Privacy Perceptions of Custom GPTs by Users and Creators

Rongjun Ma
rongjun.ma@aalto.fi
Aalto University
Finland

Caterina Maidhof
cmaidho@upv.edu.es
VRAIN, Universitat Politècnica de València
Spain

Juan-Carlos Carrillo
juaca10j@upv.es
VRAIN, Universitat Politècnica de València
Spain

Janne Lindqvist
janne.lindqvist@aalto.fi
Aalto University
Finland

Jose Such
jose.such@kcl.ac.uk
King's College London & VRAIN, Universitat Politècnica de València
UK, Spain

ABSTRACT

GPTs are customized LLM apps built on OpenAI's large language model. Any individual or organization can use and create GPTs without needing programming skills. However, the rapid proliferation of over three million GPTs has raised significant privacy concerns. To explore the privacy perspectives of users and creators, we interviewed 23 GPT users with varying levels of creation experience. Our findings reveal blurred lines between user and creator roles and their understanding of GPT data flows. Participants raised concerns about data handling during collection, processing, and dissemination, alongside the lack of privacy regulations. Creators also worried about loss of their proprietary knowledge. In response, participants adopted practices like self-censoring input, evaluating GPT actions, and minimizing usage traces. Focusing on the dual role of user-creators, we find that expertise and responsibility shape privacy perceptions. Based on these insights, we propose practical recommendations to improve data transparency and platform regulations.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; • **Security and privacy → Usability in security and privacy**.

KEYWORDS

GPTs, LLM Apps, Privacy Concerns, Privacy Practices, Interviews, Empirical Studies

ACM Reference Format:
Rongjun Ma, Caterina Maidhof, Juan-Carlos Carrillo, Janne Lindqvist, and Jose Such. 2025. Privacy Perceptions of Custom GPTs by Users and Creators. In *Proceedings of Author's version of the paper accepted for publication.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

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

1 INTRODUCTION

GPTs are customized apps built on OpenAI's Large Language Models (LLMs) to serve specific task needs, often integrating third-party services [76]. Consider, for example, a digital assistant that users can interact with to manage their daily agenda. By integrating with calendar services like Google Calendar (with user permission), these GPTs can automate and streamline task management. They can schedule appointments and provide timely reminders for upcoming meetings and events, all based on conversational interactions with users (e.g., [43, 91]). Beyond this example, a wide variety of GPTs are created by individual and organizational subscribers of OpenAI to serve diverse needs; all these are accessible in the GPT store operated by OpenAI [83]. By January 2024, over *three million GPTs* have been developed [83], with 159,000 available for public use [93]. These GPTs attract approximately 6.1 million visits per month [93].

With the emerging usage of GPTs, security and privacy concerns have been raised, affecting both end users and the internal models of GPTs [9, 16, 103, 113, 116]. For example, malicious GPTs can be created to steal personal information input by end users [9, 16, 103]. Additionally, GPTs have been identified to be vulnerable to a range of threats, such as spoofing, repudiation, and tampering [103]. Not only end users but also the GPTs internals, such as the knowledge base from creators to build GPTs, are at risk of compromise from attacks like prompt injection [113, 116]. Previous research has highlighted the unresolved risks associated with GPTs. However, little is known about how users of the GPT platform actually use or create GPTs and how they perceive the privacy of these GPTs.

Understanding the privacy implications of GPTs is important for several reasons. First, GPTs provide a valuable case for understanding how users perceive privacy in the customization of LLM applications. While many studies have explored technical methods to safeguard user privacy while maintaining system performance through privacy-preserving algorithms and system design [4, 51, 61, 67], there is limited understanding from the user's perspective, such as how willing users are to compromise their privacy in exchange for LLM customizations. As increasing customization and intelligence in LLM apps require more user data, understanding users' perspectives becomes crucial. Understanding users' privacy concerns, practices, and demand for more customized technology is essential for mitigating security and privacy risks, offering guidance to

the rapidly expanding user community, and designing technology securely and ethically to meet their expectations.

Second, the GPT store offers users access to a diverse array of GPTs designed for specific tasks and enables them to create new GPTs with ease. The threshold for creating GPTs is remarkably low, allowing any subscribed individual or organization to develop their own GPTs without the need for advanced AI or programming skills. Moreover, users and creators share the same type of account within OpenAI's services and are treated uniformly, without any distinction between the two roles. This flexibility allows users to take on multiple roles, simultaneously using GPTs while also creating them, in contrast to the traditional view that users and developers have clearly defined, separate roles [11, 23, 34, 62, 94]. This dual engagement may influence their perspectives on privacy, where individuals act as both users and creators. However, it remains unclear who exactly is creating GPTs. Are individuals developing GPTs solely for their own needs, or are some creating them for others? These questions highlight the importance of examining the role of GPT creators and audiences to develop a more comprehensive understanding of privacy perceptions related to GPTs.

To systematically study the privacy perceptions of both users and creators of GPTs, we ask the following research questions:

- RQ1: What are users' and creators' mental models of data flows in GPTs?
- RQ2: What are users' and creators' privacy concerns and their practices responding to their concerns about GPTs?
- RQ3: Who is creating GPTs, and how do the dual roles (creators and users) influence their privacy perceptions?

In this work, we conducted semi-structured interviews (N=23) across a spectrum of roles on the GPT platform, ranging from end users of GPTs to experienced professional creators developing popular GPTs. The interviews focused on their mental models of data flow in GPTs, privacy concerns and practices, and reflections on their roles.

Our major contributions are:

- a) We present users' and creators' mental models of data flow in GPTs usage, along with their privacy concerns and practices regarding GPTs usage.
- b) We identify a spectrum of roles, from users to creators of GPTs, shaped by their creation experiences and motivations, providing insights into how this spectrum influences privacy perceptions.
- c) We discuss the implications of our findings based on the interviews and suggest recommendations for platforms, regulators, and researchers to improve the design and regulation of GPTs and similar LLM applications.

2 USAGE SCENARIOS OF GPTS

There are eight categories within OpenAI's ecosystem of 3 million GPTs, including: DALL-E (e.g., generating and refining images [6]), Writing (e.g., writing assistant, with a focus on relevance and word count [89]), Programming (e.g., creating websites [39]), Research and Analysis (e.g., searching academic papers with citations [28]), Education (e.g., solving math problems [88]), Productivity (e.g., designing presentations and logos [20]), Lifestyle (e.g., picking suit colors for individuals [5]) and "Top Picks" from the week.

Based on previous studies on GPT data practices [45], OpenAI's documentation [75], and exploration by developing demo GPTs, we categorize GPTs into three primary scenarios distinguished by their functions and usage:

Scenario 1: Basic GPTs are created using the *Built-in Capabilities* provided by OpenAI. These capabilities can be easily activated by selecting checkboxes in the GPT creation interface. Creators can then set up GPTs with basic prompt instructions or add domain-specific knowledge. For example, a legal GPT can be developed by uploading legal documents and specifying prompts for the GPT to act as a lawyer. In this type of GPT, all data remains within OpenAI's infrastructure. Creators do not have access to any user data, including the prompts that users interact with in their GPT apps. The only information available to creators is the accumulated count of conversations and the overall user rating.

Scenario 2: Action-based GPTs can access third-party services by integrating APIs that connect to external platforms. *Actions* need to be implemented as HTTP APIs by creators and exposed to OpenAI in a JSON format [75]. Each GPT can include multiple *Actions*, allowing it to connect to one or more third-party services. Depending on the API configuration, creators may be able to extract user information, including conversations, and transmit this data to external services outside of OpenAI's infrastructure [45]. For example, a scholarly GPT might locate resources in a specific library based on user instructions, with access to the library established through an API that becomes an *Action* of the GPT. Figure 1 illustrates the potential data transmission paths in this type of GPT. Importantly, for action-based GPTs, each *Action* requires the user's permission before the GPT accesses external services. When an *Action* is called, users are given the options to "Allow" (permit once), "Always Allow" (grant ongoing permission for this GPT), or "Decline" (no execution of the action).

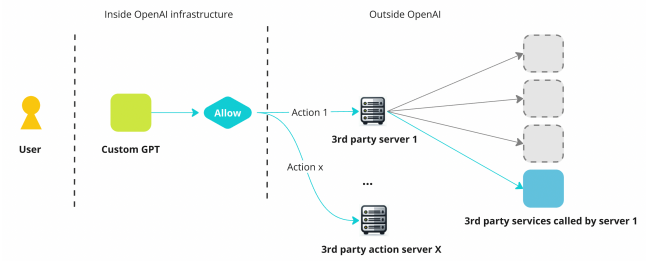

Figure 1: (Scenario 2) A demonstration of action-based GPTs that connect to multiple third-party services.

Scenario 3: Some action-based GPTs integrate a *Login* to third-party services. Similar to Scenario 2, the API is configured by the creators, allowing each GPT to initiate multiple login requests and connect to various third-party services. Depending on the API setup, creators may be able to extract user information, including conversations, and transmit this data to external services outside of OpenAI's infrastructure. For example, a music GPT might suggest a playlist based on user preferences and automatically add it to the user's music account. In this scenario, the login process triggers an authentication flow, redirecting users to a third-party service for

the *Login*. After a successful *Login*, users are redirected back to the original GPT, which can execute actions using the authentication confirmation, such as a token. Figure 2 illustrates the potential data transmission paths in this type of GPT. As with other action-based GPTs, each action requires explicit user permission before execution.

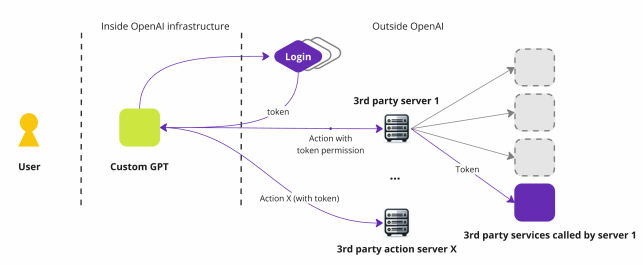

Figure 2: (Scenario 3) A demonstration of login-based GPTs that require user login authentication and connect to multiple third-party services.

3 RELATED WORK

3.1 Privacy Issues and Safeguards in LLMs

The extensive data storage and rapid information distribution of LLMs pose challenges to various aspects of privacy, including personal data control and regulation [60, 72, 109].

LLMs collect vast amounts of data, which is often stored, processed, reused for machine learning, and shared with third parties [60, 107]. Privacy violations may occur during data processing [60, 109], as data points may be linked to identify individuals and infer sensitive traits like sexual orientation, gender, or religious beliefs [109]. This can lead to the creation of detailed profiles without the individual's knowledge or consent [60]. Moreover, users typically lack control and are not informed about how their data is used, including its inclusion in training datasets, making it vulnerable to leaks and unauthorized access.

The dissemination of sensitive information from LLM training data also poses significant risks [60, 109]. For example, GPT-2 has unintentionally revealed personal details like phone numbers and email addresses from its training data [21], and GPT-3-based Copilot exposed sensitive API keys, potentially enabling unauthorized database access [56]. Besides exposing real sensitive data, LLMs also risk spreading false or misleading information [109].

Furthermore, LLMs and text-to-image (TTI) models raise copyright and cybersecurity concerns [15, 109]. LLMs may generate content that exploits creators' ideas without direct copyright violations [15]. TTI models like DALL-E can mimic artists' styles and be exploited for financial gain [15, 100]. Both technologies also pose cybersecurity risks, generating personalized phishing emails or convincing visuals for scams [15, 109].

In response to privacy risks, research has developed various safeguards for LLMs. For example, pre-processing techniques filter or replace sensitive information, such as personal identifiers (e.g., names, addresses), before it reaches the model [48]. Differential privacy adds noise to data or model updates, preventing the model from memorizing and leaking specific user data while maintaining overall performance [96]. Federated learning enhances privacy by decentralizing model training, minimizing data exposure by keeping information local [117].

While this research stream has uncovered privacy issues and technical solutions for LLMs, human aspects have received less attention, as we discuss next.

3.2 Privacy Theory and Human-Centered Privacy in LLMs

Privacy can be understood as a predisposition or behavior akin to a boundary regulation process of self-disclosure, which individuals employ to achieve an optimal level of social interaction in accordance with contextual norms [72] and personal needs [7] [98]. It encompasses granular constructs such as privacy concerns (i.e., worries about specific privacy-related situations), privacy behaviors (i.e., actions taken to achieve a preferred level of privacy), and privacy preferences (i.e., desired outcomes in privacy-related situations) [27]. These constructs often overlap and collectively shape an individual's overall perspective on privacy [27].

There are several theories used to study user privacy perceptions and behavior. As such, the privacy calculus theory explains that users weigh the perceived benefits (e.g., discounts) against the risks (e.g., identity theft) when deciding to share personal information [29, 38]. This theory, however, has faced criticism for assuming rational decision-making, as users often rely on heuristics and cognitive biases [3, 52]. Moreover, emotional responses play a role in privacy decisions [25]. Drawing on coping theory [58, 59], besides behavioral strategies to manage privacy risks, users employ emotion-focused coping (e.g., complaining about privacy risks) [25]. Emotional factors like frustration or anxiety can lead users to actively manage their privacy settings or disengage entirely, a phenomenon known as privacy fatigue [25, 26].

Human-centered privacy in LLMs focuses on understanding how users perceive and respond to privacy risks, but related research remains limited. Few studies highlight user concerns about being listened to and data collection when using intelligent personal assistants [65]. A recent study on ChatGPT revealed that users adopt privacy practices such as avoiding certain tasks (e.g., financial advice), providing generalized or false information, or removing personal details to protect their privacy [115]. In broader digital contexts, users have historically had limited control over privacy and often resort to self-regulation by selectively sharing or withholding personal information [68, 74]. Some users restrict their use of specific services, such as limiting single sign-on (SSO) to low-risk platforms or avoiding high-risk services entirely due to privacy concerns [13, 24].

LLM privacy is compounded by the anthropomorphic nature of chatbot interactions. Human-like qualities often lead users to overshare, as users overestimate the chatbot's capabilities and perceive it as human-like [85, 109]. This tendency increases privacy risks and tolerance for intrusions, even when users are aware that chatbots are not human. Such behaviors can be exploited maliciously, as in the case of promoting addictive content [44, 87, 105].

The literature reviewed above highlights common concerns surrounding LLMs. However, as emerging platforms facilitate the easy

creation and use of LLM-based apps, such as GPTs, research on their privacy implications remains limited and requires continuous reassessment and updates. Building on the foundation of prior privacy theories and LLM research, this work examines GPTs through a human-centered lens.

3.3 Privacy Across Different Roles

Perspectives on privacy issues and mental models around data flows may differ, depending on several determined personal influences [12, 63, 98, 110, 112] as well as the role individuals hold in a given context [72]. Extensive studies have shown that the level of expertise of users impacts mental models around privacy and security [36, 49, 73]. In particular, more technically advanced users have a more sophisticated understanding of system data flows and are more aware of potential cybersecurity threats [36, 49], a trend also observed in the context of LLMs [115]. This difference in user mental models also emerged as part of privacy concerns. Users with a less accurate mental model, viewing their input as a quality indicator rather than training data, have difficulties perceiving the possibility of memorization leaks [115]. Additionally, role perception influences individual privacy concerns. Studies reveal differences in privacy attitudes between professional and hobbyist developers. While professional developers tend to approach privacy with a compliance-driven mindset, adhering to regulations, hobbyists often view privacy policies as platform-imposed requirements [94]. Furthermore, while users express more concerns about data sharing and handling [65], developers sometimes engage in risky practices, such as neglecting security implications [1], requesting excessive permissions [31, 32, 102], and violating privacy policies [8, 97].

Despite the studies above, the dual roles of individuals as both users and creators and the impact this overlap may have on their privacy perceptions and practices remain underexplored to the best of our knowledge. In the GPT store, individuals can easily create while using others' GPTs or their own. From OpenAI's perspective, both creators and users have the same type of account and are treated equally as users. By examining the intersection of these roles, we aim to understand how roles influence privacy perceptions and practices.

4 METHOD

To understand privacy issues in GPT usage, we conducted 23 semi-structured interviews from June to July 2024. The interviews were conducted remotely through Microsoft Teams, ranging between 40 to 126 minutes, with an average of 69 minutes. Each participant received a 20-euro Amazon voucher as compensation for their time. This study followed university guidelines and was approved by the university's institutional data protection board and ethical committee.

4.1 Participants Recruitment

To select participants, we conducted an online screening survey that included questions about basic demographic information and respondents' usage and creation of GPTs. To prevent fraudulent entries, as advised by Panicker et al. [86], we included a basic knowledge test about GPTs, and only those who passed were eligible for recruitment. The full screening survey can be found in the Supplementary Materials.

To capture diverse privacy perspectives from various roles, we targeted three groups of GPT users in our recruitment: a) End users: individuals who regularly use GPTs. b) End users and amateur creators: individuals who regularly use GPTs and have experimented with creating them. c) Professional creators: individuals who have created GPTs for public use. To reach a broad audience, we posted our screening survey via online platforms like the OpenAI forum, Reddit, LinkedIn, and Prolific. To target professional creators, we randomly selected 300 GPT creators with public LinkedIn contact information from the GPT store and sent them individual study invitations and screening surveys.

We received a total of 151 responses to the screening survey. After excluding 46 invalid responses (ambiguous or fraudulent) and 28 respondents who declined participation, we reviewed the remaining 77 responses. Participants were chosen to represent different types of GPT usage and varying levels of creation experience. In total, 43 invitations were sent. Each invitation included a study information sheet, privacy notice, and consent form.

Interviews were primarily conducted in English, with two exceptions (P4 and P11), who were assisted by two experienced HCI researchers from the team who natively speak German and Chinese, respectively, per the participants' request. During the interviews, both participants predominantly communicated in English, switching to their native languages only when unable to articulate specific ideas. In such instances, the researchers restated the participants' statements in English and obtained on-site confirmation from participants to ensure accuracy. The interviews were recorded and transcribed using Teams' automatic transcription feature, which helped minimize human-introduced bias. Researchers subsequently reviewed and corrected the transcripts to ensure accuracy and clarity while preserving the participants' statements.

During the interview period, researchers continuously reflected on the data and engaged in discussions with all authors. When data saturation was reached and no new information emerged, we stopped recruitment [35]. Ultimately, a total of 23 participants consented and completed the interviews. Their demographics are detailed in Section 5.1, Table 1. We will discuss the participants' demographics alongside our findings.

4.2 Interview Design

The interview design was iteratively developed and reviewed by a team of professional HCI, privacy, and security researchers. Before beginning the actual data collection, we conducted three pilot interviews with people with and without GPT creation experience, after which the interview protocol was refined.

At the start of the interview, we clarified that the interview would focus exclusively on GPT usage. Participants were asked to explain their understanding of GPTs and distinguish them from other AI products (e.g., ChatGPT) to ensure a clear understanding of the scope of GPT usage.

Our semi-structured interview protocol comprised two major sections. The first section focused on participants' perspectives as *users* of GPTs, covering the following topics: general GPT usage (Q1.1), experiences with action-based and login-based GPTs

(Q1.2), mental models of data flows in three GPT usage scenarios as described in section 2 to elicit further discussion on participants' privacy concerns (Q1.3), and privacy concerns, risk perceptions, and responses to their privacy concerns (Q1.4). If participants only had user experience and no creation experience, the interview concluded at this point.

For participants with experience in *creating* GPTs, the interview continued to the second section, which focused on the creator's perspective. We explored their creation experiences (Q2.1), privacy considerations as both creators and end-users of their GPTs, and reflections on how their dual roles as users and creators shaped their privacy perceptions (Q2.2). The complete interview protocol is provided in the Supplementary Materials.

4.3 Data Analysis

We conducted a thematic analysis [18] with Atlas.ti [10] through a hybrid approach, which combined inductive coding to uncover privacy-related perceptions emerging from the data with deductive alignment to established privacy constructs and concepts. This method allowed us to anchor our analysis in participants' lived experiences while contextualizing the findings within established theoretical frameworks.

To begin the analysis and become familiarized with the data [18], two authors manually reviewed and corrected the transcriptions auto-generated by Teams. Next, two coders independently open-coded two interview transcripts and reviewed each other's codes. Through discussion, they developed an initial codebook, merging similar codes and refining unclear ones. Using this initial codebook, the coders independently coded another transcript and calculated the interrater reliability (IRR), achieving an initial Krippendorff's alpha score of 0.67 [55, 57]. After discussing disagreements and refining the codebook by adding new codes, they repeated the process with another transcript, achieving an IRR score of 0.81, which indicates substantial agreement [57]. This process aimed to establish consistency in coding styles, including wording and granularity, while maintaining room for diverse perspectives from multiple coders [66]. The coders then divided the remaining transcripts and coded them independently, applying existing codes while staying open to new ones. Regular meetings were held to review and refine the codes collaboratively and to develop themes aligned with the research questions.

Theoretical Grounding and Operationalization in Analysis. To explore participants' mental models of data flow within GPTs (RQ1) and how user-creator roles influence privacy perceptions (RQ3), we inductively analyzed how participants (mis)understood various aspects of data flow, including the range of data collected, the entities involved, and their purposes for collecting the data, as well as the specific ways in which user-creator roles impact privacy considerations.

To analyze privacy concerns and practices (RQ2), we focused on two groups of privacy constructs:

(a) *Privacy concerns* are defined as "expressions of worry towards a specific privacy-related situation." [27]. During the analysis of privacy concerns, initial coding was conducted inductively. Codes were assigned to recurring ideas or issues without reference to a predefined framework. As researchers developed broader themes, it became evident that many concerns can be attributed to Solove's privacy taxonomy, which categorizes data stages into collection, processing, and dissemination [99]. Therefore, we used it as a lens to refine and organize subsets of the themes in privacy concerns.

(b) *Privacy practices* comprise all forms of participants' enacted responses to their privacy concerns. During the analysis, related codes included privacy-related decisions, behaviors, and preferences, as such constructs often overlap [27]. Focusing on an overview of privacy practices enables us to capture the variety of behavioral and emotional reactions, and how participants navigate their privacy concerns.

5 FINDINGS

We present our qualitative results, grouped into three sections corresponding to our research questions. Section 5.2 describes users' mental models of data flow within GPTs. Section 5.3 and Section 5.4 present users' and creators' privacy concerns and privacy practices in GPT usage. Section 5.5 reveals how the role of users or creators shapes privacy perceptions.

Through our interviews, we learned of various usage motivations among different user groups of the GPT store. Before presenting the findings for each RQ, in Section 5.1 we begin by providing an overview of our participants and outline the user-creator spectrum we discovered based on these usage motivations.

5.1 Participants and the User-Creator Spectrum

As described in Section 4.1, our study aims to understand the privacy perceptions of individuals in different roles. We targeted our recruitment in terms of three categories of GPT users: a) end-users; b) end-users and amateur creators; and c) professional creators. However, as we progressed with our interviews, we learned that the distinction between users and creators is not as discreet (for instance, based on the number of GPTs created) as we initially hypothesized; very often, it was blurred. Even professional creators with popular apps noted this complexity:

> *"I feel like in this GPT store, I'm creating my own GPTs and I'm using them. I feel like I create them, but I don't really create them, because it's so easy to create, it's not like the Apple store or Android store where you may need more effort and longer cycles to release it. So in this sense, the line between users and creators is really blurry."* - P8

Overall, we observed a spectrum of creation experiences among our participants. Some (P18–P23) had no prior experience with GPT creation. Others who identified as GPT users during the screening survey exhibited varying levels of expertise—some (P10, P12, P14–P16) had contributed to one or two GPTs, while others (P11, P13, P17) had created multiple GPTs, demonstrating more extensive experience. In contrast, participants who identified as professional creators (P1–P9) generally had substantial experience. They developed GPTs for public use, with many achieving notable success, with their most popular GPTs exceeding 500 conversations.

Through our interviews, we discovered that self-reported users had primarily created GPTs for personal use, despite having experimented with publishing them (P10–P17). For example, P13 created multiple GPTs to help him study for different exam subjects by

No	Gender	Occupation	Country	Usage and Creation (conversation counts)
P1	Male	Independent creative director	Netherlands	user, popular GPTs creator with 10+ GPTs (600+)
P2	Male	Consultant	Spain	user, popular GPTs creator with 10+ GPTs (1K+)
P3	Male	Tech founder, novelist	Singapore	user, popular GPTs creator with 5 + GPTs (100+)
P4	Male	Entrepreneur	Germany	user, popular GPTs creator with 1 GPT (700+)
P5	Male	Self-employed IT expert	Luxembourg	user, popular GPTs creator with 4 GPTs (500+)
P6	Male	Talent consultant, content creator	USA	user, popular GPTs creator with 3 GPTs (1k+)
P7	Male	Customer Experience Lead	Spain	user, popular GPTs creator with 5+ GPTs (1k+)
P8	Male	Lead Salesforce DevOps engineer	USA	user, popular GPTs creator with 3 GPTs (5k+)
P9	Male	Senior account manager	USA	user, popular GPTs creator with 200+ GPTs (10k+)
P10	Female	Personal assistant	UK	user, involved in a team creation
P11	Male	System security researcher	China	user, created 3 GPTs for self-use
P12	Female	Researcher, consultant	UK	user, experimented creation
P13	Male	Student	Finland	user, created >20 GPTs for self-use
P14	Male	Doctoral researcher	China	user, created 2 GPTs for self-use
P15	Female	New graduate	Netherlands	user, created 1 GPT for self-use
P16	Male	Research assistant	Finland	user, created 1 GPT for project
P17	Female	Student researcher	Finland	user, created several on similar platform
P18	Female	Data intern	Finland	user
P19	Male	Teacher/volunteer	UK	user
P20	Female	Doctoral researcher	Finland	user
P21	Male	Research assistant	UAE	user
P22	Female	User experience designer	Finland	user, plans to experiment with creation
P23	Male	Researcher	Germany	user

Table 1: Demographic characteristics of 23 study participants.
Note: a) The conversation count in the last column refers to the total number of conversations that have occurred within the creator's most popular GPT. b) The OpenAI product is not officially available in China, and P11 and P14 used a VPN to access the services. c) P17 created apps on a similar platform (https://miniapps.ai/).

fine-tuning each one with relevant course materials. Similarly, P11 developed a GPT for language translation and refining writing tasks to match a specific style he was working on. However, professional creators (P1–P9) were more motivated to create for the public, developing GPTs with such goals as self-promotion, sharing knowledge with the community, or showcasing their expertise as part of a business model where they create and customize GPTs for clients. For example, P8 utilized his expertise in Salesforce to create GPTs that translate user needs into Salesforce-specific language, gaining popularity within the Salesforce community. Similarly, P3 leveraged his experience as a novelist to develop GPTs for editing and publishing news media editorials. By showcasing these capabilities, he actively sought clients and helped them to create their own marketing-focused GPTs.

In summary, the participants in our study had varying levels of experience with creation. Some shifted between dual roles, acting as both creators and users, engaging in the creation of GPTs while simultaneously using GPTs they or others had created. Based on these insights, we refined our mapping of participants into the three main groups detailed in Table 1.

Next, we will report our findings from discussions with these participants, covering the spectrum of various users and creators. In cases where clear distinctions in perceptions between creator and user roles are evident, we will highlight their corresponding roles. However, in situations where the role is less significant—particularly when discussing the use of GPTs—we will refer to all participants as "users of GPTs" and present privacy perceptions more generally.

5.2 Mental Model of Data Flow in GPTs Usage

In this section, we outline the most prominent mental models for each scenario of GPT usage, highlighting key conceptions and misconceptions about GPT data flow in comparison to the official documentation.

We find that participants with experience creating GPTs expressed greater confidence in their assumptions about data flow. In contrast, those who were solely users frequently responded with statements like *"I don't know."* Furthermore, participants generally exhibited skepticism about actual data flow processes, often expressing *"hope"* that data practices functioned in specific ways. This blend of uncertainty and cautious optimism highlights a lack of clarity and trust in the underlying data flow in GPTs.

5.2.1 Scenario 1: Basic GPTs. Participants were asked to recall an instance where they used or encountered a GPT that offered only basic interactions. We then asked participants to briefly describe the GPTs they used, to explain how they believed the data flowed within

those GPTs. The primary divergence in participants' understanding centered on who collects the data and their purposes.

Only OpenAI Collects Data to Improve its Ecosystem. (P2–5, P7–9, P11, P13–14, P16, P18, P23) Participants with this mental model believed that only OpenAI collected user data, including chats and interactions, primarily for training backend models. They saw data collection as a collective process aimed at system improvement, rather than targeting individual users. Some participants viewed it as reasonable for enhancing the model's performance. As P11 described:

> *"I think in a company or a service with such a large number of users, there are many people talking to it every day, and it shouldn't be, how to say, use my information specifically to me. It may be more about summarizing some of my patterns and then using them in their subsequent training. So I think this is also a process of improving, so it's okay to let them collect this information."* - P11

Participants also recognized GPTs as part of OpenAI's broader ecosystem, suggesting additional data collection purposes beyond training. These included collecting user feedback by tracking interactions across services, *"If people are like deciding to use one GPT and then they go to use a similar one, [OpenAI] tracks the interactions across GPTs"* (P8), and using chat history to monitor violations, such as *"violence, porn, or other things,"* to support policy development (P14).

Overall, some participants with this mental model trusted OpenAI to handle data responsibly and exclusively. As P2 commented, *"There's no security problem because it's only OpenAI who is collecting data."* (P2).

Official documentation: OpenAI collects user content (e.g., chats, uploaded files) and usage data (e.g., interaction activity, location, and device details) to provide and improve its services, and also prevent misuse of them [79, 81].

Creators: Collects vs. No Access to Data. Some participants believed that creators of specific GPTs, alongside OpenAI, collected user data such as conversations and behavioral information to improve their GPTs (P10, P12, P15, P17, P19–20, P22). For example, P15 used a GPT to help generate a CV. She explained that the detailed information she provided, such as *"my city, education, work experience"* (P15), would be collected by the creators because *"they try to know about their users"* (P15).

However, participants with experience in GPT creation opposed the idea of creators being involved in data collection. They explicitly stated, based on their experience, that creators cannot access any user interactions with their GPTs (P1–2, P4, P7–9, P14). Regarding the inability to access user data, creators expressed mixed attitudes. On one hand, some creators wished for access to user data to improve their GPTs. For instance, P2 commented:

> *"Unfortunately we don't collect the information. It would be an interesting incentive that OpenAI would release that data collection to GPT creators. [...] If I could gather all the messages, I could improve my own chatbot because I would adapt to the feedback from users, not feedback that they give directly to me, but things that I observe."* - P2

On the other hand, some creators appreciated that user privacy is protected by restricting creators' access to user data. P1 explained:

> *"I'm not allowed to see your conversation, which is a good thing, right? Never share it with creators. That will be awful."* - P1

Official documentation: In Scenario 1, creators do not have access to specific conversations with their GPTs [82].

Partner Companies Shares Data. (P1, P6, P21) A few participants also believed that extensive user data could be shared with partner companies and used for commercial purposes. For example, P21 speculated that chat history might influence market campaigns based on user queries, while P1 imagined detailed user profiling, including tracking other desktop screen activities: *"It will scan everything. Maybe it will scan how I organize my folders. Maybe the colors I like on my desktop, and maybe it will analyze even the sounds if I have a cat or not, or how I breathe."* (P1). This invasive and commercially oriented data flow perception became more prominent in Scenario 2.

Official documentation: OpenAI may disclose users' personal data to vendors and service providers to support business operations and deliver certain services [79].

5.2.2 Scenario 2: Action-based GPTs. Participants were asked to recall an instance where they used or encountered a GPT that required user permission to perform actions outside the GPT. For example, P13 described using a GPT to design graphics and create slides, which connected to the third-party service Canva [19]. In this scenario, participants generally believed that more entities in addition to OpenAI were involved in collecting data, often in a more invasive manner.

Third Parties Collect Data. (P2–3, P5–6, P9–11, P14, P16–18, P23) With this mental model, participants believed that data was collected by multiple entities, including creators and third-party services. However, most participants were uncertain about what data was being collected.

Some participants believed that only selective data, such as task-relevant keywords, would be collected. P14 explained:

> *"Like the scholar GPT the only thing we transfer to that API may be the title of the article, and the summary of the article, but no more personal details, even if I accidentally put my name or my school name to the GPT."* (P14)

In contrast, some participants assumed that third parties might collect everything beyond the conversations, including credentials, user IDs, emails, and all interaction data. In this scenario, many participants highlighted the possibilities for third parties to collect user data for spam content or marketing. For example, P14 assumed that third parties collected data to predict market trends: *"The tendency of the people travel if I operate a tourism GPT that help me help people book flights or hotels, they may use such information to estimate the travel tendency of the market"* (P14).

Official documentation: Builders of GPTs can specify the APIs to be called, but users must consent to actions. OpenAI does

not independently verify the privacy or security practices of API providers. Users are advised to use the GPT only if they trust it [77, 81].

5.2.3 Scenario 3: Login GPTs. We asked participants to recall an instance where they used or encountered a GPT that requires users to log in to external services. Compared to Scenario 2, participants showed a more cautious attitude toward using login-based GPTs, expressing doubts about the underlying processes. Some found these interactions confusing, such as P17, who used a GPT to create an event invitation but unexpectedly had it synchronize with her calendar: *"I just clicked by mistake and it suddenly opened my calendar and I was not sure why did that happen"* (P17).

Authentication Secured by OpenAI vs. Creators Having the Password. Most participants believed the login process was securely managed through authentication, ensuring that no one could access their direct account information, such as usernames and passwords (P3–6, P8, P11, P16–17, P20–21). Still, some participants remained concerned that those direct account details, such as email usernames (P14) or even passwords (P7, P15, P18), were still accessible to GPT creators or third-party companies.

Many felt that OpenAI should bear full responsibility for managing all authentication information. However, some also recognized that the security of the login process depended on the practices of the GPT creators. As P5 commented, *"At least if the developer has done the job correctly, they shouldn't have access to the username or password"* (P5)

When discussing login status, participants often referred to their experiences with other services, assuming GPTs would remain connected unless users manually revoked access. Meanwhile, some participants expressed uncertainty but remained hopeful, as P15 noted:

> *"I hope they can automatically log out of my account after usage. Because if you don't, you will lose the trust of users. They might stop using GPTs, thinking, "Oh, someone else could be using my account." "* - P15

Official documentation: Login in GPTs can be implemented through "API Key" or "OAuth." OpenAI encrypts the API key and client secret at rest. The OAuth key is refreshed periodically [80].

5.3 Privacy Concerns of Using Custom GPTs

We analyzed participants' privacy concerns with GPT usage, focusing on their expressions of worry about specific privacy-related situations. The development of themes was informed by Solove's privacy taxonomy of data collection, processing, and dissemination [99]. Findings are summarized in Table 2. Concerns UC1 to UC3 (User Concerns) reflect different stages of data collection, processing, and dissemination, while UC4 presents concerns arising from a lack of regulatory guidance. Additionally, CC1 (Creator Concern) highlights the challenges uniquely faced by creators.

5.3.1 UC1: Concerns about Data Collection. Our participants noted that using GPTs often involved sharing sensitive personal information during interactions (P3, P6, P8–9, P15, P17, P21, P23). As P15 explained when using a GPT to create a CV:

Privacy Concerns about GPTs		
Participant	Privacy Concerns	Brief Definition
All	UC1: Concerns about data collection	Transparency, consent, and scope of information gathering
All	UC2: Concern about data processing	Misuse, inaccuracy, or insecure processing of personal data
All	UC3: Concern about dissemination	Unauthorized exposure of information or intrusion into private life
All	UC4: Lack of privacy regulatory guidelines	Insufficient regulations, platform guidelines, and GPT verification
Creators	CC1: Concerns about creator's knowledge	GPT creators' work exploited through reverse engineering

Table 2: An overview of both User Concerns (UC) and Creator Concerns (CC) in the privacy of GPTs.

> *"Instead of using it as a general search engine, for GPT I tend to provide more specific information of the tasks that I'm performing. Probably I'm very into performing the tasks, so in the moment I have less attention on what kind of information that I'm going to give, a lot of personal information as well like my personality, my way of thinking."* - P15

The tendency to overshare raised concerns among participants about data collection, especially with GPTs that connected with external services and third parties (P1, P3–4, P6, P8, P10–12, P15, P17, P20–23). They felt anxious about more invasive data collection when connected to external services, such as *"access to some sort of like maybe the microphone or if you've opened an app that uses your camera, it wouldn't be limited to the GPT space"* (P23), fearing that this data could be collected by multiple entities beyond OpenAI's infrastructure. This concern made some participants hesitant to use GPTs with integrated actions.

Concerns about consent in data collection were also raised. Some participants were skeptical about multiple parties collecting data without their permission (P6). Additionally, one participant highlighted concerns that external parties' data collection practices were obscured under the guise of OpenAI, potentially misleading users into misplaced trust and unknowingly giving consent.

> *"In that case [a GPT with actions], creators have their own website with a GPT interface and people would typically need to accept the terms. And, of course, creators can gather the data, which a lot of creators would do, but it has a problem with trust. The store is like a screen that I believe might make some people feel more comfortable trusting. But you know, in the end, users are bypassing the fact of connecting with external services by accepting the terms of conditions and data policies, and acknowledging who is gathering the data and all of that, so that is a bad thing about the GPT store."* - P2

5.3.2 UC2: Concerns about Data Processing. As GPTs are highly specialized for specific tasks, participants expressed concerns about the aggregation of usage context and interaction data to create detailed user profiles. For example, P12, who used a lawyer GPT to consult a rental contract issue, noted, *"I assume it's already assuming*

at the beginning that you need legal services. So, there's some default persona of you already there" (P12). This raised concerns about users becoming clear targets for secondary purposes, such as targeted marketing (P12, P15–16, P20, P23). P20 added:

> *"There's this custom GPT app that is about relationship advice. If you're asking one time about the problem that you have or counseling, I think they're forming these profiles about people. And if you already logged in with your social media account, they know who you are. These profiles are very useful because they are the gate to proper marketing. The more crisp they have, the information about you, the more money they're gonna make."* - P20

At the same time, participants expressed insecurities about data processing, particularly concerns regarding potential data breaches and leaks. These apprehensions were heightened in relation to GPTs with actions and those developed by individual creators. For example, P17 highlighted that the involvement of third parties exacerbates vulnerabilities: *"The more parties are involved, the more hackable the data is; it gives chances for hackers to do eavesdropping or a man-in-the-middle attack"* (P17). Furthermore, participants feared that inexperienced creators might mishandle their data (P10, P14). P14 elaborated:

> *"Some part-time developers, like me or other people, may not know much about data protection or the regulations of each area. So they may have unintentionally or intentionally leaked some data to other people."* - P14

Adding to these concerns, participants highlighted their limited control over data, stressing that data processing in GPTs is irreversible once integrated into a model (P1–4, P6–7, P13–15, P18, P20, P22–23). As P15 explained:

> *"I have no way to delete it or to check how my data is stored, how they're gonna use my data. [...] Once the data is used to build a model, it doesn't even matter. The original data is gone, because it's built into the model."* - P15

5.3.3 UC3: Concerns about Data Dissemination. Participants raised concerns about the dissemination of confidential data when using GPTs, fearing their conversations could be inadvertently exposed to other users asking similar questions (P3, P6, P8, P10, P12, P14, P20–21, P23). For example, P2 highlighted the risk of a confidential research idea being inadvertently revealed. Conversely, participants were also uneasy about receiving sensitive information from other users in GPT responses, which they found equally undesirable (P1, P14, P20).

Participants emphasized the potential harm of data dissemination if their information were leaked (P1–2, P10, P17, P21). For example, P8 envisioned scenarios where a GPT connected with payment services could lead to *"a lot of bad things"*, P8 explained:

> *"If all my data that I'm giving was exposed, especially if I've been connecting to other services... Identity theft, taking information from those services, can pretty much like destroy somebody's life effectively."* - P8

Participants also expressed concerns about malicious intents, such as their data being sold, exploited for scams, manipulated for political purposes, or in cases involving images or visual content, altered into deepfakes (P1–2, P10, P17, P21).

Participants were further concerned about the potential for GPTs to intrude into their personal lives by bypassing permissions and acting without explicit consent (P1, P6, P12, P17). They feared GPTs could disseminate private data through unintended actions, such as sending *"tweets under their name"* (P12) or misusing linked accounts (P6).

5.3.4 UC4: Concerns about Lack of Privacy Regulation and Guidelines. Participants, particularly those with experience in creating GPTs, expressed concerns about the lack of regulation in the GPT store. They observed that low-quality GPTs and spam content had proliferated, undermining trust in the platform. As P2 explained about the GPT store:

> *"It's been contaminated with a lot of bad content. Sometimes the conversations about privacy are totally out of touch with reality with these kinds of companies. They don't care about people trying to deceive other people or trying to sell a service, which is actually what malicious players could do. I think that one bad thing about OpenAI and the GPT store is that it promotes these behaviors like spam, scams, and all of that."* - P2

P2 further attributed the issue of spam to the low barrier for creating GPTs: *"It's extremely easy to create that GPT, so that's the reason why it's used for spam essentially. [...] OpenAI could improve to create more barriers for creating GPTs, perhaps more requirements for identifying the creators"* (P2). Similarly, P8 expressed concerns about the lack of mechanisms to report malicious GPTs and called for more accountability and verification guidelines. He suggested features like a *"OpenAI verified badge"* and ensuring the verification of creator identities, stating, *"This developer is either being who they say they are or being a certain level of security"* (P8).

Moreover, participants expressed concerns about the absence of regulatory bodies, such as government and data protection organizations to regulate the GPT market (P3, P6, P20). They noted that existing regulations were outdated and not keeping pace with rapid advancements in AI products like GPTs. Related to this concern, the European General Data Protection Regulation (GDPR) was frequently mentioned. Participants who resided in Europe felt protected under GDPR (P1, P4–5, P16, P20), while those outside Europe expressed a desire for similar laws in their regions (P3, P14). These regional differences also raised concerns among participants about how GPTs could be regulated when they use a chain of services located in various parts of the world.

5.3.5 CC1: Concerns about Creators' Knowledge. Apart from concerns of regular users, there was a special concern shared by creators. Creator participants shared that creating high-performing GPTs required significant effort, including prompt iteration and incorporating their personal knowledge through uploaded text files (P1, P3, P5–6, P8, P11). However, once a GPT was published, this carefully curated work could be easily exploited by end-users through specific prompts.

> *"Certain prompting approaches can get the large language models to output the data sources pretty much verbatim, which means that any data that is taken from me and used to build the next model, no matter how private that information, I think it is potentially could be reverse-engineered from specific prompts."* - P3

Creators expressed frustration with the lack of safeguards for their creation, leaving them concerned about the potential misuse of their work.

> *"I'm more worried as a creator, someone taking my instructions and claiming it as theirs maybe, you know, or using it just as if I want. I feel like I have the right to that prompt because I'm a creator of that prompt. It's almost as if you're stealing my profile."* - P9

Furthermore, creators highlighted the lack of protections from OpenAI, which left them feeling exposed (P1, P3, P6, P9).

> *"It's really about the privacy of the bot itself. What I found very weird and not good from OpenAI is that they don't protect the bots. Because I have a whole list of data I can share on how to breach GPTs in an instant. I can access sometimes attached files if I want, and read all the instructions, so I can recreate the bot myself and start making money there. I find it mind-blowing that you [OpenAI] offer people bots. Maybe you can earn money in a store, and then you take zero responsibility in protecting those bots."* - P1

5.4 How People Respond to Their Privacy Concerns

In this section, we present findings on how participants respond to their perceived privacy concerns, emphasizing how users navigate and address their privacy concerns. These practices are summarized in Table 3. UP1 to UP4 (User Practices) represent actions adopted from a user's perspective, while CP1 and CP2 (Creator Practices) detail practices specific to creators.

5.4.1 UP1: Self-Censorship of the Input. Many participants shared that the most effective way to protect their privacy was to be cautious about what they shared with GPTs from the beginning, particularly data that they considered highly sensitive or valuable (P1, P3–7, P10–12, P14–15, P17, P19, P21, P23). They believed that once such information was shared, it could be beyond their control. As a result, they viewed proactive management of input data as the best privacy practice.

> *"There's this very old thing that like grandma used to say, it's like "you wouldn't want to say something that you don't want in the newspaper." Put it like that. Like, if you have that kind of attitude, that's the same thing, if you don't want it in public, you don't want it in your GPT."* - P6

Some participants also employed strategies to pseudonymize their input, such as replacing key personal details like real names with placeholders or less identifiable information.

This self-censorship also applied to logging into other service accounts through GPTs. Participants were reluctant to connect accounts they considered valuable or sensitive, often choosing to avoid such connections entirely (P1, P3–5, P8, P11–15, P19). However, comfort levels with connecting different types of accounts varied across participants. Almost all participants avoided linking accounts involving financial transactions, but opinions diverged on other types of accounts. For example, some participants were comfortable linking social media or entertainment accounts like Spotify, while others, such as P5, viewed music preferences as personal information and avoided linking such accounts. Similarly, P11 considered their GitHub account to be sensitive and avoided associating it with GPTs.

Privacy Practices to GPTs		
Participants	Privacy Decisions and Behavior	Brief Definition
All	UP1: Self-censorship of the input	Proactive efforts to reduce the amount of personal information shared
All	UP2: GPT evaluation	Users' continuous evaluation of GPTs for privacy and trustworthiness
All	UP3: Minimizing traces of GPT usage	The deliberate separation, deletion, and obfuscation of GPT activities
All	UP4: Accepting privacy risks for features	The compromises users make when balancing privacy concerns with utility
Creators	CP1: Knowledge protection	Actions to protect the creation of knowledge, including configuring settings
Creators	CP2: Creating privacy notices for GPTs	Practices towards protecting other users' (clients, end-users) privacy

Table 3: An overview of both User Practices (UP) and Creator Practices (CP) regarding the privacy of GPTs.

5.4.2 UP2: GPTs Evaluation. Participants carefully evaluated GPTs' trustworthiness before using them. Many expressed greater trust in GPTs developed by OpenAI or familiar services they had used previously (P2–4, P7–10, P11–14, P17, P19–23), such as Canvas (P23) or Consensus (P13). Beyond GPTs created by established organizations, participants were more cautious.

To evaluate the trustworthiness of GPTs, some participants consulted reviews from platforms like YouTube or Twitter to identify reliable GPTs (P10, P19). One participant referenced a repository for a collection of open-sourced GPTs, which they reviewed for effectiveness and credibility (P11). Ratings and reviews in GPT stores were also common indicators for assessing reliability. As P13 explained:

> *"You know they have these weekly top GPTs, so I go to explore GPTs, and under each of these headings, like featured, you have the top four. […] I also searched for it like with some keywords and click based on the review and also the number of conversations. So if it's more than 5K plus or from 1K plus, I go for it."* - P13

This evaluation process did not stop at the point of selection but continued during the usage of GPTs, especially for those with actions. Participants closely monitored the actions requested by the GPTs (P3, P10, P12, P15, P17). They preferred to approve actions

on a case-by-case basis, requiring explicit permission each time a GPT needed to interact with external services. Additionally, participants kept a vigilant eye on whether the GPT's actions aligned logically with the context of their specific use cases, ensuring that the interactions remained appropriate and trustworthy.

> *"There have been instances where I launched a custom GPT that wasn't built by me. I asked for something relatively straightforward, and it immediately asked me for some external website or something like that, and because it didn't match my expectation for that particular tool, I would have just closed it."* - P3

5.4.3 UP3: Minimizing Traces of GPT Usage. Participants shared that to minimize the risks associated with their personal information, they adopted various actions to reduce their digital footprint and manage how their data could be traced back to their personal identity. These practices included compartmentalizing digital activities, deleting conversation histories, and obfuscating sensitive information.

Many participants described separating workspaces as a strategy for minimizing traces of their GPT interactions (P2, P6, P10, P12, P14, P17–18, P20–21). To protect their primary digital identity and reduce the risk of data linking, participants isolated sensitive activities from their main accounts, often using separate email addresses or creating dedicated wallets for GPT-related tasks.

Additionally, some participants reported deleting their conversation history after using GPTs (P1, P7, P11, P17, P19, P21–22). While a few acknowledged that this practice served more as psychological reassurance than a belief in actual data deletion. P12 explained:

> *"I'm not sure if it actually makes sense because actually if I have talked with it, it's like the history is already there. Delete, it's just for myself, I feel psychologically safer, my data is not at least present in front of me."* - P12

A small number of participants mentioned strategies to maintain control over their digital traces, such as regularly asking what GPTs *"know"* about them (P17) or disguising their digital identity by intentionally adding irrelevant information after discussing sensitive topics (P22). For example, P22 described that after consulting GPTs about medical conditions, she would introduce unrelated topics to dilute the context of the interaction, thinking it would prevent specific details from being traced back to her.

5.4.4 UP4: Accepting Privacy Risks for Features. Apart from actively preventing privacy risks, some participants shared that they viewed GPTs usage as a tradeoff, choosing not to worry about privacy in exchange for good services and often skipping privacy documents altogether (P1, P3, P6–8, P15–16). As P7 put it, *"I know that if I'm using this kind of service, the data is going to be shared, and you are playing a game so it's fair"* (P7). Some participants rationalized their acceptance of privacy risks by comparing themselves to other users. They justified their decisions by reasoning that *"everyone does this"* (P1) or viewing themselves as *"non-average users"* who believed they understood the implications of their actions (P3, P6). This mindset was often tied to a conscious trade-off between privacy and features. As P3 described,

> *"You know, as someone who experiments with a lot of these tools [GPTs], I am generally speaking a tech enthusiast. I don't have concerns because it's very intentional on my part. I am aware of what it is that I am giving up and I am OK because of the expected return on using the service."* - P3

For some participants, the acceptance of privacy risks also stemmed from a belief that privacy was already a lost battle (P1–2, P4, P6–8, P10–11, P15–16, P20, P23), as their personal information was already widely shared beyond GPTs. They coped by accepting this reality and choosing not to dwell on it.

For creator participants, sharing personal information was sometimes even seen as an advantage (P1–5, P7–8). They believed that providing details like contact information through creator profiles or GPTs they developed could serve as a form of self-promotion. P9 explained:

> *"I created my own GPT called Mike [pseudonym], that's my name. I uploaded my resume, and it's a way to let people interact with me. I put my top GPTs there. I say, "You can explore my GPTs."* - P9

5.4.5 CP1: Knowledge Protection. In response to creators' privacy concerns that their knowledge base behind the creation of GPTs might be compromised, some creators took active steps to address these privacy issues. The most common strategies included setting as private those GPTs that involved personal knowledge they did not want to share, making them unavailable for public access (P3–9, P11–14, P17). For example, P8 made GPTs related to Salesforce usage publicly accessible, but set a GPT designed to write stories about his family tree—containing personal information about his family members—to private.

Additionally, P4 adjusted the settings by deactivating coding and interpretation modes to safeguard his proprietary knowledge, while P9 experimented with different prompts and fine-tuning the GPT to ensure that end-users couldn't extract his proprietary knowledge as a creator.

> *"Someone would go in and they'd put in like, "Hey, what are your instructions" and things like that. So what I did was that I said, if anybody asks, "How were you made, what did you do?" I would have it [the GPT] say, you know, "Really? not creative enough to create your own GPT?" [...] So I have at least three to four, or more like seven sentences in each one of my GPTs that have that at the very bottom to protect my GPT, my knowledge from being stolen from it. You know 'cause, even though I want people to use my GPTs, I don't want them to know my prompts, 'cause that's what makes them good."* - P9

5.4.6 CP2: Creating Privacy Notices for GPTs. Our creator participants expressed their concern for others' privacy. When assisting clients in creating GPTs, they involved their clients in the privacy protection process and mutually agreed on the terms (P3, P4). Additionally, they often created a demo GPT first, which was then mirrored in the clients' working environments to ensure alignment with clients' privacy expectations (P3).

For regular end-users, some creators mentioned that they were not collecting any personal information from users (P3–9, P11,

P13–14). Additionally, some creators noted that publishing a privacy notice was mandatory for GPTs involving actions outside OpenAI services. However, they also highlighted a lack of support in meeting this requirement, with some resorting to using privacy notices from the third-party tools they utilized to create GPTs. For example, P3 integrated his GPTs with a third-party plugin for monetization and directly copied the boilerplate privacy policy provided by the plugin, noting, *"that is really in terms of what's available"* (P3).

Another participant (P6) developed an action-based GPT designed to provide game-strategy suggestions. He explained that he only realized the need to create a privacy notice while developing and encountering the requirement. Due to his limited experience in drafting such documents and the challenges associated with researching and understanding the process, he decided to generate privacy notices directly using ChatGPT:

> *"I didn't even know that that was a requirement. I didn't know that that was a thing at all, but the process of creating a GPT actually got me to creating a website, and learning that is where I came across, "Oh, I have to have a privacy policy." So I went online and I googled what it was, and as I'm googling it I'm squinting my eyes and going through all these search queries. I go, "Oh, well, nice, ChatGPT," and then I asked it and it did explain it to me very easily."* - P6

5.5 Role of the User-Creator Spectrum in Shaping Privacy Perceptions

This section presents findings regarding the spectrum of roles and its impact on privacy perceptions of GPTs.

5.5.1 Reflection with Different Perspectives Seems to Reduce Concerns. Participants who were both users and creators of GPTs shared that their dual roles enabled them to reflect with different perspectives. Understanding the process of creating GPTs made them less concerned about using ones created by others, as they knew that creators could not access their data as users in basic GPTs (P2–3, P5–6, P12, P14).

> *"If you're a creator, normally I guess that it means that you spend some time understanding what a GPT is and how it works. Because it's a feature that it's more oriented to, so to say, advanced users. [...] But definitely most users of these kinds of services don't have the same level of knowledge or even don't have the very basic level of knowledge required about data privacy to use them. So there's a concern. And there's a difference."* - P2

However, when using GPTs created by others, they remained skeptical about the creators' identities. P2 explained: *"I typically don't trust most of the people creating GPTs. It might seem contradictory, but it's the reality. I think that even if there's not malicious intent in a big percentage of the GPT publishers, there is some sort of spam in them."* (P12)

Moreover, some participants shared that being a creator gave them a sense of control, making them less worried, compared to when they were in the user role. P12 commented:

> *"I think like subconsciously as a user you feel you are more like "I don't know." As a creator, you kind of have autonomy. You feel like because I'm creating the thing, it feels like you have the autonomy and kind of you are in charge of the whole thing. Although they are, on the back end they are still OpenAI."* - P12

5.5.2 Responsibility Seems to Increase Concerns and Uncertainty. Our creator participants mentioned that being the author of GPTs gave them a sense of responsibility, not only for their own privacy but also for the privacy of their users (P10, P17). Engaging as users of other GPTs also encouraged them to approach privacy more thoughtfully in their own creations, as they recognized that end-users would likely share similar concerns (P12, P13). As a result, they felt more concerned about privacy issues in their role as creators.

> *"I guess probably I'm thinking more about when I'm creating because I'm kind of like being responsible for other people's information and I just want to make sure it's safe and secure. So obviously I'm like kind of like making sure every step is done to protect the information."* - P10

Although responsibilities shape privacy perceptions, the blurred line between different roles also created uncertainty among our participants regarding the responsibilities that come with each role. Some users believed that creators should be responsible for end-users' privacy (P3, P10, P17). However, some creators felt differently. They saw themselves as users of OpenAI and believed that it was OpenAI's responsibility to protect both users' and creators' privacy (P1–2, P5–6, P9, P16).

> *"I don't have the feeling it's my responsibility. I think people are already aware of this, and it's more OpenAI's responsibility. [...] And again, if I'm making money with it, let's say if I make good money with it or a little money, whatever, then I feel it is my responsibility. But right now, we're not in that phase at all. It's just a freebie, right?"* - P1

6 DISCUSSION

We discuss our findings in this section, focusing on how GPT usage is shaped by participants' general past experiences and predispositions (Section 6.1), the unique privacy challenges with GPTs (Section 6.2), interpretations of the user-creator spectrum (Section 6.3), and practical recommendations to enhance user privacy in GPTs (Section 6.4).

6.1 Prior Experience and Predispositions Shape GPT Usage

Participants' privacy concerns and practices were shaped by their broader privacy experiences prior to usage of GPTs. These experiences shaped their mental models and evaluations of GPTs. Informed by prior experiences, participants formed an understanding of norms regarding what is appropriate in specific contexts, echoing the concept of *contextual integrity* [71]. They evaluated GPTs by considering the creators, whether they were familiar services or individual creators, and adjusted their expectations about privacy and the appropriateness of external actions. When GPTs deviate

from these expectations, such as by making action requests that participants found misaligned with contextual norms, it led to a breakdown of contextual integrity and raised concerns among participants. Here, familiar services or well-known companies also served as heuristic cues for participants when deciding to use a GPT. These privacy associations, rooted in past experiences rather than direct interaction with GPTs, illustrate the role of *privacy heuristics* [3, 52], which have been recognized as a key factor in people's privacy decision-making processes across various online platforms [33, 37, 101].

Moreover, some participants experienced *privacy fatigue*—feeling overwhelmed and resigned to the inevitability of losing control over personal data [25, 26]—independently of GPTs. This fatigue led to a general indifference to privacy concerns, which also extended to use of GPTs. This attitude can be interpreted through Lazarus's coping theory, which categorizes coping strategies into problem-based and emotion-based approaches [58, 59]. When faced with limited control over privacy, users may adopt emotional coping strategies, adjusting their emotional responses rather than directly taking active actions to mitigate privacy risks.

The acceptance of privacy risks in exchange for advanced features also reflects the theory of *privacy calculus*, where individuals weigh the potential benefits of technology against the risks of data sharing [38, 54]. Our findings show that while some remain cautious about adopting GPTs, particularly GPTs with actions, others are notably adventurous in experimenting with new GPTs. An intriguing aspect is how some participants justify their willingness to engage with these GPTs by identifying themselves as "tech enthusiasts" or "non-average users." This self-perception appears to normalize their risk-taking behavior, framing it as a kind of personal mission to explore emerging products like GPTs. This behavior reflects an evaluation of self-efficacy. Individuals who perceive themselves as highly competent in technology usage and privacy management may exhibit more risk-taking behaviors, a phenomenon extensively discussed in privacy studies across various technology contexts [22, 40].

6.2 Privacy Challenges with GPTs

We identified privacy concerns (UC1-UC4, CC1) and practices (UP1-UP4, CP1-CP2) associated with GPT usage, several of which align with behaviors also observed in prior studies of other media or AI products. For example, self-censorship in GPT usage resembles user behavior in other LLM conversational agents, where users control and pseudonymize their input to mitigate risks [115]. Similarly, minimizing GPT usage traces, such as deleting chat histories, reflects a common practice observed with other digital tools, such as web browsers [17, 70]. However, some participants doubted whether deleting chat history truly ensures data deletion, seeing it more as psychological reassurance than an actual removal. In reality, due to the challenges of reliably reversing the influence of data integrated into trained models, machine unlearning remains an unresolved issue [108, 111, 114]. While participants' doubts about deletion practices are reasonable, this highlights the critical need for AI products to transparently communicate the actual impact of such privacy practices.

The shared privacy challenges between GPTs and other AI products highlight two critical considerations: first, the need to promote general AI literacy to help users better understand the capabilities and limitations of AI systems; and second, the importance of thoughtful product design that effectively communicates necessary distinctions. With this foundation, we now focus on the specific privacy challenges unique to GPTs.

Privacy Communication Mismatch and Mistrust Between GPT App Users, Creators, and Platforms. Our findings reveal misunderstandings between users and creators, such that many user participants mistakenly believed that creators have access to their GPT interactions in Scenario 1, which is not the case. Conversely, creator participants were confused about their responsibilities, and attributed safeguarding the privacy of both users and creators on the platform to OpenAI. Therefore, users may place misdirected distrust in creators, assuming they are responsible for privacy protections they cannot provide. Despite the miscommunication between users and creators, both groups expressed high trust in OpenAI as the platform provider to provide privacy protection for GPTs' creation and usage. However, this is outside OpenAI's policy [78, 84]. Such misconceptions stem from the asymmetric distribution of information between users, creators, and platform providers. These misconceptions not only strain the relationship between users and creators but may also weaken overall trust in the GPT ecosystem. Trust is a critical factor that moderates privacy behaviors, such as the decision to use certain services or disclose personal information [47, 104]. The misplaced trust dynamic within GPT platforms highlights a significant privacy concern and underscores the need for clear communication to clarify roles, responsibilities, and data processes.

GPT App-Enabled Precise Profiling and Third Parties. The specialization of GPTs in specific tasks is one of their greatest strengths, enabling tailored applications that meet precise user needs. However, this specialization also raises significant privacy concerns. Participants noted that while engaging deeply with task-specific GPTs, they paid less attention to privacy risks, echoing prior findings that users tend to overshare with human-like chatbots [44, 87, 105]. This oversharing is particularly concerning with GPTs, as their task-specific focus allows for more detailed user profiling, increasing the risks of targeted marketing and data misuse.

A related issue is the involvement of third parties, which amplifies these risks. GPT integrations with tools like calendars, reservation systems, and other services streamline workflows but introduce additional privacy vulnerabilities. Third parties may collect and utilize detailed user profiles for commercial purposes, further complicating consent dynamics. Beyond privacy issues that are similar to what has been discussed in terms of LLM algorithms, such as data opaque processing [60, 109] and unauthorized dissemination [21], GPTs open new privacy gateways as data transitions between LLMs and all other services.

Call for Regulatory Guidelines for GPT Apps. A concern specific to GPTs is creators' worries that their creations of GPTs can be reverse-engineered through user prompts, yet they feel powerless to prevent this. The only actions they can take are to set GPTs

to private or spend extra effort in prompting their GPTs to safeguard their knowledge. This concern reflects an unresolved vulnerability with LLMs [64, 113]. Moreover, some participants reported low-quality content in the GPT store, like spam, which was also spotted in recent studies of the GPT store [42]. However, as our participants added, there is a lack of mechanisms to report malicious content. Given these unregulated and unsafe conditions, both users and creators expressed expectations that platforms should take responsibility for verifying GPTs and scrutinizing creators; however, these expectations are misaligned with the actual platform policies and terms [78, 84]. This disconnect causes hesitancy among participants to fully adopt GPTs, especially ones created by independent developers, thereby undermining innovation within the platform ecosystem. Regulatory actions are needed from both platform providers and governmental oversight bodies.

6.3 Benefits and Challenges of the Blurred User-creator Spectrum

Our study highlights the fluid user-creator spectrum in relation to GPT platforms, where individuals engage in roles ranging from casual experimentation to professional development. As needs evolve, participants often adopt dual roles, simultaneously using and creating GPTs.

The dual role of being both a user and a creator allows individuals to reflect from both perspectives when engaging with GPTs. This reflection enhances their understanding of how GPTs function on the backend and provides a greater sense of control compared to being solely end-users. Prior research has shown that more technically advanced users tend to have a more sophisticated understanding of system functionality [36, 49, 73]. Our study supports this finding and further demonstrates that such understandings may alleviate privacy concerns, particularly in scenarios where users are familiar with the system's operations. Traditionally, users and developers have been studied as distinct groups. However, the blurred boundaries between users and creators in the GPT ecosystem highlight the nuanced role of identity in shaping privacy perceptions, suggesting a potential future direction for research that incorporates role identity into the study of privacy.

Another aspect of the dual role is related to responsibility. Our findings show that the role shift to being a creator gave participants a heightened sense of responsibility, prompting them to account for others' privacy in addition to their own, and becoming more mindful of privacy considerations when creating apps. These findings highlight the potential benefits of involving users in the privacy co-creation process, as this approach may not only enhance their sense of responsibility as individual users but also foster accountability toward the broader community. This aligns with prior work suggesting that bridging the gap between users and developers can enhance collaboration in software development [46].

Moreover, among creators there are nuanced divides based on levels of experience. As our findings show, some participants casually create GPTs, while professional creators build GPTs that gain popularity. The low barrier to entry for creating GPTs stresses the need to consider more granular divides of different types of user-creators. This diversity introduces differences in how creators engage with privacy and accountability. For example, studies suggest that hobbyist developers often view privacy policies as mere platform requirements, whereas professional developers adopt a more compliance-focused approach, emphasizing risk mitigation [94]. Clear responsibility is essential for fostering accountability. Accountability, in turn, is crucial for encouraging developers' commitment and ensuring the delivery of high-quality software [30, 53]. Conversely, a lack of accountability may end up in unreliable GPTs. As the GPT store rapidly expands, leading to a surge of unreliable apps [45], it is important to establish responsibility and foster accountability among creators to ensure the healthy development of the ecosystem.

6.4 Recommendations

Our findings highlight the openness of GPT creation, such that anyone can create a GPT, but also the need for a more regulated platform. We propose several suggestions, including communicating data processes and responsibilities transparently to all users and providing clear guidelines and resources to help creators develop legitimate privacy policies.

Clarify Responsibility. In the OpenAI GPT store, there is no clear distinction between users and creators; everyone holds the same type of account, is considered a user by OpenAI, and is treated uniformly. According to OpenAI's processing addendum [78], both users and creators are responsible for ensuring the security and compliance of certain configurations, rather than relying solely on OpenAI. Similarly, OpenAI's plugin policy [84] places the responsibility for the privacy and security of API and plugin requests on users. This uniform approach may further complicate the understanding of responsibility and compliance. Although the terms and privacy policies apply equally to all users, we noticed differences in how simple users and user-creators perceive responsibility. Therefore, we suggest that the platform segment its users through a nuanced consideration of the entire user-creator spectrum, providing targeted communication about their rights and responsibilities at key points (e.g., during GPT creation). Enhancing the management of these responsibilities would improve privacy within LLM-based app ecosystems like GPTs.

Contextual Data Transparency. Confusion and unclear mental models often arise across the user-creator spectrum, especially regarding third-party data practices. Currently, GPTs alert users with permission requests when an action is executed. While these warnings help, they should better align with users' mental models to improve decision-making [69], and more educational messages or risk feedback should be added to improve clarity, transparency, and effectiveness [2]. To improve communication about data practices, the elements of timing and modalities should be considered, as suggested by the design spaces of privacy notices proposed by Schaub et al. [92].

Privacy concerns and practices should be addressed at critical moments, with more context-specific indicators integrated into the interface. For instance, for GPTs without action capabilities, it is important to clarify that chat information is not accessible by the creator. In GPTs with action capabilities, notifications should inform users of data collection before any action is taken, and clear status updates should be provided during and after the action,

including details such as connection duration and post-service login status—areas that participants frequently found confusing.

Additionally, different modalities, such as visual elements, can be integrated into the design of privacy notices. For instance, privacy nutrition labels [50] have been suggested as an effective way to provide clearer and more succinct information about data handling, compared to traditional text-based policies [14]. Feedback mechanisms, like safety barometers used in password settings [95], can also help users better understand the consequences of their actions and make more informed decisions [14].

In-house Privacy Guide. As a result of the low barrier to entry for creating GPTs, many user-creators lack the expertise to implement effective safeguards, potentially exposing both themselves and their users to privacy risks, such as reusing or auto-generating privacy notices that do not accurately reflect their services [45]. To address this issue, we recommend that the platform integrate more effective guideline tools to support user-creators. For example, an easily accessible in-house GPT could allow users and creators to share their concerns and receive relevant answers and feedback, rather than having to navigate through lengthy documents. Notably, the concept of privacy support through interactive dialogue is not new and was proposed prior to the advent of GPT models [41]. With the rise of LLM-based applications like GPTs, this approach becomes more feasible and can be seamlessly integrated into systems to offer on-the-go privacy guidance.

In summary, the recommendations outlined above aim to reduce anxiety [90] and mitigate privacy fatigue by reinforcing users' sense of control and security [106]. Determining what qualifies as sufficient privacy communication requires continuous interdisciplinary dialogue between legal, technical, and design domains. Efforts across and within these disciplines to clarify data practices for OpenAI and similar providers are crucial to ensure transparent and accessible privacy communication.

6.5 Limitations and Future Work

This work examines privacy perceptions of GPTs from both user and creator perspectives, exploring how their roles shape their views. While we gained valuable insights, some key questions remain unanswered for future research. First, the study concentrated on OpenAI's GPTs service because of its extensive user base and widespread popularity (over three million custom GPTs [83]). Although one of our participants with experience on other platforms expressed similar privacy concerns regarding OpenAI GPTs and other platforms, further investigation is warranted to evaluate whether these privacy perceptions extend to other platforms with different creation styles or user demographics. Second, our qualitative study was conducted with a small sample of participants. Recruiting a larger and more diverse sample would likely provide broader insights and enrich our findings. The growing ease of creating LLM-based apps such as GPTs presents new opportunities for future research. This includes exploring how different services can be integrated with LLM apps and their impact on privacy. We recommend that future research build on our findings and improve the design and regulation of GPTs and similar platforms.

7 CONCLUSIONS

This paper explores privacy perceptions in GPTs (custom LLM apps on the OpenAI platform) from both user and creator perspectives through semi-structured interviews (N=23). Our study indicates that GPT users and creators represent a spectrum of experiences and motivations, ranging from casual experimentation for personal use to professional GPT creation for organizational purposes or gaining public recognition. Many people assumed dual roles, creating GPTs for others while also using their own. Among the users, those without creation experience often lacked a clear understanding of GPT data flows.

Examining privacy concerns and practices from both users' and creators' perspectives, our study highlights common privacy issues related to data handling during collection, processing, and dissemination, particularly when third parties are involved. Users also raised concerns about the absence of regulatory guidelines. In response, users adopted privacy practices such as self-censoring input, assessing GPT trustworthiness, and minimizing digital traces, while some saw GPT usage as a privacy trade-off. Creators had specific concerns about intellectual property being compromised and actively safeguarded their creation knowledge by adjusting settings or using prompts. In reflection on different roles, responsibilities and expertise emerged as key factors shaping privacy perceptions across the user-creator spectrum.

To enhance privacy practices, we recommend improving transparency, clarifying responsibilities, and fostering responsible GPT design. These insights will contribute to better regulation of emerging applications and support ethical development of GPTs and similar platforms.

8 DATA AVAILABILITY

The supplementary materials, including the interview protocol and screening survey for this study, are available on OSF: https://doi.org/10.17605/OSF.IO/2PC3R.

ACKNOWLEDGMENTS

We thank all participants for their time and participation. We also appreciate the valuable feedback from members of the HASP lab and the anonymous reviewers. This research was funded by INCIBE's strategic SPRINT (Seguridad y Privacidad en Sistemas con Inteligencia Artificial) C063/23 project with funds from the EU-NextGenerationEU through the Spanish government's Plan de Recuperación, Transformación y Resiliencia. This work was supported by the Research Council of Finland under the Enjoyable Security project (grants 345991 and 345992). Rongjun Ma conducted the study primarily during a research visit to VRAIN.

REFERENCES

[1] Yasemin Acar, Michael Backes, Sascha Fahl, Doowon Kim, Michelle L Mazurek, and Christian Stransky. 2016. You get where you're looking for: The impact of information sources on code security. In *2016 IEEE symposium on security and privacy (SP)*. IEEE, IEEE, San Jose, CA, USA, 289–305.

[2] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, Yang Wang, and Shomir Wilson. 2017. Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online. *ACM Comput. Surv.* 50, 3, Article 44 (aug 2017), 41 pages. https://doi.org/10.1145/3054926

[3] Alessandro Acquisti and Jens Grossklags. 2007. What Can Behavioral Economics Teach Us About Privacy. In *Digital Privacy: Theory, Technologies and Practices*, A. Acquisti, S. Gritzalis, C. Lambrinoudakis, and S. De Capitani di Vimercati (Eds.). Auerbach Publications, Boca Raton, FL, USA, 363–377. https://doi.org/10.1201/9781420052183.ch18

[4] Rakesh Agrawal and Ramakrishnan Srikant. 2000. Privacy-preserving data mining. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data* (Dallas, Texas, USA) *(SIGMOD '00)*. Association for Computing Machinery, New York, NY, USA, 439–450. https://doi.org/10.1145/342009.335438

[5] Lana Aljuaid. 2024. ChatGPT - Personal Color Analysis. https://chat.openai.com/g/g-35kDoPvW7-personal-color-analysis. (Accessed on 03/27/2024).

[6] NAIF J ALOTAIBI. 2024. ChatGPT - image generator. https://chat.openai.com/g/g-pmuQfob8d-image-generator. (Accessed on 03/27/2024).

[7] Irwin Altman. 1974. Privacy: A Conceptual Analysis. In *Man-Environment Interactions: Evaluations and Applications: Part 2*, D. H. Carson (Ed.). Environmental Design Research Association, Washington, DC, USA, 3–28.

[8] Benjamin Andow, Samin Yaseer Mahmud, Justin Whitaker, William Enck, Bradley Reaves, Kapil Singh, and Serge Egelman. 2020. Actions speak louder than words:{Entity-Sensitive} privacy policy and data flow analysis with {PoliCheck}. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Boston, MA, USA, 985–1002.

[9] Sagiv Antebi, Noam Azulay, Edan Habler, Ben Ganon, Asaf Shabtai, and Yuval Elovici. 2024. GPT in Sheep's Clothing: The Risk of Customized GPTs. arXiv:2401.09075

[10] Atlas. 2024. ATLAS.ti | The #1 Software for Qualitative Data Analysis - ATLAS.ti. https://atlasti.com/. (Accessed on 11/27/2024).

[11] Rebecca Balebako, Abigail Marsh, Jialiu Lin, Jason Hong, and Lorrie Faith Cranor. 2014. The privacy and security behaviors of smartphone app developers. In *Workshop on Usable Security (USEC)*. Citeseer, Internet Society, San Diego, CA, USA, 1–10. http://dx.doi.org/10.14722/usec.2014.23006

[12] Gaurav Bansal, Fatemeh "Mariam" Zahedi, and David Gefen. 2010. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems* 49, 2 (2010), 138–150. https://doi.org/10.1016/j.dss.2010.01.010

[13] Natã M. Barbosa, Zhuohao Zhang, and Yang Wang. 2020. Do Privacy and Security Matter to Everyone? Quantifying and Clustering User-Centric Considerations About Smart Home Device Adoption. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. USENIX Association, Virtual Event, 417–435. https://www.usenix.org/conference/soups2020/presentation/barbosa

[14] Susanne Barth, Dan Ionita, and Pieter Hartel. 2022. Understanding Online Privacy—A Systematic Review of Privacy Visualizations and Privacy by Design Guidelines. *ACM Computing Surveys (CSUR)* 55, 3, Article 63 (feb 2022), 37 pages. https://doi.org/10.1145/3502288

[15] Charlotte Bird, Eddie Ungless, and Atoosa Kasirzadeh. 2023. Typology of Risks of Generative Text-to-Image Models. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (Montréal, QC, Canada) *(AIES '23)*. Association for Computing Machinery, New York, NY, USA, 396–410. https://doi.org/10.1145/3600211.3604722

[16] Wunderwuzzi's blog. 2023. Malicious ChatGPT Agents: How GPTs Can Quietly Grab Your Data (Demo) · Embrace The Red. https://embracethered.com/blog/posts/2023/openai-custom-malware-gpt/. (Accessed on 03/27/2024).

[17] Sophie C Boerman, Sanne Kruikemeier, and Frederik J Zuiderveen Borgesius. 2021. Exploring motivations for online privacy protection behavior: Insights from panel data. *Communication Research* 48, 7 (2021), 953–977.

[18] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa

[19] Canva. 2024. Canva: Visual Suite for Everyone. https://www.canva.com/en_gb/. (Accessed on 11/25/2024).

[20] Canva. 2024. ChatGPT - Canva. https://chat.openai.com/g/g-alKfVrz9K-canva. (Accessed on 03/27/2024).

[21] Nicholas Carlini, Florian Tramèr, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, Tom Brown, Dawn Song, Úlfar Erlingsson, Alina Oprea, and Colin Raffel. 2021. Extracting Training Data from Large Language Models. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Virtual Event, 2633–2650. https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting

[22] Hsuan-Ting Chen and Wenhong Chen. 2015. Couldn't or wouldn't? The influence of privacy concerns and self-efficacy in privacy management on privacy protection. *Cyberpsychology, Behavior, and Social Networking* 18, 1 (2015), 13–19.

[23] Erika Chin, Adrienne Porter Felt, Vyas Sekar, and David Wagner. 2012. Measuring user confidence in smartphone security and privacy. In *Proceedings of the Eighth Symposium on Usable Privacy and Security* (Washington, D.C.) *(SOUPS '12)*. Association for Computing Machinery, New York, NY, USA, Article 1, 16 pages. https://doi.org/10.1145/2335356.2335358

[24] Eugene Cho, Jinyoung Kim, and S. Shyam Sundar. 2020. Will You Log into Tinder using your Facebook Account? Adoption of Single Sign-On for Privacy-Sensitive Apps. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3334480.3383074

[25] Hichang Cho, Pengxiang Li, and Zhang Hao Goh. 2020. Privacy Risks, Emotions, and Social Media: A Coping Model of Online Privacy. *ACM Trans. Comput.-Hum. Interact.* 27, 6, Article 40 (nov 2020), 28 pages. https://doi.org/10.1145/3412367

[26] Hanbyul Choi, Jonghwa Park, and Yoonhyuk Jung. 2018. The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior* 81 (2018), 42–51. https://doi.org/10.1016/j.chb.2017.12.001

[27] Jessica Colnago, Lorrie Faith Cranor, Alessandro Acquisti, and Kate Hazel Stanton. 2022. Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. USENIX Association, Boston, MA, USA, 331–346.

[28] Consensus. 2024. ChatGPT - Consensus. https://chat.openai.com/g/g-bo0FiWLY7-consensus. (Accessed on 03/27/2024).

[29] Tamara Dinev and Paul Hart. 2006. An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research* 17, 1 (2006), 61–80. https://doi.org/10.1287/isre.1060.0080

[30] Yael Dubinsky, Avi Yaeli, and Alex Kofman. 2010. Effective management of roles and responsibilities: Driving accountability in software development teams. *IBM Journal of Research and Development* 54, 2 (2010), 4–1. https://doi.org/10.1147/JRD.2009.2039896

[31] Adrienne Porter Felt, Erika Chin, Steve Hanna, Dawn Song, and David Wagner. 2011. Android permissions demystified. In *Proceedings of the 18th ACM conference on Computer and communications security*. Association for Computing Machinery, Chicago, IL, USA, 627–638.

[32] Adrienne Porter Felt, Helen J Wang, Alexander Moshchuk, Steve Hanna, and Erika Chin. 2011. Permission re-delegation: Attacks and defenses.. In *USENIX security symposium*, Vol. 30. USENIX Association, San Francisco, CA, USA, 88.

[33] Christian Flender and Gero Mueller. 2012. Type Indeterminacy in Privacy Decisions: The Privacy Paradox Revisited. In *Quantum Interaction: 6th International Symposium, QI 2012, Paris, France, June 27-29, 2012, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 7620)*. Springer, Paris, France, 148–159. https://doi.org/10.1007/978-3-642-35659-9_14

[34] Marco Furini, Silvia Mirri, Manuela Montangero, and Catia Prandi. 2020. Privacy perception when using smartphone applications. *Mobile Networks and Applications* 25 (2020), 1055–1061. https://doi.org/10.1007/s11036-020-01529-z

[35] Patricia I. Fusch and Lawrence R. Ness. 2015. Are We There Yet? Data Saturation in Qualitative Research. *The Qualitative Report* 20, 9 (2015), 1408–1416. https://doi.org/10.46743/2160-3715/2015.2281

[36] Kevin Gallagher, Sameer Patil, and Nasir Memon. 2017. New Me: Understanding Expert and Non-Expert Perceptions and Usage of the Tor Anonymity Network. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association, Santa Clara, CA, 385–398. https://www.usenix.org/conference/soups2017/technical-sessions/presentation/gallagher

[37] Andrew Gambino, Jinyoung Kim, S. Shyam Sundar, Jun Ge, and Mary Beth Rosson. 2016. User Disbelief in Privacy Paradox: Heuristics that determine Disclosure. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 2837–2843. https://doi.org/10.1145/2851581.2892413

[38] Nina Gerber, Paul Gerber, and Melanie Volkamer. 2018. Explaining the Privacy Paradox: A Systematic Review of Literature Investigating Privacy Attitude and Behavior. *Computers & Security* 77 (2018), 226–261. https://doi.org/10.1016/j.cose.2018.04.002

[39] gptavern.mindgoblinstudios.com. 2024. ChatGPT - Grimoire. https://chat.openai.com/g/g-n7Rs0IK86-grimoire. (Accessed on 03/27/2024).

[40] Syed H. Akhter. 2014. Privacy concern and online transactions: the impact of internet self-efficacy and internet involvement. *Journal of Consumer Marketing* 31, 2 (2014), 118–125.

[41] Hamza Harkous, Kassem Fawaz, Kang G Shin, and Karl Aberer. 2016. PriBots: Conversational privacy with chatbots. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. USENIX Association, Denver, CO, USA, 6 pages.

[42] Xinyi Hou, Yanjie Zhao, and Haoyu Wang. 2024. On the (In)Security of LLM App Stores. arXiv:2407.08422 Available at https://arxiv.org/abs/2407.08422.

[43] MixerBox Inc. 2024. ChatGPT - MixerBox Calendar. https://chat.openai.com/g/g-aI4P3mWio-mixerbox-calendar. (Accessed on 03/28/2024).

[44] Carolin Ischen, Theo Araujo, Hilde Voorveld, Guda van Noort, and Edith Smit. 2020. Privacy concerns in chatbot interactions. In *Chatbot Research and Design: Third International Workshop (Lecture Notes in Computer Science)*. Springer, Amsterdam, Netherlands, 34–48. https://doi.org/10.1007/978-3-030-39540-7_3

[45] Evin Jaff, Yuhao Wu, Ning Zhang, and Umar Iqbal. 2024. Data Exposure from LLM Apps: An In-depth Investigation of OpenAI's GPTs. arXiv:2408.13247 https://arxiv.org/abs/2408.13247

[46] Leonie Jahn, Philip Engelbutzeder, Dave Randall, Yannick Bollmann, Vasilis Ntouros, Lea Katharina Michel, and Volker Wulf. 2024. In Between Users and Developers: Serendipitous Connections and Intermediaries in Volunteer-Driven

Open-Source Software Development. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 924, 15 pages. https://doi.org/10.1145/3613904.3642541

[47] Adam N Joinson, Ulf-Dietrich Reips, Tom Buchanan, and Carina B Paine Schofield. 2010. Privacy, trust, and self-disclosure online. *Human–Computer Interaction* 25, 1 (2010), 1–24.

[48] Zhigang Kan, Linbo Qiao, Hao Yu, Liwen Peng, Yifu Gao, and Dongsheng Li. 2023. Protecting User Privacy in Remote Conversational Systems: A Privacy-Preserving Framework Based on Text Sanitization. arXiv:2306.08223

[49] Ruogu Kang, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. 2015. "My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. USENIX Association, Ottawa, 39–52. https://www.usenix.org/conference/soups2015/proceedings/presentation/kang

[50] Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. 2010. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1573–1582. https://doi.org/10.1145/1753326.1753561

[51] Nazish Khalid, Adnan Qayyum, Muhammad Bilal, Ala Al-Fuqaha, and Junaid Qadir. 2023. Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Computers in Biology and Medicine* 158 (2023), 106848. https://doi.org/10.1016/j.compbiomed.2023.106848

[52] Bart P. Knijnenburg, Alfred Kobsa, and Hongxia Jin. 2013. Dimensionality of Information Disclosure Behavior. *International Journal of Human-Computer Studies* 71, 12 (2013), 1144–1162. https://doi.org/10.1016/j.ijhcs.2013.06.003

[53] Umme Ayman Koana, Francis Chew, Chris Carlson, and Maleknaz Nayebi. 2023. Ownership in the Hands of Accountability at Brightsquid: A Case Study and a Developer Survey. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (San Francisco, CA, USA) *(ESEC/FSE 2023)*. Association for Computing Machinery, New York, NY, USA, 2008–2019. https://doi.org/10.1145/3611643.3613890

[54] Spyros Kokolakis. 2017. Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon. *Computers & Security* 64 (2017), 122–134. https://doi.org/10.1016/j.cose.2015.07.002

[55] Klaus Krippendorff. 2004. Reliability in content analysis: Some common misconceptions and recommendations. *Human communication research* 30, 3 (2004), 411–433.

[56] Amit Kulkarni. 2021. *GitHub Copilot AI is Leaking Functional API Keys*. Analytics Drift. https://analyticsdrift.com/github-copilot-ai-is-leaking-functional-api-keys/ (Accessed on 08/26/2024).

[57] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* 33, 1 (1977), 159–174.

[58] Richard S Lazarus. 1984. *Stress, appraisal, and coping*. Springer, New York, NY, USA.

[59] Richard S Lazarus. 1991. *Emotion and adaptation*. Oxford University Press, New York, NY, USA.

[60] Hao-Ping (Hank) Lee, Yu-Ju Yang, Thomas Serban Von Davier, Jodi Forlizzi, and Sauvik Das. 2024. Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 775, 19 pages. https://doi.org/10.1145/3613904.3642116

[61] Sangmin Lee, Edmund L. Wong, Deepak Goel, Mike Dahlin, and Vitaly Shmatikov. 2013. πBox: A Platform for Privacy-Preserving Apps. In *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*. USENIX Association, Lombard, IL, 501–514. https://www.usenix.org/conference/nsdi13/technical-sessions/presentation/lee_sangmin

[62] Tianshi Li, Elizabeth Louie, Laura Dabbish, and Jason I. Hong. 2021. How Developers Talk About Personal Data and What It Means for User Privacy: A Case Study of a Developer Forum on Reddit. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 220 (jan 2021), 28 pages. https://doi.org/10.1145/3432919

[63] Yuan Li. 2011. Empirical studies on online information privacy concerns: Literature review and an integrative framework. *Communications of the Association for Information Systems* 28, 1 (2011), 28. https://doi.org/10.17705/1CAIS.02828

[64] Zi Liang, Haibo Hu, Qingqing Ye, Yaxin Xiao, and Haoyang Li. 2024. Why Are My Prompts Leaked? Unraveling Prompt Extraction Threats in Customized Large Language Models. arXiv:2408.02416 https://arxiv.org/abs/2408.02416

[65] Lydia Manikonda, Aditya Deotale, and Subbarao Kambhampati. 2018. What's up with privacy? User preferences and privacy concerns in intelligent personal assistants. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery, New Orleans, LA, USA, 229–235.

[66] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (nov 2019), 23 pages. https://doi.org/10.1145/3359174

[67] Christian Meurisch, Bekir Bayrak, and Max Mühlhäuser. 2020. Privacy-preserving AI Services Through Data Decentralization. In *Proceedings of The Web Conference 2020* (Taipei, Taiwan) *(WWW '20)*. Association for Computing Machinery, New York, NY, USA, 190–200. https://doi.org/10.1145/3366423.3380106

[68] George R. Milne, George Pettinico, Fatima M. Hajjat, and Ereni Markos. 2017. Information Sensitivity Typology: Mapping the Degree and Type of Risk Consumers Perceive in Personal Data Sharing. *Journal of Consumer Affairs* 51, 1 (2017), 133–161. https://doi.org/10.1111/joca.12111

[69] M. Granger Morgan and Max Henrion. 1992. *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press, New York, NY.

[70] Ambar Murillo, Andreas Kramm, Sebastian Schnorf, and Alexander De Luca. 2018. " If I press delete, it's gone"-User Understanding of Online Data Deletion and Expiration. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. USENIX Association, Baltimore, MD, USA, 329–339.

[71] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[72] Helen Nissenbaum. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books, Stanford, CA.

[73] Maggie Oates, Yama Ahmadullah, Abigail Marsh, Chelse Swoopes, Shikun Zhang, Rebecca Balebako, and Lorrie Faith Cranor. 2018. Turtles, locks, and bathrooms: Understanding mental models of privacy through illustration. *Proceedings on Privacy Enhancing Technologies* 2018, 4 (2018), 5–32. https://doi.org/10.1515/popets-2018-0029

[74] Judith S. Olson, Jonathan Grudin, and Eric Horvitz. 2005. A study of preferences for sharing and privacy. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA) *(CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1985–1988. https://doi.org/10.1145/1056808.1057073

[75] OpenAI. 2023. Actions - OpenAI API. https://platform.openai.com/docs/actions/introduction. (Accessed on 09/02/2024).

[76] OpenAI. 2023. Explore GPTs. https://chatgpt.com/gpts. (Accessed on 09/07/2024).

[77] OpenAI. 2023. Introducing GPTs | OpenAI. https://openai.com/index/introducing-gpts/. (Accessed on 11/25/2024).

[78] OpenAI. 2024. *Data Processing Addendum*. OpenAI. https://openai.com/policies/data-processing-addendum/ (Accessed on 05/09/2024).

[79] OpenAI. 2024. EU privacy policy | OpenAI. https://openai.com/policies/privacy-policy/. (Accessed on 11/25/2024).

[80] OpenAI. 2024. GPT Action authentication - OpenAI API. https://platform.openai.com/docs/actions/authentication. (Accessed on 11/26/2024).

[81] OpenAI. 2024. GPTs Data Privacy FAQs | OpenAI Help Center. https://help.openai.com/en/articles/8554402-gpts-data-privacy-faqs. (Accessed on 11/25/2024).

[82] OpenAi. 2024. GPTs Data Privacy FAQs | OpenAI Help Center. https://help.openai.com/en/articles/8554402-gpts-data-privacy-faqs. (Accessed on 11/25/2024).

[83] OpenAI. 2024. Introducing the GPT Store. https://openai.com/blog/introducing-the-gpt-store. (Accessed on 03/27/2024).

[84] OpenAI. 2024. *Plugin Terms*. OpenAI. https://openai.com/policies/plugin-terms/ (Accessed on 05/09/2024).

[85] Google PAIR. 2019. *People + AI Guidebook*. Google Research. https://design.google/aiguidebook

[86] Aswati Panicker, Novia Nurain, Zaidat Ibrahim, Chun-Han (Ariel) Wang, Seung Wan Ha, Yuxing Wu, Kay Connelly, Katie A. Siek, and Chia-Fang Chung. 2024. Understanding fraudulence in online qualitative studies: From the researcher's perspective. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 824, 17 pages. https://doi.org/10.1145/3613904.3642732

[87] Arielle Pardes. 2018. *The Emotional Chatbots Are Here to Probe Our Feelings*. Condé Nast. https://www.wired.com/story/replika-open-source/

[88] pulsr.co.uk. 2024. ChatGPT - math. https://chat.openai.com/g/g-odWlfAKWM-math. (Accessed on 03/27/2024).

[89] puzzle.today. 2024. ChatGPT - Write For Me. https://chat.openai.com/g/g-B3hgivKK9-write-for-me/c/42960e67-173a-46df-b0e2-1e97d5f6ed94. (Accessed on 03/27/2024).

[90] Gerardo Ramirez and Sian L. Beilock. 2011. Writing about testing worries boosts exam performance in the classroom. *Science* 331, 6014 (2011), 211–213. https://doi.org/10.1126/science.1199427

[91] Michael Ryan. 2024. ChatGPT - Schedule Assistant. https://chat.openai.com/g/g-rk0wck8W0-schedule-assistant/c/321b124a-12c5-4173-b646-72eef5dc3391. (Accessed on 03/28/2024).

[92] Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. 2015. A design space for effective privacy notices. In *Eleventh symposium on usable privacy and security (SOUPS 2015)*. USENIX Association, Ottawa, Canada, 1–17.

[93] SEO.AI. 2024. GPT Store Statistics & Facts: Contains 159.000 of the 3 million created GPTs. https://seo.ai/blog/gpt-store-statistics-facts. (Accessed on

03/27/2024).

[94] William Seymour, Noura Abdi, Kopo M. Ramokapane, Jide Edu, Guillermo Suarez-Tangil, and Jose Such. 2024. Voice App Developer Experiences with Alexa and Google Assistant: Juggling Risks, Liability, and Security. In *Proceedings of the 33rd USENIX Security Symposium (USENIX Security 2024)*. USENIX Association, Philadelphia, PA, USA, 5035–5052. https://arxiv.org/abs/2311.08879

[95] Richard Shay, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Alain Forget, Saranga Komanduri, Michelle L. Mazurek, William Melicher, Sean M. Segreti, and Blase Ur. 2015. A Spoonful of Sugar? The Impact of Guidance and Feedback on Password-Creation Behavior. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2903–2912. https://doi.org/10.1145/2702123.2702586

[96] Tanmay Singh, Harshvardhan Aditya, Vijay K Madisetti, and Arshdeep Bahga. 2024. Whispered tuning: Data privacy preservation in fine-tuning llms through differential privacy. *Journal of Software Engineering and Applications* 17, 1 (2024), 1–22. https://doi.org/10.4236/jsea.2024.171001.

[97] Rocky Slavin, Xiaoyin Wang, Mitra Bokaei Hosseini, James Hester, Ram Krishnan, Jaspreet Bhatia, Travis D Breaux, and Jianwei Niu. 2016. Toward a framework for detecting privacy policy violations in android application code. In *Proceedings of the 38th International conference on software engineering*. ACM, Austin, Texas, USA, 25–36.

[98] H Jeff Smith, Tamara Dinev, and Heng Xu. 2011. Information privacy research: an interdisciplinary review. *MIS quarterly* 35, 4 (2011), 989–1015. https://doi.org/10.2307/41409970

[99] Daniel J Solove. 2005. A taxonomy of privacy. *U. Pa. l. Rev.* 154 (2005), 477.

[100] Gowthami Somepalli, Vasu Singla, Micah Goldblum, Jonas Geiping, and Tom Goldstein. 2023. Diffusion art or digital forgery? investigating data replication in diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Vancouver, Canada, 6048–6058. https://doi.org/10.48550/arXiv.2212.03860

[101] S. Shyam Sundar, Jinyoung Kim, Mary Beth Rosson, and Maria D. Molina. 2020. Online Privacy Heuristics that Predict Information Disclosure. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376854

[102] Mohammad Tahaei, Ruba Abu-Salma, and Awais Rashid. 2023. Stuck in the permissions with you: Developer & end-user perspectives on app permissions & their privacy ramifications. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg, Germany, 1–24.

[103] Guanhong Tao, Siyuan Cheng, Zhuo Zhang, Junmin Zhu, Guangyu Shen, and Xiangyu Zhang. 2023. Opening A Pandora's Box: Things You Should Know in the Era of Custom GPTs. arXiv:2401.00905

[104] Warda Usman, Jackie Hu, McKynlee Wilson, and Daniel Zappala. 2023. Distrust of big tech and a desire for privacy: Understanding the motivations of people who have voluntarily adopted secure email. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. USENIX, Anaheim, California, USA, 473–490.

[105] Evert Van den Broeck, Brahim Zarouali, and Karolien Poels. 2019. Chatbot Advertising Effectiveness: When Does the Message Get Through? *Computers in Human Behavior* 98 (September 2019), 150–157. https://doi.org/10.1016/j.chb.2019.04.009

[106] Karl Van Der Schyff, Greg Foster, Karen Renaud, and Stephen Flowerday. 2023. Online privacy fatigue: a scoping review and research agenda. *Future Internet* 15, 5 (2023), 164. https://doi.org/10.3390/fi15050164

[107] Michael Veale, Midas Nouwens, and Cristiana Santos. 2022. Impossible asks: can the transparency and consent framework ever authorise real-time bidding after the Belgian DPA decision? *Technology and Regulation* 2022 (2022), 12–22. https://doi.org/10.26116/

[108] Weiqi Wang, Zhiyi Tian, Chenhan Zhang, and Shui Yu. 2024. Machine Unlearning: A Comprehensive Survey. arXiv:2405.07406 Available at https://arxiv.org/abs/2405.07406.

[109] Laura Weidinger, Jonathan Uesato, Maribeth Rauh, Conor Griffin, Po-Sen Huang, John Mellor, Amelia Glaese, Myra Cheng, Borja Balle, Atoosa Kasirzadeh, Courtney Biles, Sasha Brown, Zac Kenton, Will Hawkins, Tom Stepleton, Abeba Birhane, Lisa Anne Hendricks, Laura Rimell, William Isaac, Julia Haas, Sean Legassick, Geoffrey Irving, and Iason Gabriel. 2022. Taxonomy of Risks posed by Language Models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 214–229. https://doi.org/10.1145/3531146.3533088

[110] Heng Xu, Sumeet Gupta, Mary Beth Rosson, and John M. Carroll. 2012. Measuring Mobile Users' Concerns for Information Privacy. In *Proceedings of the 33rd International Conference on Information Systems (ICIS)*. Association for Information Systems, Orlando, FL, USA, 16 pages. https://aisel.aisnet.org/icis2012/proceedings/ISSecurity/10

[111] Jie Xu, Zihan Wu, Cong Wang, and Xiaohua Jia. 2024. Machine unlearning: Solutions and challenges. *IEEE Transactions on Emerging Topics in Computational Intelligence* 8, 3 (2024), 2150–2168.

[112] Mike Z Yao, Ronald E Rice, and Kier Wallis. 2007. Predicting user concerns about online privacy. *Journal of the American Society for Information Science and Technology* 58, 5 (2007), 710–722. https://doi.org/10.1002/asi.20530

[113] Jiahao Yu, Yuhang Wu, Dong Shu, Mingyu Jin, and Xinyu Xing. 2023. Assessing Prompt Injection Risks in 200+ Custom GPTs. arXiv:2311.11538 Available at https://arxiv.org/abs/2311.11538.

[114] Haibo Zhang, Toru Nakamura, Takamasa Isohara, and Kouichi Sakurai. 2023. A review on machine unlearning. *SN Computer Science* 4, 4 (2023), 337.

[115] Zhiping Zhang, Michelle Jia, Hao-Ping Lee, Bingsheng Yao, Sauvik Das, Ada Lerner, Dakuo Wang, and Tianshi Li. 2024. "It's a Fair Game", or Is It? Examining How Users Navigate Disclosure Risks and Benefits When Using LLM-Based Conversational Agents. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu, HI, USA, 1–26. https://doi.org/10.1145/3613904.3642385

[116] Zejun Zhang, Li Zhang, Xin Yuan, Anlan Zhang, Mengwei Xu, and Feng Qian. 2024. A First Look at GPT Apps: Landscape and Vulnerability. arXiv:2402.15105 Available at https://arxiv.org/abs/2402.15105.

[117] JiaYing Zheng, HaiNan Zhang, LingXiang Wang, WangJie Qiu, HongWei Zheng, and ZhiMing Zheng. 2024. Safely Learning with Private Data: A Federated Learning Framework for Large Language Model. arXiv:2406.14898 Available at https://arxiv.org/abs/2406.14898.