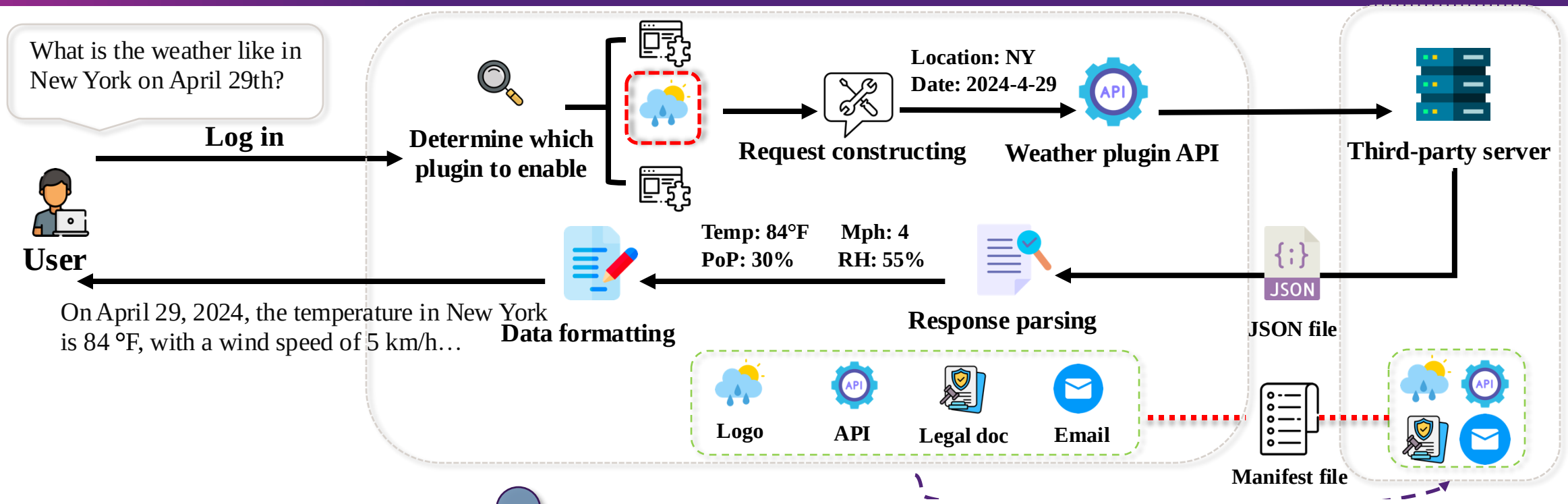
$$\exists p \in \mathsf{P},\ m \neq \emptyset$$



The process of getting the plugin manifest file

We totally detect 368 (35.7%) plugins that leak manifest files, which enables external access to the plugin's configuration.

RQ3: Security and privacy issues



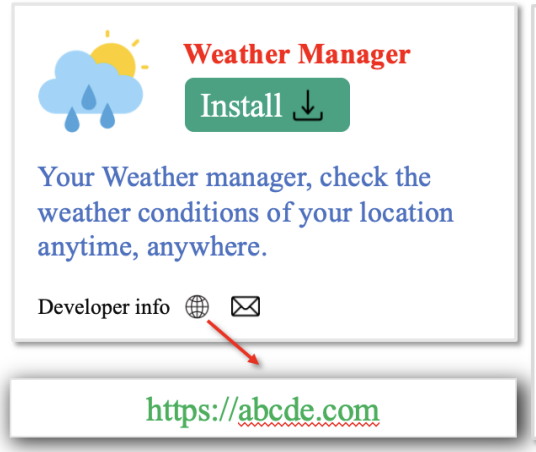
Data consistency layer

Expo sure 2: Discrepancies Metadata

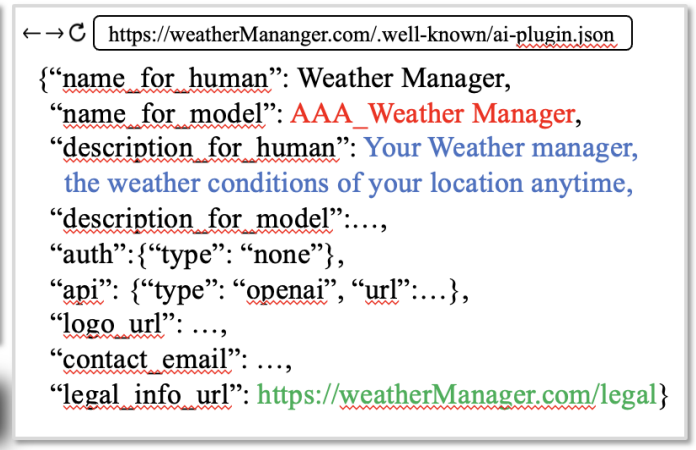
$\exists p \in P, (sim(a, am) < \vartheta 1) \wedge$
$(sim(d, am) < \vartheta 2) \vee (sim(u, am) < \vartheta 3)$

We have identified a total of 69 plugins where the metadata submitted to OpenAI does not match the metadata provided to users.

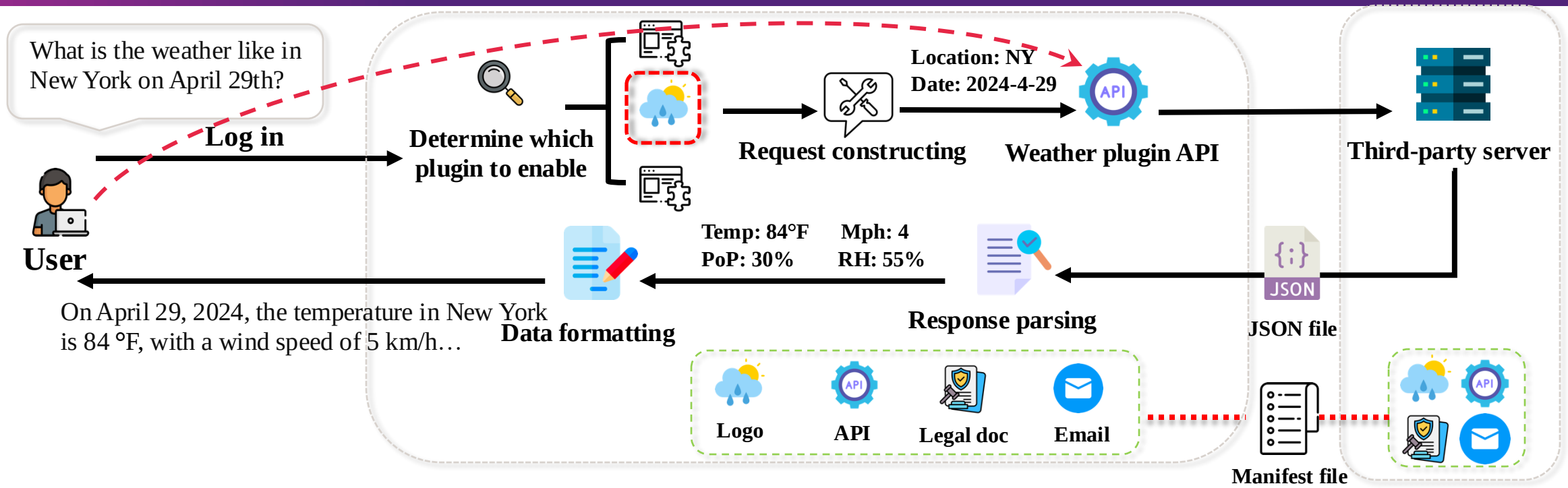
GPT plugin UI



Plugin manifest webpage



RQ3: Security and privacy issues



API-authorized testing layer

Exposure 3: Single Authentication External API Calls

$\exists p \in \mathrm{P}, km \neq \emptyset \vee \neg hm,$
$\exists a \in \mathrm{A}, \forall x, rx \neq \emptyset \wedge rx \notin \mathrm{N}$

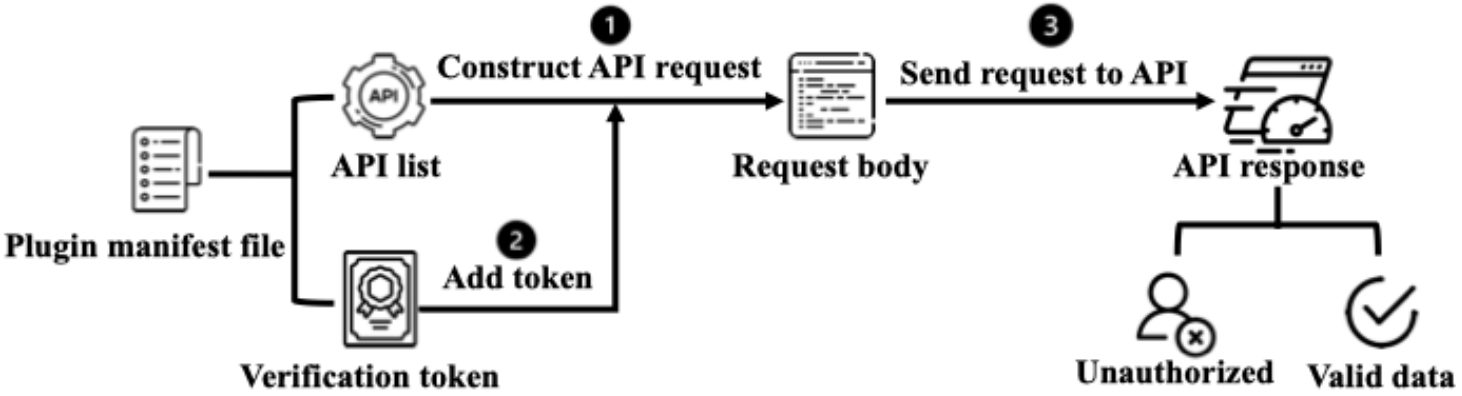
Exposure 4: Multi-Authentication External API Calls

$\exists p \in \mathrm{P}, km \neq \emptyset \vee hm,$
$\exists a \in \mathrm{A}, \forall x, rx \neq \emptyset \wedge rx \notin \mathrm{N}$

Exposure 5: Token Leakage

$\exists p \in \mathrm{P}, km, tm \neq \emptyset \vee hm,$
$\exists a \in \mathrm{A}, \forall x, rx \neq \emptyset \wedge rx \notin \mathrm{N}$

RQ3: Security and privacy issues



The process of verifying the external accessibility of APIs at the API-authorized testing layer

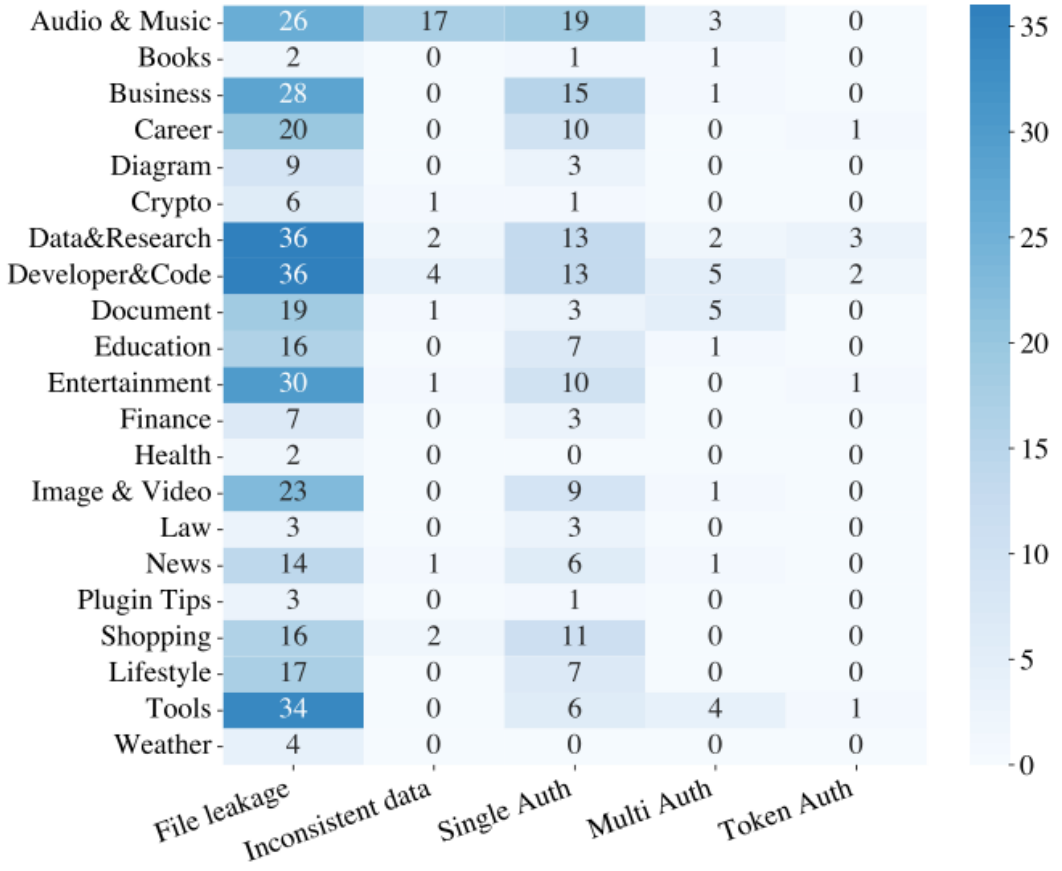
The distribution of API responses

API responsiveness	Auth types			Reasons			
	none	others	Token†	Change‡	Unauthorized	Client errors	Rate limiting
respondable	141	32	8	5	-	-	-
non-respondable	87	72	-	-	55	62	42

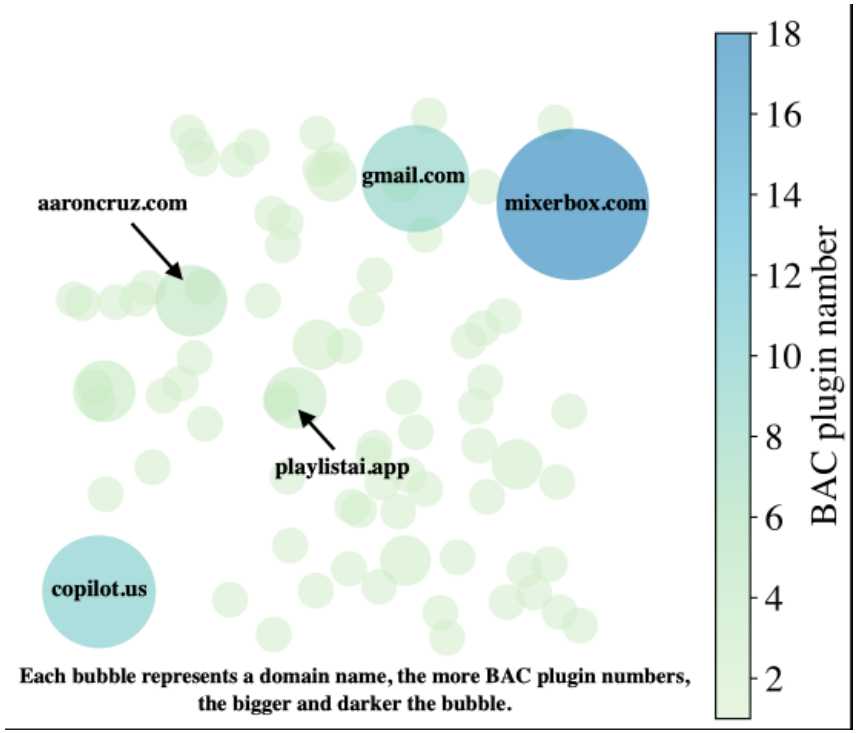
† Token: Include verification token in the API request body.
‡ Change: Manually requestable.

173 (52.1%) plugins can retrieve valid information from their specialized API for OpenAI, including 141 with auth type is none. For plugins with authentication requirements, 24 plugins can return valid information, while 8 plugins are able to retrieve valid information after adding verification OpenAI token.

RQ3: Security and privacy issues associated with the integration

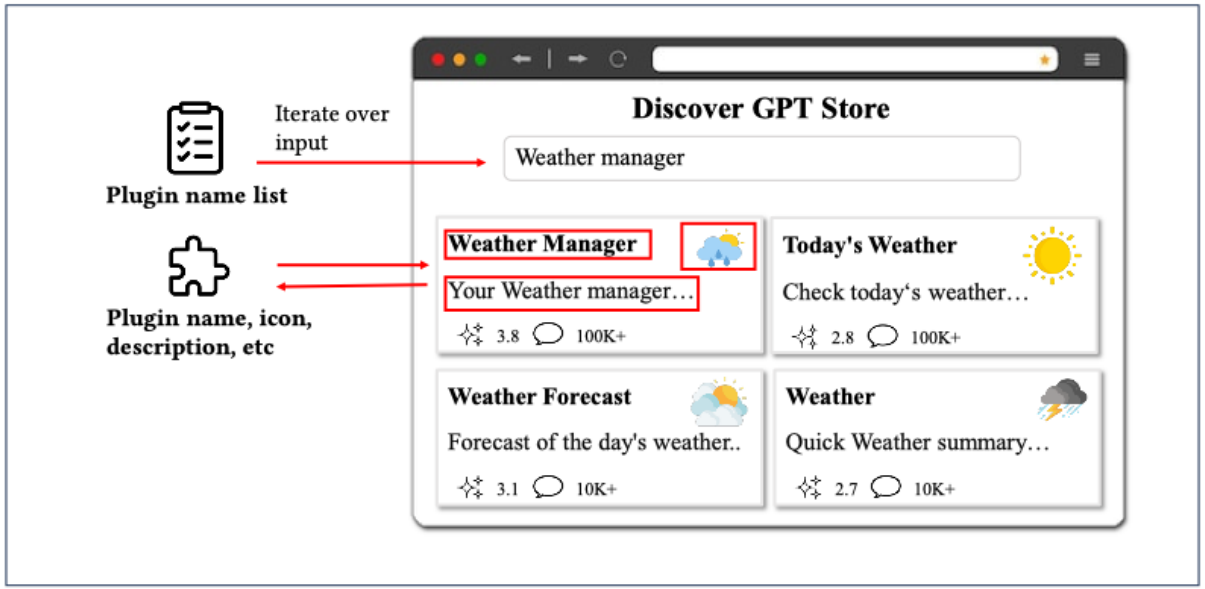


Five types exposures in different plugin categories



The distribution of developers' email domains for the plugins found to have BAC vulnerabilities

Legacy Plugins in GPT Store



The process of detecting plugins in the GPT store

We discover that out of 1038 plugins, 417 are still available in the GPT store as GPTs. Among them, 70 have previously leaked manifest files, and 41 are still externally accessible through external API requests.

Conclusion

- **A comprehensive characterization of ChatGPT plugins.**

  We summarize functionalities provided by these apps, offering an overview to app users.

- **A systematic security assessment and practical impact.**

  We reveal the deployment and runtime execution mechanisms of ChatGPT plugins for the first time.
  Based on that, we propose a three-layer security assessment to evaluate the resource and data exposure associated with ChatGPT plugins.

- **Revealing the status quo and development trajectory of ChatGPT plugin store.**

  Our findings indicate that the ChatGPT app ecosystem is still in a nascent stage in providing rich functionalities comparable with its mobile counterparts. It also lacks a mature regulatory mechanism to enforce user privacy compliance and security standards.

Thank you

Chuan Yan(*c.yan@uq.edu.au*)

TrustLab : *https://trustlab.uqcloud.net/*

UQ TrustLab

TrustLab@UQ is a research team at The University of Queensland, Australia. Our primary objective is to develop theories, techniques, and systems that can ensure the trustworthiness of complex software and intelligence systems. It strives to enhance security, robustness and reliability of software and intelligence systems, and make them more trustworthy for the individuals and organizations that rely on them.