Bot Among Us: Exploring User Awareness and Privacy Concerns About Chatbots in Group Chats

Kai-Hsiang Chou
National Taiwan University
b07705022@csie.ntu.edu.tw

Yi-An Wang
National Taiwan University
b09902007@csie.ntu.edu.tw

Chong Kai Lau
National Taiwan University
b10202012@ntu.edu.tw

Mahmood Sharif
Tel Aviv University
mahmoods@tauex.tau.ac.il

Hsu-Chun Hsiao
National Taiwan University
Academia Sinica
hchsiao@csie.ntu.edu.tw

Abstract

As chatbots become increasingly integrated into group conversations on instant messaging platforms, concerns arise about their impact on user privacy. While prior research has examined chatbot risks in one-on-one interactions, little is known about how users perceive and respond to privacy threats in group settings, where chatbots may silently access messages and metadata. To address this gap, we conducted an online survey (N=374) across five popular messaging platforms—WhatsApp, Discord, Telegram, Viber, and LINE—to evaluate user awareness, understanding of chatbot access, privacy concerns, and behavioral responses. We found that many users were unaware of bots in their group chats and significantly underestimated their data access: only 41.7% correctly identified what messages chatbots could access. Privacy concerns also rose sharply after users learned about actual bot permissions. Based on our findings, we propose a five-stage model that captures how users detect, interpret, and respond to chatbot-related privacy risks. We further analyzed the designs of platforms with official chatbot support through this model and found mismatches between design choices and user expectations. Finally, we offer design recommendations to improve transparency and user control in group chatbot-interactions.

Keywords

chatbots, group chats, messaging platforms, user awareness

1 Introduction

Instant messaging platforms have become essential communication tools in both personal and professional contexts. As these platforms evolve, chatbots—i.e., automated software agents that can interact with users—have become increasingly integrated into group chat environments across messaging services [12, 41]. These chatbots provide various functionalities ranging from content moderation [19, 38] to entertainment [8, 32], enhancing user experience and enabling automated tasks [13, 29] within group conversations.

Prior work has examined the security and privacy risks posed by chatbots and third-party apps in messaging platforms, identifying issues such as overprivileged bots [4, 12] flawed data access controls [3, 41], and ambiguous data-access policies [41] that can lead to privilege escalation and data breaches. Prior work has also examined users' privacy concerns and perceptions of chatbots, primarily in one-on-one settings, highlighting issues such as risks of self-disclosure and data misuse [15], loss of control over shared information [33], and the potential manipulations of users' willingness to disclose personal data [14, 18].

However, to the best of our knowledge, no prior work has examined how users interact with chatbots in group settings and perceive the associated privacy risks. Group chatbots differ from one-on-one chatbots in three ways. First, in one-on-one interactions, chatbots are the direct recipients of user messages, but in group chats, messages are typically directed at other group members, not the chatbot. Second, users may join group chats without realizing that a chatbot is present. Third, users often have no control over whether a specific chatbot is included in the conversation. Understanding how users perceive these chatbots and what privacy expectations they have is crucial for designing systems that respect user privacy while maintaining functionality.

Our research addresses this gap by investigating two research questions:

- RQ1: How do users interact with chatbots in group settings, and how do they perceive, understand, and respond to the associated privacy risks? What expectations do users hold regarding chatbot behavior and data access?
- RQ2: To what extent do current chatbot platform designs align with users' mental models and privacy expectations? Where misalignments exist, what design interventions can better support user understanding and privacy protection?

To investigate how users perceive, understand, and respond to the privacy risks posed by chatbots in group chat environments, we conducted an online survey with 374 participants who have used chatbots on popular instant messaging platforms, including WhatsApp, Discord, Telegram, Viber, and LINE. Our study explores four core dimensions of user experience: (1) whether users are aware of the presence of chatbots in their group chats; (2) how accurately they understand the scope of chatbot data access, including message content and metadata; (3) what privacy concerns and behavioral responses they exhibit; and (4) what normative expectations they hold regarding what chatbots should be allowed to access.

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license visit https://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.
Proceedings on Privacy Enhancing Technologies 2026(1), 296–320
© 2026 Copyright held by the owner/author(s).
https://doi.org/10.56553/popets-2026-0016

To complement our user study, we conducted a systematic evaluation of chatbot implementations across three messaging platforms that officially support chatbots: Discord, Telegram, and LINE. We excluded WhatsApp and Viber as they lack official chatbot integration. Our analysis focused on assessing how well these platforms align with user expectations and privacy concerns identified in our survey. Through the analysis, we identify gaps between user expectations and current implementation practices. For instance, while users may expect chatbots to be clearly labeled in chat rooms, some platforms do not provide such indicators. Our findings highlight specific areas where design improvements could enhance user privacy without undermining chatbot functionality.

This work makes four key contributions:

(1) We show that many users have limited awareness of chatbots in their group conversations and misconceptions about their data access capabilities. Specifically, only 41.7% of participants correctly understood what messages chatbots can access. We identify six major mismatches between users' perceptions and the actual capabilities and behaviors of chatbots. These gaps increase users' privacy risks when interacting with chatbots.
(2) Based on our empirical findings, we systematize a five-stage model that describes how users detect, interpret, and respond to chatbot-related privacy risks.
(3) We evaluate current chatbot designs on Discord, Telegram, and LINE, three platforms officially supporting chatbot integrations, and find that some design choices conflict with user expectations emerged from our user study, especially in terms of transparency and privacy notifications. For example, Telegram does not disclose a chatbot's message access capabilities in the user interface until after the bot has been added to a group.
(4) We present eight design recommendations for improving chatbot transparency and control mechanisms.

2 Data Access Capabilities of Chatbots

Messaging platforms vary greatly in their support and regulation of chatbots in groups. This section examines the data access capabilities granted to chatbots on platforms mentioned in our study.

2.1 Platforms Officially Supporting Chatbots

Several messaging platforms officially support chatbots in group conversations, offering APIs and SDKs [10, 22, 37] for programmatic message exchange. To users, a bot appears as a regular group participant, with its messages visible in the conversation history.

Chatbots often have different privileges than human users. Three capabilities are particularly relevant from a privacy standpoint: access to message content, message metadata (e.g., sender's identifier), and group metadata (e.g., group members' identifiers). These capabilities vary by platform.

On LINE, chatbots are treated like regular users. They can read all messages, access metadata such as sender identifiers and profile images, and view the full list of group members and their profiles. There is no distinction in permissions between chatbots and users.

Telegram enforces stricter controls. By default, chatbots run in *Privacy Mode*, which limits access to messages that mention the bot or include predefined commands. Developers can disable this mode to allow full message access. In both modes, chatbots can access metadata such as user identifiers, profile images, and usernames. Although they cannot directly view the full member list, they can see active participants and gradually infer group composition.

Discord imposes the most restrictive default permissions. Chatbots are notified of all new messages by default, but can only access the content if the messages mention them or contain their registered command [17]. To access the full message content or the member list, chatbots must pass a platform-level review [11]. Some metadata, such as the sender's nickname, role, global identifier, and profile image, remains accessible, regardless of review status.

2.2 Userbots on Platforms Without Official Chatbot Supports

Some platforms, such as WhatsApp and Viber, do not offer official support for chatbot integration in group conversations. In these cases, programmatic participation is typically achieved using automated client applications, commonly known as *userbots*.

Since userbots are real user accounts operated by software, they have the same access as human users, including full message content, metadata, and group membership. Their detectability depends largely on how much they reveal their automated behavior. Open-source tools make building userbots easier [5, 26].

However, using userbots generally violates the Terms of Service of the platforms [23, 31]. These platforms have taken a strict stance against them. For instance, WhatsApp has explicitly warned that it may take legal action against automated or bulk messaging [24], and actively works to detect and ban userbots [25].

3 User Study

To understand the privacy risks of chatbots in group chat settings, we conducted a user study to explore the usage patterns and privacy perceptions associated with chatbots. We engaged with participants across demographic groups to understand the frequency with which they interact with chatbots in group settings and their awareness of the data access permissions these chatbots have.

Our user study aimed to investigate RQ1, which we further divide into five sub-research questions.

- **SQ1.** What do users use chatbots for, and in what contexts do they interact with them?
- **SQ2.** Are users aware of the presence of chatbots in group chats? Do they actively look for the presence of the chatbot?
- **SQ3.** Do users understand what the chatbot has access to?
- **SQ4.** How do users perceive the privacy risks of chatbots? How do they respond to these risks?
- **SQ5.** What are users' expectations about what chatbots should have access to?

3.1 Study Design

We conducted our study online, using a structured questionnaire distributed via Qualtrics. Most questions used 5-point Likert scales to measure participants' preferences and familiarity, and some open-ended questions allowed participants to justify or elaborate on their answers. The main questionnaire had five sections, preceded by a

comprehension check and followed by a demographic survey. The full questionnaire is presented in Appendix C.

To ensure that participants understood the study context, our questionnaire began with a comprehension check that asked them to distinguish between chatbots in group chats on instant messaging services (e.g., Discord moderation bots) and other chatbots not relevant to our study (e.g., ChatGPT). Participants who failed this check were disqualified and their responses were not recorded, as this indicated they likely misunderstood the consent form, which specified the type of chatbot being studied.

The first section (Appendix C.2) gathered data on participants' use of instant messaging services. Participants selected one messaging platform they used regularly (WhatsApp, Discord, Telegram, Viber, or LINE) and reported whether they had prior experience with chatbots in group chats. All subsequent questions focused on their experiences within that specific platform. The study focused on globally popular platforms [7] where chatbots are known to be present. These questions provided background on participant familiarity and filtered out those with minimal exposure to chatbots.

The second section (Appendix C.3) examined participants' awareness of chatbots in both private and public group chats. We asked whether they had noticed chatbots, how frequently they checked for them, and what types of group chats they had encountered chatbots in. To account for differences across social contexts, we classified group chats into two categories based on context and purpose: private groups, which involve close friends or family and focus on intimate conversations, and public groups, which are broader and often include unknown participants. Participants reported their behaviors in both contexts. This section provided insights into how users recognize and understand chatbot presence in different settings. The first and second sections together provided data for answering SQ1 and SQ2.

In the third section (Appendix C.4), we explored participants' knowledge about chatbot data access. We asked them to estimate message types, message metadata, and group-related information that chatbots could access. Afterward, we disclosed the actual data access capabilities of chatbots on their selected messaging platform and collected participants' reactions and privacy concerns. This allowed us to assess gaps in user understanding and address SQ3.

In the fourth section (Appendix C.5), we measured participants' privacy concerns about chatbot data access in both public and private group contexts. We asked participants about their privacy concerns and how the presence of chatbots affected their willingness to share sensitive or controversial information in group chats. This information helps to understand users' behavioral responses to perceived privacy risks and contributes to answering SQ4.

The fifth section (Appendix C.6) examined participants' expectations about chatbot data access. We asked what types of information they believed chatbots should be able to access and presented two specific scenarios to explore whether their preferences changed based on context. These responses helped address SQ5.

We validated the survey design through two rounds of pilot studies. In the first round, we recruited eight volunteers from our institution, excluding members of our research team, to complete the survey and provide feedback. In the second round, we recruited 30 participants from Prolific to further refine the survey. Based on their responses, we adjusted some questions for clarity and added a comprehension check at the beginning. This addition addressed the issue where some participants mistakenly assumed that the term "chatbot" referred to conversational agents such as ChatGPT.

Ethics. Our study was approved by the Institutional Review Board (IRB) of our institution. In addition, all participants were required to consent to the study online, confirming that they understood the scope of the study and their rights as participants. Participants could withdraw at any time without consequence.

3.2 Data Analysis

Quantitative Analysis. In addition to descriptive statistics, we applied statistical tests to identify significant differences and associations in the data. For paired ordinal data, such as changes in privacy concerns, we used the Wilcoxon signed-rank test to assess median differences. To measure correlations between self-reported knowledge and actual understanding of permissions, we used Spearman's rank-order correlation. All tests were two-tailed unless stated otherwise, with a significance level of $\alpha = 0.05$.

Qualitative Analysis. We conducted thematic analysis of responses to open-ended questions about chatbot interactions in group chats. This analysis was conducted by three researchers using inductive coding techniques. They iteratively developed and refined coding categories, which were then consolidated into broader themes. Regular meetings were held to resolve any discrepancies. Since three researchers met to resolve conflicts, we did not report the inter-rater reliability. The complete codebooks are provided in Appendix D.

3.3 Recruitment

We recruited participants through the Prolific platform. We established specific eligibility criteria for our participants: they must (1) be regular users of instant messaging services, (2) have seen or used chatbots on instant messaging services, (3) be at least 18 years old and meet the legal age of adulthood in the participant's location, and (4) be fluent in English. We did not restrict participant geography. Data was collected from October 4 to October 7, 2024. Recruitment continued until we reached the target sample size of 385, as determined by a power analysis (95% confidence level, 5% margin of error). In total, 482 participants were recruited, of which 374 provided valid submissions. Due to an administrative error, some invalid submissions were discovered only after data collection ended, leaving fewer valid responses than planned. As a result, the margin of error of our study is approximately 5.1%. The demographic of the participants are shown in Table 4 in the Appendix A. We excluded responses from individuals who failed any attention checks or had a reCAPTCHA score below 0.8. Specifically, 84 participants were disqualified due to failed attention checks, and 24 did not meet the reCAPTCHA score requirement. Some participants reported difficulties using Qualtrics' drag-and-drop features, leaving certain questions at default values and rendering their responses invalid. We identified 39 submissions likely affected by this issue.

Participants who completed the survey with at most one failed attention check received compensation of £3.50. Participants who

were screened out or experienced technical issues were compensated at a rate of £0.10 per minute spent on the survey. Among the valid submissions, the average completion time was 26.0 minutes.

4 Results and Findings

We collected valid responses from 374 participants.[1] Among all participants with valid submissions, the majority (244 participants, 65.2%) choose WhatsApp as their most commonly used messaging services that they have seen chatbots on. Discord (83, 22.2%) and Telegram (40, 10.7%) come the second and the third. Six participants (1.6%) reported Viber and one reported LINE. Regarding chatbot usage, 146 participants (39.0%) reported noticing chatbots in group chats on a daily basis, while 94 participants (25.1%) said they interacted with chatbots daily. The full results are shown in Fig. 1. Note that this may be an overestimate, as those who do not know they have encountered chatbots would not participate in the study.

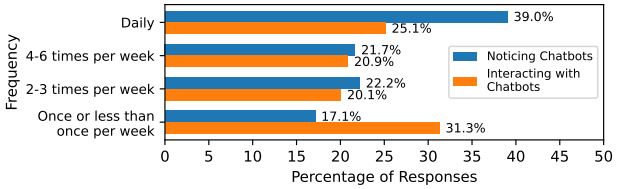


Figure 1: Frequency of noticing and interacting with chatbots

4.1 SQ1. Users' Experiences With Chatbots

To understand the social context in which participants use or encounter chatbots, we examined the types of groups in which they appear. We classified chat groups into private groups and public groups, and hypothesized that users would avoid using chatbots in private groups due to privacy concerns. However, surprisingly, our results show that users do use chatbots in private groups. When asked if they noticed chatbots in public and private groups, 327 participants (87.4%) reported noticing chatbots in public groups, while 204 participants (54.6%) noticed them in private groups.

We also conducted a thematic analysis of open-ended responses to understand how chatbots are used. The most common purposes were informational functions (e.g., Q&A support), group moderation (e.g., spam prevention, user verification), productivity tools (e.g., reminders), and entertainment (e.g., games). A detailed analysis and discussion can be found in Appendix B.1. Looking into the bots named by participants, we found that particularly those for moderation, require full access to messages and metadata, while others operate with minimal data. This variation underscores the importance of user awareness and accurate mental models regarding chatbot data access and privacy implications.

[1]The analysis script, anonymized dataset, and qualitative coding results are available at https://github.com/csienslab/bot-among-us.

Key Takeaways for SQ1

- Users interact with chatbots even in private groups, which often consist of close contacts such as family and friends.
- Chatbots in group chats serve various purposes, primarily productivity tasks and group moderation, highlighting the need for different levels of access.

4.2 SQ2. User Awareness & Checking Strategies

Given privacy concerns, it is essential for users to be aware of the presence of chatbots in group chats. This awareness can be achieved either by users actively seeking out chatbots or by ensuring that the presence of chatbots is made clear. We aimed to understand whether users consider the potential presence of chatbots, how they detect them, and the reasons behind their actions, or lack thereof, when verifying the presence of chatbots. In our user study, we examined these behaviors in both private and public group chat settings.

4.2.1 When to Check the Presence of Chatbot. When asked if they check for the presence of chatbots when joining a new group, less than half of participants said they always or usually check, while about 10% said they never check, regardless of whether the group is private or public. Participants were also asked when they typically check for chatbots. Besides checking when first joining a group (about 75% for both public and private groups), another common trigger was when a bot is added (about 50%), as shown in Fig. 2. Interestingly, we found no significant differences in chatbot checking behavior between private and public groups, even if the presence of chatbots in private groups may raise greater privacy concerns.

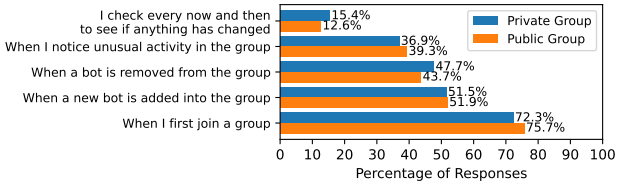


Figure 2: Percentage distribution showing user behavior in checking for chatbots across private and public groups.

4.2.2 Why Not Check the Presence of the Chatbot? In the open-ended questions, we ask participants to describe their chatbot checking behavior. Through thematic analysis, we identified several reasons why users choose to check or not check for the presence of chatbots in a group. While we present the analysis of why users check for chatbots in Appendix B.2, our primary interest lied in the opposite question: why do some users choose *not* to check for chatbots? Since failing to do so can lead to uninformed decisions, it is crucial to understand the rationale behind this choice.

Lack of Concern About Chatbots. One of the most straightforward reasons is that some participants are not particularly concerned about the potential threats posed by chatbots. For instance, P79 states, "I do not consider them a threat and I will continue to chat freely despite a chatbot being in use."

Low Priority. For some participants, identifying chatbots was simply not a priority. P102 explained, "my priority is engaging with fellow human members." P106 also said they join groups "to interact with real human beings," not to focus on chatbots.

Ubiquity of Chatbots. Participants reported that chatbots have become so common that they no longer feel the need to actively check for their presence. As P85 observed, "I know they are there so I do not look or check for them, [...] their presence is so common that I often don't register their presence."

This ubiquity may create a habitual blind spot, particularly in public groups. Users could become desensitized to the presence of chatbots, losing opportunities to notice and avoid malicious ones, even when clear indicators are available.

Chatbots Are Easy to Notice. Some participants felt no need to check for chatbots, believing they are easily identifiable. Participants mentioned that chatbots often greet new members or respond immediately to certain messages. As P308 noted, "I don't really go out of my way to check for chatbots [because] they are normally noticable [sic]." Moreover, chatbots on some platforms are explicitly labeled. P310 shared, "on the discord interface I immediately see if there is a chatbot or not."

However, this assumption is likely flawed. Some chatbots, whether adversarial ones eavesdropping on conversations or benign ones moderating groups, may remain silent most of the time, making them undetectable to users who assume chatbots are always conspicuous. This reflects a mismatch between users' expectations and the actual visibility of chatbots, leading users to potentially share sensitive information before realizing a bot is present.

Trust. Some participants do not verify chatbots because they trust that other group members or administrators are effectively managing the situation. For example, P77 stated that they are comfortable with the presence of chatbots because "I trust my friends to choose safe to use bots."

Key Takeaways for SQ2

- A substantial portion of users (about 30%) of users do not check for the presence of chatbots when joining a group.
- Some users mistakenly assume that chatbots are always easy to spot, making them less aware of more "quiet" chatbots.

4.3 SQ3. Understanding of Chatbot Permissions

Knowledge about what chatbots have access to is crucial for group members to make informed decisions on what to share in the group. To understand how much users know about chatbot's permission, we asked the participants to self-report their knowledgeability of chatbot permissions. Then, we ask users several questions about the permission of chatbots on their most commonly used platforms. We ask what kinds of messages, what message metadata, and what group metadata chatbots have access to. These shed light on whether users have correct understanding and what the most common misconceptions are.

4.3.1 User's Understanding and Common Misunderstandings. We ask users to report their understanding of what chatbots can access on their chosen platforms. Specifically, we examine their perceptions of what types of messages, message metadata, and group metadata chatbots can access. We then discuss common misunderstandings and their implications.

Message Type. When asked what kinds of messages a chatbot can access, only 41.7% (156) of participants answered correctly. On WhatsApp and Discord, the most common incorrect response was *only messages needed for the normal operation of the chatbot.* This assumption seems reasonable, as users without technical knowledge about chatbots might expect them to access only the messages necessary for their function. However, this reflects a concerning mismatch between users' perceptions and the actual scope of chatbot access, as users tend to underestimate data exposure due to their intuitive but inaccurate understanding of how chatbots operate.

Another common misconception, especially among Telegram users, is the belief that all messages are sent to the chatbot, except for those that are later deleted. In reality, once a message is sent, the chatbot retains access to it, and there is no mechanism in the current paradigm to force the chatbot to "forget" received messages. This reflects a mismatch between user expectations and actual data retention practices. Users may mistakenly believe that deleting a message removes all traces of its existence, including from the chatbot's memory, which is not the case.

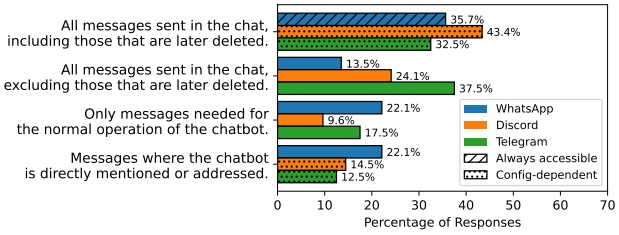
Message Metadata. Regarding the second question about what message metadata chatbots can access, most participants correctly understand that chatbots can read message contents. However, other types of information are less obvious to them. The most commonly overlooked detail is user profile information. Only about 28% of WhatsApp and Discord users correctly recognize that chatbots have access to their profiles. Another frequently missed piece of information is the sender's username or user identifier. While only 51.2% of WhatsApp users correctly recognize this, a much higher percentage, 85.5%, of Discord users do.

The mismatch between users' expectations and the actual extent of chatbot access to message metadata suggests that many users mistakenly believe chatbots cannot link messages to their identity. As a result, they may assume they are anonymous and become more willing to share sensitive information, even though chatbots on many platforms do have access to sender information.

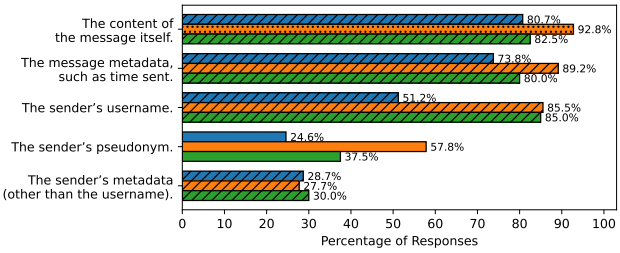
Group Metadata. Finally, in the third question, we asked what group metadata chatbots can access. Most participants correctly recognized that chatbots know the group name. However, other types of information are less obvious. The most commonly overlooked detail is the profile information of group members. Only 39.8% of WhatsApp users and 32.5% of Telegram users realize that chatbots can access group members' profile pictures.

4.3.2 User's Self-Reported and Actual Knowledgeability. Before showing participants the chatbot permissions of selected platforms, we asked them to rate their knowledge of chatbot permissions using a Likert scale. Alarmingly, nearly half (46.8%) of the participants indicated that they were either only slightly knowledgeable or not knowledgeable at all. Only 13 participants (3.5%) considered themselves extremely knowledgeable about chatbot permissions.

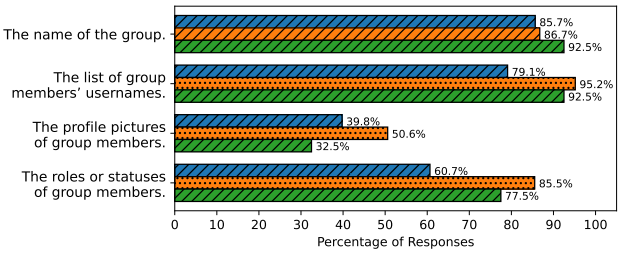
Next, we examined whether participants' self-reported knowledge, which reflects their confidence in understanding chatbot



(a) User perceptions of the types of messages accessible to chatbots on different platforms.



(b) User expectations regarding the scope of chatbot data access.



(c) User perception of group metadata accessible to chatbots on different platforms.

Figure 3: User understanding of chatbot data access across message types, message metadata, and group metadata.

permissions, correlated with their actual understanding. To do this, we checked whether their beliefs matched the actual chatbot permissions on their chosen platform. For the first question about what types of messages chatbots can access, we simply checked whether the participant answered correctly. For the next two questions, we calculated the number of options they correctly selected or left unchecked. Because some platforms offer different permission configurations (e.g., Telegram's Privacy Mode), we selected the mode that best matched the participant's response. We then used Spearman's rank correlation to analyze the results.

For the first two questions, which addressed message types and message metadata, we found no positive correlation between self-reported and actual knowledgeability. The correlation coefficients were $\rho = -0.058$ ($p = 0.27$) and $\rho = 0.079$ ($p = 0.13$), respectively. That is, we did not find evidence showing that participants' confidence about their understanding of chatbot permissions matches the reality. However, for the third question on message group metadata, there was a significant correlation between self-reported and actual knowledge ($\rho = 0.15$, $p < 0.01$). More details about the statistical inferences are in Appendix B.8.

Key Takeaways for SQ3

- Users often underestimate what data chatbots can access, especially which messages are shared with them.
- Nearly half of the users report low confidence in their understanding of chatbot data access, and we found no evidence showing that their perceived knowledge match actual understanding of message content and metadata access.

4.4 SQ4. Privacy Concerns About Chatbots

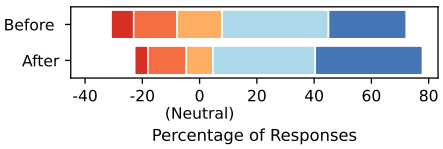
Understanding users' privacy concerns about chatbots would provide a guideline for improving the privacy design. In our questionnaire, we first ask the participants to self-report their concerns without providing further information about chatbots. Then, we present the chatbot permissions, and ask the participants about their comfort with these designs.

We made a potentially misleading statement when describing chatbot permissions on Discord. In our presentation, we stated that Discord bots have access to all messages. In reality, bots can receive all messages by default, but they cannot access message content unless they pass a platform-level review [11]. Without this review, bots can only read the content of messages that either mention them directly or use pre-registered commands. Because our wording may have led participants to believe that chatbots could access all message content, we excluded Discord responses from our quantitative analysis of privacy concerns. For the qualitative analysis, we included Discord user responses only when they were unrelated to whether bots have full message access. In Appendix B.4, we present a sensitivity analysis showing that removing Discord user data does not affect our main conclusions.

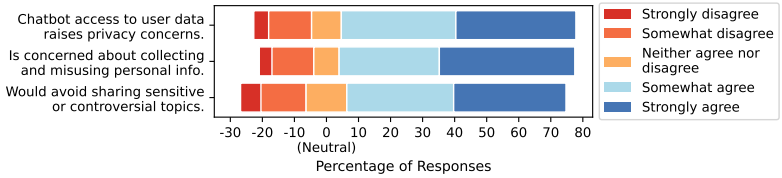
4.4.1 Privacy Perception. Our survey shows that many users are concerned about chatbot privacy in group chats. Over two-thirds (73.2%) somewhat or strongly agree that chatbots accessing user data raises privacy issues. Similarly, 73.5% somewhat or strongly agree that they are concerned about the collection and misuse of personal information shared in group chats. These results suggest that users have privacy concerns and become more cautious once they understand what chatbots can access.

To measure how participants' perceptions changed after learning about chatbot permissions, we asked the same question before and after presenting the permissions: *Chatbots from group chats with access to users' data as checked above would raise my privacy concerns.* As shown in Fig. 4a, privacy concerns increased after participants learned about the permissions. Initially, 64.3% somewhat or strongly agreed with the statement. Afterward, this rose to 73.2%, with strong agreement increasing from 27.1% to 37.5%. A Wilcoxon signed-rank test showed that the change in concerns exists ($p < 0.001$, $r = 0.308$, 95% CI = $[-2, 3]$).

4.4.2 Thematic Analysis of Privacy Concerns. To understand what users believe the privacy concerns are, we asked participants to



(a) Privacy concerns before and after learning chatbot permissions.



(b) Privacy concerns about chatbots.

Figure 4: User privacy concerns measured on a 5-point Likert scale from "strongly disagree" (left) to "strongly agree" (right).

describe their privacy concerns in response to an open-ended question. Four major themes emerged: lack of transparency, excessive or sensitive data collection, data misuse (including data sales, scams, advertising, and legal repercussions), and risks of data breaches. A summary of the thematic analysis is presented in Table 1, and a more detailed discussion is provided in Appendix B.5.

4.4.3 Responses to Privacy Concerns. We investigate how users respond to the privacy concerns of chatbots. After learning the chatbot's potential privacy threats, we ask the participants whether they would continue using chatbots and why. Many participants (42.3%) claimed that they would continue using chatbots.

We also examine whether users would change their behavior after learning about potential privacy threats from chatbots. We focus on self-censorship, i.e., avoiding controversial topics or sensitive information to reduce risk. As shown in Fig. 4b, 68.4% (199 participants) reported they would stop talking about controversial topics due to the presence of chatbots in the group.

For participants who continued using chatbots, we asked about their reasoning. By thematic analysis, we find some recurring patterns on rationale to continue using chatbots and mitigation to the privacy concerns.

Utility. Unsurprisingly, many users continued using the chatbot despite recognizing potential risks, as they found its practical value outweighed their privacy concerns.

Privacy Cynicism. Some users are less concerned about privacy because they believe their personal information is not valuable or that privacy is not important enough for them to care about. Several participants simply stated they have nothing to hide.

Privacy Helplessness. Some participants expressed their feelings of powerlessness about privacy online. They believe that they cannot avoid privacy risks and that data collection is inevitable. For example, P250 commented, "there is nothing I can do to stop the chatbots from accessing this information for as long as I still need to use the app, it's just a matter of acceptance." P73 echoed similar thoughts by saying that "my personal information is being used by so many companies already [...] I became numb to that because otherwise I would have to worry about everything I do online."

Forced Acceptance. Some participants reported privacy resignation since they feel they have no way to avoid chatbots. P63 remarked, "if you want to participate in public servers, you can't opt out of chatbots." This issue is exacerbated by a mismatch between group members' expectations of mutual consent and the actual ability of administrators to add chatbots without input from others. Unlike general privacy helplessness regarding online surveillance, this resignation specifically relates to group participation design and the lack of user agency within group settings.

Groups as Non-Private Context. Some participants mention that groups are inherently non-private contexts, so chatbots reading the messages is not an issue. For example, P71 stated, "If I am writing a message to a group, I assume it can be public knowledge, so I only share information about me that I am okay with people knowing."

However, this assumption may overlook chatbots' ability to aggregate and infer sensitive information, even from seemingly trivial data [34]. This false sense of control can make users underestimate how user profiling systems can compromise their privacy in ways that are not immediately apparent.

Changing Behaviors as Risk Mitigation. Some users indicated they would adjust their behavior to mitigate privacy risks. This involves being more careful about the information they share. As P131 noted, "I will be [aware] what I should share on group chats and what not to share." P192 also mentioned the importance of awareness and making informed decisions, stating that "the user will understand what should be shared versus what should not if the user fears a privacy breach."

Trust as Risk Mitigation. Some users continue using chatbots because they trust the platform or developers. They might believe that the developers would not do harm. For example, P346 mentioned "I believe my information is confidential and I trust they will keep my information private." Besides trusting the developer, P261 also mentioned that they trust that the group admins of "official and recognized groups" would take good care of their privacy.

Transparency as Risk Mitigation. Some participants mention that they believe the privacy risk can be mitigated by transparency. For example, P89*[2] mentioned that "I don't have a problem [...] as long as there is a form of disclosure before i join the group." P291* also mentioned "if the chatbot is clear about its permissions and intentions, [...] I wouldnt[sic] mind it."

[2]An asterisk denotes a Discord user whose response was likely unaffected by our potentially misleading statements.

Table 1: Thematic Analysis of User Privacy Concerns About Chatbots

Theme	Description	Example
Lack of Transparency	Users are confused about what data chatbots collect and how it is used.	"I do not understand how they protect our data or use it to be more specific besides helping us." – P95
Excessive or Sensitive Data Collection	Chatbots may collect more data than necessary, including sensitive information, often without clear consent.	"I feel like they should [...] not collect data at the time when they weren't called on to be used." – P16
Data Misuses: Selling to Third Parties	Fear that chatbot providers might sell user data to external companies or data brokers.	"I have issues if the data collected is [...] later sold out to data brokers." – P216
Data Misuses: Scams	Concern that chatbots could steal sensitive data for scams, phishing, or identity theft.	"A chatbot that may be stealing your passwords or committing identity theft." – P195
Data Misuses: Ads and Ad-motivated Tracking	Chatbots tracking conversations to create profiles for targeted advertising.	"Only concern for me is whether or not my information is collected and used for advertisement purposes." – P196
Data Misuses: Censorship and Legal Enforcement	Concern about governments using chatbot-collected data to monitor, target, or prosecute people.	"Having data that can be subpoenaed by authorities to prosecute people." – P158*
Data Breaches	Chatbots can be hacked, resulting in a data breach.	"it could be a target for hacking." – P300

Key Takeaways for SQ4

- Users show greater privacy concerns after learning about chatbot permissions, indicating low initial awareness.
- Users attempt to mitigate privacy concerns through behavioral changes, relying on trust in developers or platforms, and seeking transparency about chatbot permissions.

4.5 SQ5. User Expectation of Data Access

In Sec. 4.3, we showed that users' expectations about chatbot access permissions often differ from actual capabilities. To explore this discrepancy further, we examine what users believe chatbots should be able to access if users controlled these permissions. Following our previous structure, we discuss three categories: message content, message metadata, and group metadata. We asked participants to select which message types they would allow chatbots to access and to rank their preferences for message and group metadata access. To understand how preferences change in different contexts, we also designed two scenarios and asked the same questions.

4.5.1 *General Expectations.*

Message Type. Our results show that most users (41.6%) prefer chatbots to access only the messages necessary for their normal operation. This is consistent with our initial hypothesis that users want chatbots to operate with the least amount of access required. However, a significant portion (31.1%) prefer chatbots to access messages only when they are directly mentioned. This suggests that while users value privacy and limited access, they also want the autonomy to control chatbot interactions, allowing them to decide when and how chatbots engage in conversations. In contrast, only 14.5% of participants prefer full access to all messages, indicating a general reluctance to grant broad permissions. A visualization of the responses is provided in Fig. 11 in Appendix B.6.

Message Metadata. Our results show a strong preference for limiting chatbot access to message metadata. Participants were most comfortable with chatbots accessing basic metadata, such as timestamps, but strongly opposed access to more sensitive information, such as sender metadata. A visualization of the responses is provided in Fig. 12 in Appendix B.6.

Interestingly, users ranked real usernames as more acceptable for chatbot access than pseudonyms, even though pseudonyms provide greater privacy by preventing tracking across groups. We suspect this is because pseudonyms feel less intuitive to users, and their privacy benefits are not immediately clear. Beyond usernames, other sender metadata such as profile picture or contact information was the least preferred, highlighting strong concerns about exposing too much personal information to chatbots.

Group Metadata. When it comes to group metadata, users generally accept chatbot access to basic information such as the group name. Participants are more comfortable with chatbots accessing a list of group members' usernames, but far less willing to allow access to more personal information, such as profile pictures. These findings suggest that while users recognize the need for some access to group metadata for chatbot functionality, they prefer to limit exposure to personally identifiable details. A visualization of the responses is provided in Fig. 13 in Appendix B.6.

4.5.2 *Influences of Contexts.*

To understand how different contexts affect user expectations, we designed two specific scenarios and asked participants to answer the same questions again. The first scenario, Reminder, allows users to add items in a specific format that includes a time and a message. The chatbot then sends reminder messages to group members at the specified time. The second scenario, Moderation Helper, is a set of tools that help with group moderation. For example, it sends welcome messages to new members, silences or bans users who send spam, and tracks member activity for future review. These two scenarios were chosen based on popular Discord chatbots and represent two extremes: the Reminder bot requires minimal permissions, while the Moderation Helper bot requires access to nearly all messages and some account identifiers to function properly.

For message access, most participants believed chatbots should access only messages relevant to their functionality in both scenarios, although no current platforms support content-sensitive permission models. Participants' second-ranked choices aligned with permissions required under current platform designs. In the Reminder scenario, participants expected the chatbot to access only messages where it was explicitly tagged. In the Moderation Helper scenario, they expected the chatbot to have access to all messages.

For message and group metadata, while slight differences in preference ratios existed across scenarios, overall trends remained consistent. Generally, participants were more comfortable with chatbots accessing non-personal metadata. Access to personally identifiable metadata, even for moderation purposes, was consistently the least preferred option.

Key Takeaways for SQ5

- Users prefer chatbots to access only the minimum necessary data, with a clear bias against full message access.
- Even when chatbot functionality varies, users' preferences consistently favor restricting access to sensitive metadata.

4.6 Privacy Expectation Mismatches

Based on our user study, we find that users' expectations about chatbot data access often diverge from the actual capabilities and behaviors of chatbots. To systematically conceptualize these gaps, we draw upon Contextual Integrity (CI) theory [28], which defines privacy as the appropriate flow of information, governed by actors (senders, recipients, subjects), attributes (types of information), and transmission principles (constraints on flow).

We identify six major types of expectation mismatches, each representing a distinct violation of CI. These mismatches reflect the most prominent gaps revealed in our study, including: what types of messages chatbots can read, whether consent is obtained from all group members, how long data is retained, whether chatbots can identify senders, whether they can track users across groups, and whether their presence is clearly visible. Table 2 summarizes these mismatches and their corresponding CI violations. While not exhaustive, these six categories provide a systematic framework for understanding chatbot-related privacy risks from a user perspective and inform future research, design, and policy efforts that aim to better align chatbot functionality with user privacy expectations.

5 Platform-Design Analysis

After examining how users perceive, understand, and respond to the privacy risks posed by group chatbots, we turn to RQ2: to what extent do current chatbot platform designs align with users' understandings and expectations, where any misalignments occur, and how these gaps can be addressed.

We begin by introducing Group-Chatbot Privacy Adaptation Pathway in Sec. 5.1, which captures users' privacy behaviors in Sec. 4. This five-stage model captures how users detect, interpret, and respond to chatbot-related privacy risks. Using this model, we then examine how current platforms guide users through these stages in Sec. 5.2. Finally, in Sec. 5.3, we present design recommendations to better support user privacy in group chats.

5.1 Adaptation Pathway

To systematically guide platform design interventions, we present a five-stage process model that captures key points at which users may either progress toward privacy-protective behaviors or experience breakdowns. The pathway identifies critical stages from initial bot awareness to eventual behavioral adaptation. Each stage is grounded in empirical evidence from our user study, and we explain how user behaviors and misconceptions map onto each stage. This allows us to identify where users fail to act on privacy concerns and how design interventions can help, as we will discuss in Sec. 5.3. This pathway also aligns with the Security & Privacy Acceptance Framework (SPAF) [9], which posits that three interdependent factors, namely awareness, motivation, and ability, must be addressed to promote the adoption of security and privacy practices. Table 3 summarizes the stages and provides examples of failure scenarios. In the following, we discuss how each stage appears in our user study and how each stage aligns with SPAF's three factors.

Awareness Trigger. The first prerequisite for privacy adaptation is noticing that a chatbot is present in the group. Our user study indicates that this trigger often fails: about a quarter of users reported not checking for bots when joining new groups (§4.2.1). Moreover, many users incorrectly assumed that bots would always be obvious, leading to complacency (§4.2.2). This stage corresponds to SPAF's *awareness* category, which refers to whether individuals understand the threats relevant to the data they wish to protect and the recommended security practices.

Capability Understanding. When noticing a chatbot, users must correctly understand what the bot can access. Our results show that this step is highly error-prone: only 41.7% of participants accurately identified the types of messages that bots could access, with widespread underestimation of bot capabilities (§4.3), reflecting an incomplete or folkloric understanding of data flows in group chats. Inaccurate mental models blunt users' ability to perceive privacy risks, reducing the likelihood of protective action. This stage corresponds to SPAF's *awareness* component.

Risk Perception Activation. Accurate mental models must then activate privacy risk perception. We found that after learning about chatbot permissions, the proportion of participants expressing privacy concern rose significantly, indicating that risk perception can be effectively triggered through transparent information. However, failures were common, as some participants perceived little or no risk or normalized privacy risks (§4.4). This stage aligns with SPAF's *motivation* component, which reflects whether individuals are willing to adopt recommended practices to protect their data.

Coping Decision. Upon perceiving risk, users must decide whether to take protective action. More than a half of participants reported that they would adjust their behavior in response to chatbot presence, such as intentional self-censorship and privacy-conscious messaging. However, our study also surfaced multiple failure modes at this stage, such as privacy resignation and forced acceptance due to social dynamics (§4.4.3). This decision-making process is influenced by SPAF's *motivation* factors, such as subjective norms and perceived relative advantage.

Behavioral Response. Finally, users must turn protective intentions into action. Some participants reported adjusting their information-sharing practices after learning of the presence of chatbots, and some participants claimed that they plan to do so (§4.4.3). However, without longitudinal data or in-lab experiments, we cannot assess whether these efforts were successful. This stage corresponds to SPAF's *ability* category, which concerns users' capacity to carry out recommended privacy practices.

Table 2: Chatbot access expectation mismatches across messaging platforms in our user study.

Mismatch Type	User Expectation	Mismatch Platform Design	Contextual Integrity Violation
Scope	Bots only listen when directly addressed or for relevant tasks.	*Discord, Telegram*: Default limited, configurable full access; *WhatsApp, Viber, LINE*: Always full access.	*Violation of transmission principles*: data flows beyond what users deem appropriate
Consent Model	Group members expect mutual consent before bot inclusion.	*All platforms*: Admins can add bots without consent from all group members.	*Change in sender-recipient structure without appropriate consent*
Data Retention	Bots process input temporarily and only retain undeleted messages.	*All platforms*: Bots may store messages indefinitely without user awareness.	*Violation of expected data retention norms*: flows exceed normative timeframes
Identity Exposure	Bots only see message content, not metadata or personal identifiers.	*All platforms*: Bots have access to sender identifiers such as usernames.	*Violation of information type and subject integrity*
Cross-Context Identity	Bots are group-specific and do not track users across contexts.	*All platforms*: Users have persistent identifiers across groups.	*Violation of contextual boundaries*: messages flow across groups unexpectedly
Visibility	Bots should be visible when active and accessing content.	*Telegram, LINE*: Bots listed but sometimes unclear; *WhatsApp, Viber*: No clear bot visibility; *Discord*: Meets expectation.	*Violation of transparency norms*: hidden data recipients undermine user trust

Table 3: Group-chatbot privacy adaptation pathway

#	Stage	Ideal User Behavior	Failure Scenario
1	Awareness Trigger	Notices chatbot in the group chat	Unaware of bot; shares sensitive info
2	Capability Understanding	Understands bot's access permissions	Assumes bot cannot see sensitive messages
3	Risk-Perception Activation	Perceives realistic privacy risk	Believes there is "nothing to hide"
4	Coping Decision	Chooses to take protective action	Feels powerless and accepts risk passively
5	Behavioral Response	Adopts behavior to protect privacy	Has concerns but takes no action

5.2 Analysis of Current Platform Designs

We examine whether common messaging platforms offer enough cues to help users pass the five stages. Our analysis focuses on Discord, Telegram, and LINE, which explicitly support chatbots and appeared in our user study, as well as userbots on platforms that do not officially support chatbots.

5.2.1 *Discord.*

Awareness Trigger. Discord provides clear visual indicators to indicate the presence of chatbots. In the chat history and member list (Fig. 5a and Fig. 7b), chatbot usernames are labeled with "APP," clearly distinguishing them from human users. However, we discovered a bug: if the name of the chatbot is too long, the "APP" label disappears (Fig. 5a). A malicious chatbot could exploit this to hide its identity by using an excessively long username. This bug was fixed before we reported it to Discord.

Capability Understanding. Discord clearly explains a chatbot's permissions when the chatbot is added to a group, as shown in Fig. 7a. This helps users immediately understand what access they are granting. Additionally, users can review detailed chatbot permissions later, as shown in Fig. 7b, although to the best of our knowledge, this feature is only available in Discord's desktop client.

When a user types a message in Discord, the auto-completion for pre-registered commands indicates which chatbots will receive the message. However, aside from this feature, there are no other visual cues to remind users that chatbots can read their messages.

Discord does not clearly explain how deleted messages are handled by the platform or chatbots, nor does it disclose its limitations in revoking a chatbot's storage. This lack of transparency can lead to misunderstandings about the technical limits of message deletion.

Risk Perception Activation. As mentioned above, Discord provides explanations of chatbot permissions. However, to the best of our knowledge, there is no other information to help users assess privacy risks. Users have no idea how much information has been accessed by chatbots or how sensitive the information is.

Coping Decision. Users can adjust chatbots' permissions with fine granularity. They can also choose to remove chatbots entirely.

5.2.2 *Telegram.*

Awareness Trigger. Telegram does not show explicit indicators for chatbots in the chat history, making it harder for users to identify chatbot-generated messages. However, it differentiates chatbots in the member list by providing additional context under each chatbot's username, such as "has access to messages" or "has no access to messages," as shown in Fig. 5d.

Capability Understanding. Telegram does not inform users about a chatbot's permissions when it is added to a group. Users can only see whether the chatbot has access to all messages by checking its details in the member list after it has been added, as shown in Fig. 5d. If the chatbot has administrator permissions, Telegram labels it as "Admin," just like human administrators.

When users send messages, Telegram shows autocomplete suggestions for pre-registered commands, which also reveal which chatbots may receive the message. However, it offers no other indicators to show that chatbots are reading messages.

Telegram also fails to clearly explain the limits of message deletion. Even worse, chatbots cannot detect when messages are deleted, so even developers who want to honor users' deletion requests cannot do so, as Telegram does not notify chatbots of message deletions [36].

Risk Perception Activation. We are not aware of any additional information from Telegram to help users assess privacy risks.

Coping Decision. Telegram does not allow users to change the permissions of chatbots, except for granting or removing administrator permissions. Users can always choose to remove chatbots from the groups entirely.

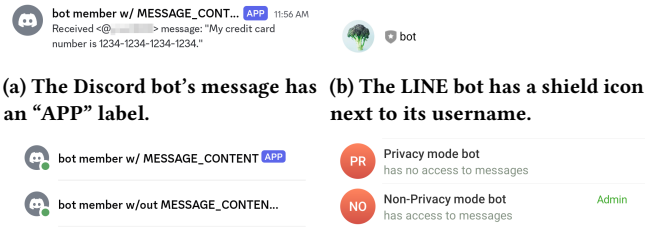
5.2.3 *LINE.*

Awareness Trigger. LINE does not provide explicit visual indicators in the chat history to distinguish chatbots from regular users. It displays a shield icon next to chatbot usernames in the group member list, as shown in Fig. 5b, but does not offer an explanation of the icon's meaning. Users can visit the chatbot's profile page, which is clearly distinct from that of a human account. This distinction assists users in identifying the member as a chatbot.

Capability Understanding. LINE does not display permission information in the user interface, leaving users unaware of what chatbots can access. This is expected as chatbots in LINE have the same permissions as regular group members. LINE provides no indication before sending messages to inform users whether chatbots can access the messages, nor clear explanations about the limitations of message deletion.

Risk Perception Activation. As with Discord and Telegram, we are not aware of any additional information provided by LINE to help users evaluate privacy risks.

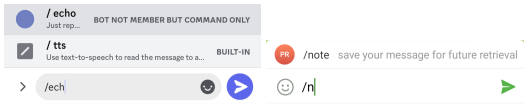
Coping Decision. Users cannot modify chatbot permissions. However, they can always choose to censor their own messages or remove chatbots from the group entirely.



(a) The Discord bot's message has an "APP" label. **(b) The LINE bot has a shield icon next to its username.**

(c) Discord bots are clearly labeled in the member list, but these labels may be mistakenly hidden. **(d) Telegram bots have their permissions written under the username.**

Figure 5: Indicators for chatbot presence across platforms.



(a) Discord **(b) Telegram**

Figure 6: Command auto-completion signals that the message will be processed by the listed chatbots.

5.2.4 *Userbots.*

Awareness Trigger. From a normal user's perspective, a userbot appears identical to a typical user because it is essentially a user client controlled by an automated program. As a result, it is difficult for users to realize that a "user" is actually a bot. Unless the userbot explicitly reveals its identity or displays obvious automated behavior, it can remain unnoticed.

Capability Understanding. Users are rarely informed about the capabilities of userbots. Since platforms treat userbots as ordinary users, they do not disclose warnings about what these bots can access beyond what they typically show when a normal user is added. In fact, aside from actively detecting and banning them, platforms have few ways to make userbots' data usage more transparent.

Risk Perception Activation. As noted above, platforms cannot provide additional explanations about userbots, and there is little other information to help users evaluate potential privacy risks.

Coping Decision. The only available action for users is to remove the userbots.

5.3 Design Recommendations

Based on our user study and platform analysis, we propose design recommendations to help messaging platforms align chatbot designs with user privacy expectations. We group these recommendations according to the stages they support.

5.3.1 *Awareness Trigger.*

Implement Visible Indicators for the Presence of Chatbots. Chatbots should always be clearly identifiable to users. Platforms should implement consistent visual cues distinguishing chatbots from regular users. For example, Discord effectively marks bot messages with specific labels and identifies bots in the group member list. This visibility helps prevent misconceptions like the belief that chatbots are always noticeable without explicit indicators (§4.2.2).

5.3.2 *Capability Understanding.*

Provide Transparent Permission Explanations. When a chatbot is added to a group, platforms should present clear, non-technical summaries of accessible information. Our study revealed participants often misunderstood what messages chatbots could read. Platforms should implement permission panels or privacy nutrition labels [20] that include: (1) message access permissions; (2) collected metadata; and (3) privacy and data retention policies. Discord demonstrates this approach, as shown in Fig. 7a. This information should be available when adding a chatbot, when new members join groups with chatbots, and when users interact with a chatbot's profile.

Clear Indicators of Chatbot Message Access Before Sending. Platforms should show clear visual indicators of which chatbots can access messages before they are sent. Our study found that users often assume chatbots only see directly mentioned messages or those needed for their function. Displaying access indicators in the message composition interface would help users make informed decisions. While Telegram and Discord offer autocomplete hints, these are meant for convenience, not privacy, and they do not warn users when chatbot commands are not used. A clear, privacy-focused indicator would help resolve confusion about chatbot access.

Clarify Message Deletion Limitations. Platforms should clearly communicate that deleting messages does not remove them from chatbot storage. Some participants mistakenly believe deleted messages become inaccessible to chatbots, indicating a misconception where users see deletion as state-changing rather than visibility-changing. Platforms should address this by warning users when they delete messages in groups with chatbots.

5.3.3 Risk Perception Activation.

Provide Privacy Reports and Privacy Nudging. Messaging platforms could offer privacy reports that summarize what information chatbots have accessed over time, along with occasional alerts to nudge group members about this access. Almuhimedi et al. [1] showed that privacy reports and permission review nudges for mobile apps increased user awareness of data collection. Similarly, messaging platforms could show how many messages were accessed, which messages, how many profiles, and what specific user data chatbots obtained, whether passively received or actively queried. These features would help users understand privacy risks and make informed choices, such as limiting bot access or removing them.

5.3.4 Coping Decision.

Implement Adaptive Permission Models Based on Content and Context. Messaging platforms should adopt permission models that consider both the content and context of group conversations. Our user study shows that users sometimes interact with chatbots in private contexts but may self-censor due to privacy concerns. Content-sensitive models can identify relevant messages (e.g., "meeting at 3pm tomorrow" for a reminder bot) while ignoring unrelated ones. Meanwhile, context-sensitive features like a privacy mode can adapt to different conversational settings. Message parsing can be implemented either on the server side or on the client side when end-to-end encryption is used. These designs reduce the need for manual tagging and give users better control over their privacy.

5.3.5 Others.

The following recommendations fall outside any specific stage of the pathway.

Recognize That Consent in Group Chats Is Collective. Platforms should recognize that privacy in group chats is a shared concern. Participants expressed concern over limited control when other members add chatbots, highlighting that one person's actions can affect everyone's privacy. While involving the whole group in such decisions would be ideal, it is rarely practical. Instead, platforms should notify all members when a chatbot is added, explain its permissions, and allow users to raise concerns or opt out of data collection. This helps ensure privacy is respected collectively, rather than overridden by a single decision-maker.

Provide Official Support for Chatbots. Messaging platforms should offer official chatbot integrations with clear privacy protections. Without such support, chatbots may operate unofficially, lacking proper oversight and transparency. Moreover, in Appendix B.3, we show that users on Discord, which officially supports chatbots, demonstrated better understanding of access capabilities than users on platforms without such support. By recognizing user demand and providing official support, messaging platforms can help developers build chatbots that better respect user privacy.

Actively Detect and Ban Userbots. As discussed in Sec. 5.2.4, userbots present unique challenges. Users may struggle to distinguish between a "user" and a bot, and platforms have limited abilities to restrict userbots' access to data or enhance data transparency. Considering the difficulties of governing userbots, we agree with WhatsApp's policy about banning userbots [24, 25] and recommend that all platforms actively detect and potentially ban userbots from their services.

6 Discussions

6.1 Three Paradigms

Through user studies and analysis of chatbot designs across platforms, we identified three common ways to frame chatbots: (1) algorithmic group member, (2) single-user service, and (3) multi-user service. Each framing highlights different aspects of chatbot design. We argue that while the first two framings are common, they may overlook some privacy considerations. Instead, the multi-user service framing should be adopted when designing chatbots.

Algorithmic Group Member. In this view, chatbots are seen as regular participants in conversations. For example, P257 in our study said, "I think they [chatbots] count as another member of the group." Similarly, LINE also treats chatbots as regular group members regarding access permissions. While this framing reflects expectations of interactive agents, it ignores their asymmetrical power. Chatbots can store and process data across multiple conversations, often without users realizing the extent of their data collection and analysis. This can lead to misunderstandings about what chatbots are capable of.

Single-user Service. Chatbots are treated like third-party single-user services integrated into messaging platforms, similar to mobile apps on mobile operating systems. For example, Discord explicitly refers to chatbots as "apps." This framing aligns with the plug-and-play experience of mobile apps and supports a clear access control mechanism. However, unlike standalone apps, group chatbots can collect data from all participants, including those who did not consent to their presence. Additionally, it overlooks the fact that chatbots can continuously monitor conversations.

Multi-user Service. This perspective views chatbots as always-on devices, like smart speakers. They may continuously listen for trigger words or full conversations, affecting not only direct users but also incidental users in group chats, often without their knowledge or consent. We believe this is the most accurate paradigm, as it captures the privacy risks in shared spaces.

6.2 Limitations

User Study. Our user study has four major limitations. First, because we rely on self-reports, we may miss cases where users encountered chatbots without realizing it. As a result, we cannot accurately estimate how many users actually encountered chatbots in group chats. Second, due to social-desirability bias, some participants may exaggerate their privacy concerns. We tried to prevent participants from overstating their privacy concerns by not mentioning privacy and security in the task description and by maintaining neutral language in question statements. Third, our

study has limited external validity. Participants recruited through Prolific tend to be younger, more educated, more technically inclined, and more aware of security and privacy issues than the general public, as shown in prior research [35]. Finally, the demographics of chatbot users may differ from the general population regardless of platform. These factors may limit the generalizability of our findings to broader populations.

Platform Analysis. Our ability to empirically verify certain permission-related behaviors was constrained by platform policies. These behaviors served as ground truth when evaluating user's understanding of what chatbots can access on their chosen platforms. Discord documentation states that chatbots added to over 100 servers require platform-level review to access message content and member lists [11], and we did not reach this threshold. LINE similarly restricts access to group member lists to verified or premium accounts [6], which we did not obtain. In these cases, we accepted official documentation as authoritative without further verification.

7 Related Work

7.1 Security and Privacy of Chatbots

Prior work has explored privacy and security risks in one-on-one chatbot interactions. Yang et al. [40] reviewed threats such as insecure implementations, user profiling, and data breaches, as well as solutions such as end-to-end encryption and organizational controls. Hasal et al. [16] surveyed security issues including identity authentication, lack of encryption, and GDPR compliance challenges. Many of these concerns, particularly user profiling and data protection, also apply to group chatbots.

Prior work has also examined the security and privacy risks of chatbots in group settings. Edu et al. [12] found that Discord bots are often overprivileged. Chen et al. [3] demonstrated that third-party apps in Microsoft Teams and Slack can exploit weak access controls to escalate privileges and access sensitive data. Chou et al. [4] showed that bots can access excessive message content and enable cross-group tracking, and proposed a secure group messaging protocol that preserves end-to-end encryption. Zha et al. [41] found that weak access controls in business chat apps allow privilege escalation. These studies highlight security and privacy concerns with chatbots in group chats and underscore the importance of understanding users' perceptions of these risks.

7.2 Users' Privacy Perceptions About Chatbots

Most research on users' privacy perceptions of chatbots has focused on one-on-one interactions. Gumusel [15] provides a comprehensive review, identifies key privacy concerns such as self-disclosure risks, data misuse, and security breaches, and notes challenges like weak regulations and manipulative designs. Our study shows these concerns also appear in group contexts. Ischen et al. [18] find that anthropomorphic cues in chatbots can reduce privacy concerns by building trust, even though they may lead to greater disclosure. Gieselmann and Sassenberg[14] report that users are more willing to share information when chatbots show problem-solving abilities. These studies collectively demonstrate that user perceptions of privacy significantly impact their engagement with chatbots.

Few studies have explored how users perceive chatbot-related privacy threats in group settings. To our knowledge, only Zha et al. [41] have conducted a user study on this topic. They explored whether attacks they identified on business team chat platforms, such as unprivileged users using chatbots to access shared links in private messages, raise concerns. In contrast, our work focuses on general messaging services and examines concerns around basic message-listening features, rather than advanced functionalities.

7.3 User's Privacy Perception in Related Fields

Prior work has extensively examined user perceptions of mobile app privacy. Prange et al. [30] found that only half of Android users correctly recognize granted permissions and often revoke sensitive ones when not essential. Almuhimedi et al. [1] showed that clear summaries of data access and privacy nudges can encourage users to adjust permissions. These findings support our observations regarding users' limited awareness and the need for assistance in assessing privacy risks.

Smart speakers, like chatbots that access all group messages, raise privacy concerns due to their always-on microphones. Manikonda et al. [27] found that users became more concerned about privacy after learning about this feature, echoing our findings on chatbots. Similarly, Lau et al. [21] reported widespread misunderstandings about smart speaker permissions, consistent with our observations.

8 Conclusion

In this work, we conducted a user study that revealed significant gaps in user awareness and understanding of chatbot data access in group messaging environments. Many users underestimated chatbot capabilities and were unaware of their presence. We identified six key mismatches between user mental models and platform designs, and introduced a five-stage model to explain how users detect and respond to privacy risks. Our platform analysis further showed that current designs often fail to support users in managing these concerns. Addressing these challenges is essential for protecting user privacy in chatbot-enabled group communication.

Acknowledgments

The authors used generative AI tools to revise the text, enhance its flow, and correct typos, grammatical errors, and awkward phrasing. This research was supported in part by the National Science and Technology Council of Taiwan under grants 112-2223-E-002-010-MY4, 114-2634-F-002-003MBK, and 113-2923-E-002-010-MY2; by National Taiwan University under grant 113L900901; by Len Blavatnik and the Blavatnik Family Foundation; by a Maof Prize for Outstanding Young Scientists; and by the Ministry of Innovation, Science & Technology, Israel (grant number 0603870071).

References

[1] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. 2015. Your location has been shared 5,398 times! A field study on mobile app privacy nudging. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 787–796.

[2] Apple Inc. 2024. Categories and Discoverability - App Store. https://developer.apple.com/app-store/categories/. Accessed on 2024-07-30.

[3] Yunang Chen, Yue Gao, Nick Ceccio, Rahul Chatterjee, Kassem Fawaz, and Earlence Fernandes. 2022. Experimental security analysis of the app model in

business collaboration platforms. In *31st USENIX Security Symposium (USENIX Security 22)*. 2011–2028.
[4] Kai-Hsiang Chou, Yi-Min Lin, Yi-An Wang, Jonathan Weiping Li, Tiffany Hyun-Jin Kim, and Hsu-Chun Hsiao. 2025. Bots can Snoop: Uncovering and Mitigating Privacy Risks of Bots in Group Chats. In *34th USENIX Security Symposium (USENIX Security 25)*. 6599–6618.
[5] codgoencasa. 2025. codigoencasa/builderbot. https://github.com/codigoencasa/builderbot. Accessed on 2025-05-01.
[6] LY Corporation. 2025. Messaging API reference. https://developers.line.biz/en/reference/messaging-api/. Accessed on 2025-05-29.
[7] David Curry. 2024. Messaging App Revenue and Usage Statistics (2024). https://www.businessofapps.com/data/messaging-app-market/. Accessed on 2024-02-07.
[8] Dank Memer. 2025. Dank Memer Discord Bot. https://top.gg/bot/270904126974590976 Accessed on 2025-05-01.
[9] Sauvik Das, Cori Faklaris, Jason I Hong, Laura A Dabbish, et al. 2022. The security & privacy acceptance framework (spaf). *Foundations and Trends® in Privacy and Security* 5, 1-2 (2022), 1–143.
[10] Discord Inc. 2025. Discord Developer Portal. https://discord.com/developers/docs/intro. Accessed on 2025-05-01.
[11] Discord Inc. 2025. How do I get Privileged Intents for my bot? – Developers. https://support-dev.discord.com/hc/en-us/articles/6205754771351-How-do-I-get-Privileged-Intents-for-my-bot. Accessed on 2025-05-01.
[12] Jide Edu, Cliona Mulligan, Fabio Pierazzi, Jason Polakis, Guillermo Suarez-Tangil, and Jose Such. 2022. Exploring the security and privacy risks of chatbots in messaging services. In *Proceedings of the 22nd ACM internet measurement conference*.
[13] Fernando Fierro. 2024. *Smart Chatbots for Smart Communities: ChainGPT's Solution.* https://www.chaingpt.org/blog/smart-chatbots-for-smart-communities-chaingpts-solution Accessed on 2025-05-01.
[14] Miriam Gieselmann and Kai Sassenberg. 2023. The more competent, the better? the effects of perceived competencies on disclosure towards conversational artificial intelligence. *Social Science Computer Review* 41, 6 (2023), 2342–2363.
[15] Ece Gumusel. 2025. A literature review of user privacy concerns in conversational chatbots: A social informatics approach: An Annual Review of Information Science and Technology (ARIST) paper. *Journal of the Association for Information Science and Technology* 76, 1 (2025), 121–154.
[16] Martin Hasal, Jana Nowaková, Khalifa Ahmed Saghair, Hussam Abdulla, Václav Snášel, and Lidia Ogiela. 2021. Chatbots: Security, privacy, data protection, and social aspects. *Concurrency and Computation: Practice and Experience* 33, 19 (2021), e6426.
[17] Discord Inc. 2025. Message Content Intent Alternatives – Developers. https://support-dev.discord.com/hc/en-us/articles/6383579033751-Message-Content-Intent-Alternatives. Accessed on 2025-05-24.
[18] Carolin Ischen, Theo Araujo, Hilde Voorveld, Guda Van Noort, and Edith Smit. 2019. Privacy concerns in chatbot interactions. In *International workshop on chatbot research and design*. 34–48.
[19] Sienna Jang. 2024. *Introducing Advanced Chat Moderation.* https://sendbird.com/blog/introducing-advanced-chat-moderation Accessed on 2025-05-01.
[20] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. 2009. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*. 1–12.
[21] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–31.
[22] LINE Corporation. 2025. Messaging API. https://developers.line.biz/en/services/messaging-api/. Accessed on 2025-05-01.
[23] WhatsApp LLC. 2021. Terms of Service. https://www.whatsapp.com/legal/terms-of-service. Accessed 2025-05-11.
[24] WhatsApp LLC. n.d. Unauthorized use of automated or bulk messaging on WhatsApp. https://faq.whatsapp.com/5957850900902049/. Accessed on 2025-05-12.
[25] Pedro S. Lopez. 2022. WhatsApp Banning Account. Unofficial App · Issue #1872 · pedroslopez/whatsapp-web.js. https://github.com/pedroslopez/whatsapp-web.js/issues/1872. Accessed on 2025-05-01.
[26] Pedro S. Lopez. 2025. pedroslopez/whatsapp-web.js. https://github.com/pedroslopez/whatsapp-web.js. Accessed on 2025-05-01.
[27] 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*. 229–235.
[28] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.
[29] Paradox AI. 2025. Conversational Hiring Software that Gets Work Done for You. https://www.paradox.ai/ Accessed on 2025-05-01.
[30] Sarah Prange, Pascal Knierim, Gabriel Knoll, Felix Dietz, Alexander De Luca, and Florian Alt. 2024. {"I} do (not) need that {Feature!"}–Understanding {Users'} Awareness and Control of Privacy Permissions on Android Smartphones. In *Twentieth Symposium on Usable Privacy and Security (SOUPS 2024)*. 453–472.
[31] Viber Media S.à r.l. 2025. Viber Terms of Service | Viber. https://www.viber.com/en/terms/viber-terms-use/. Accessed on 2025-05-11].
[32] Rythm. 2025. Rythm Discord Bot. https://top.gg/bot/235088799074484224 Accessed on 2025-05-01.
[33] Shruti Sannon, Brett Stoll, Dominic DiFranzo, Malte F Jung, and Natalya N Bazarova. 2020. "I just shared your responses" extending communication privacy management theory to interactions with conversational agents. *Proceedings of the ACM on Human-Computer Interaction* 4, GROUP (2020), 1–18.
[34] Robin Staab, Mark Vero, Mislav Balunović, and Martin Vechev. 2023. Beyond Memorization: Violating Privacy Via Inference with Large Language Models. *arXiv preprint arXiv:2310.07298* (2023).
[35] Jenny Tang, Eleanor Birrell, and Ada Lerner. 2022. Replication: How well do my results generalize now? The external validity of online privacy and security surveys. In *Eighteenth symposium on usable privacy and security (SOUPS 2022)*. 367–385.
[36] Telegram. 2020. Allow bots to see message deletion events – Bugs and Suggestions. https://bugs.telegram.org/c/735/. Accessed on 2025-03-15.
[37] Telegram Messenger Inc. 2025. Telegram Bot API. https://core.telegram.org/bots/api. Accessed on 2025-05-01.
[38] Toolerific AI. 2025. AI Tools for Chat Moderator. https://toolerific.ai/ai-tools-for-jobs/chat-moderator Accessed on 2025-05-01.
[39] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272.
[40] Jing Yang, Yen-Lin Chen, Lip Yee Por, and Chin Soon Ku. 2023. A systematic literature review of information security in chatbots. *Applied Sciences* 13, 11 (2023), 6355.
[41] Mingming Zha, Jice Wang, Yuhong Nan, Xiaofeng Wang, Yuqing Zhang, and Zelin Yang. 2022. Hazard Integrated: Understanding Security Risks in App Extensions to Team Chat Systems. In *NDSS*.

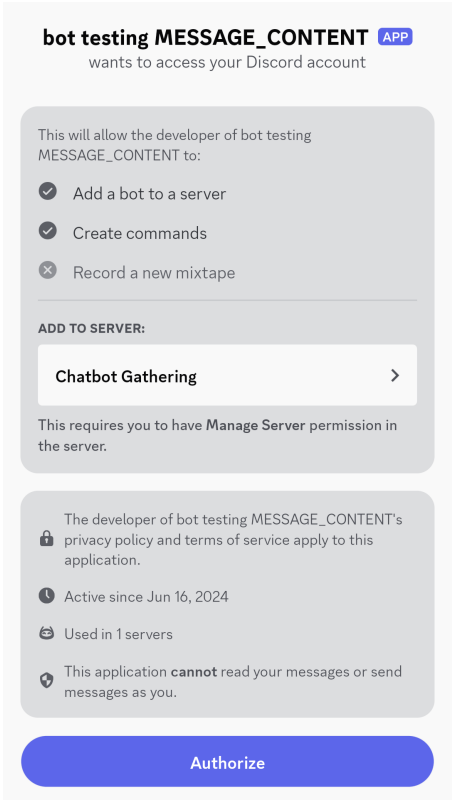
A Participant Demographics

Table 4 presents the demographic information of the participants in our user study.

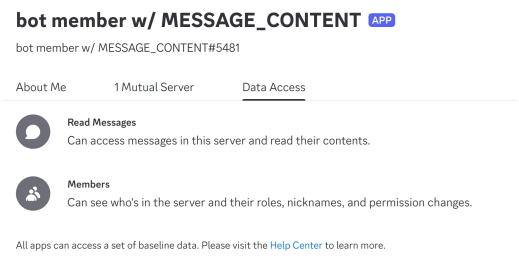
Table 4: Demographic data of participants. Providing demographic information was optional, so total counts may not sum to the total number of participants.

Category	Count
Gender	
Male	191
Female	176
Non-binary	2
Age	
18–24	131
25–34	160
35–44	48
45–54	22
55–64	6
65–74	4
75–84	2
Technical Work Experience	
Yes	235
No	131

Category	Count
Education	
Less than high school	7
High school graduate	49
Some college, no degree	52
Associate's degree	15
Bachelor's degree	181
Master's degree	56
Professional degree	4
Doctorate	7
Race and Ethnicity	
American Indian/Alaska Native	1
Asian	26
Black/African American	153
Hispanic/Latino	28
Mixed race	11
Pacific Islander	0
White/Caucasian	141
Prefer to self-describe	8



(a) Discord: Permission prompt shown when adding a chatbot to a server.



(b) Discord: Detailed permissions in bot's profile view after joining.

Figure 7: Transparency of chatbot permissions in Discord: (a) initial permission request when adding bots, (b) visibility of access permissions in group member lists.

B Supplementary Analysis of the User Study

B.1 Purposes of Chatbots in Group Chats

To better understand why participants use chatbots in group chats, we examined their reported use cases. Using the Apple App Store's app categories [2], we asked people what they use chatbots for. Most participants reported using them for productivity, entertainment and social networking, as shown in Fig. 8. We also did a thematic analysis of the the open-ended responses, and the results are listed in the Table 5. We found that the most common functions of chatbots are for productivity tasks and group moderation. These purposes are in line with common understandings of chatbots as tools or toys in group chats.

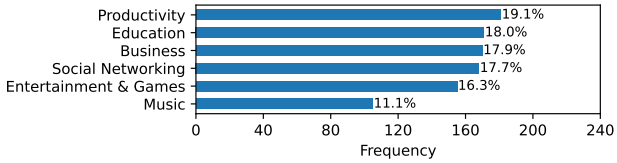


Figure 8: Purposes of chatbots using App Store's app taxonomy. Only options with more than 10% are included.

B.2 Why Checking the Presence of Chatbot

We investigate why users try to identify chatbots by conducting a thematic analysis of responses to open-ended questions. Our analysis reveals four main reasons users do so: utility, administrative needs, privacy concerns, and curiosity.

Desire to Use Chatbots. The most straightforward and commonly mentioned reasons for looking for chatbots are related to their utility. Many participants use terms like "helpful" or "convenient" to describe chatbots and explain that their presence can "benefit the members" (P256) or "facilitate discussions" (P351). For instance, P183 notes that they check for chatbots to improve productivity, stating, "I sometimes do check for chatbots in a group just to see if there are any that can help me perform certain tasks." Similarly, P32 highlights their entertainment value, saying, "including bots in private group chats myself to have fun with them along with my friends (usually bots that provide some kind of text-based game)."

Administrative Purposes. Many participants expect chatbots to serve administrative functions, such as providing group rules, offering instructions to participants, automating moderation, or detecting and removing spammers. Some participants feel obligated to check for the presence of chatbots for these reasons. For instance, P44 uses chatbots "[t]o see any rules of the group." P54 emphasizes the importance of chatbots in larger groups, explaining, "It's particularly important in larger groups where chatbots might be handling FAQs or enforcing rules." Having chatbots in groups may enhance users' sense of security and confidence. For example, P115 believes that groups with chatbots are likely to "be well-organized and free of offensive content." Additionally, some participants view chatbots as trusted sources of information about the group. P332, for instance, mentions that chatbots help them understand "how the group is ran [sic]."

Security and Privacy Concerns. Participants also mention checking for chatbots due to concerns about security and privacy. For instance, P109 states that they check for chatbots to ensure that "chatbots are not logging or accessing conversations without consent." Similarly, P134 highlights the importance of awareness, noting that "being aware of their presence can help you make informed decisions about what information you share in the group." These examples illustrate that some participants recognize chatbots' potential access to conversations and view this as a privacy risk.

Table 5: Thematic categorization of chatbot purposes in group chats, based on open-ended responses.

Category	Subcategory	Example
Informational functions	Announcements	"it is very helpful and it **keeps me updated**." (P132)
	Q&A support	"they usually **give accurate answers** straight away before any people respond." (P33)
	Topic monitoring	"we have chatbots that help us **monitor [cryptocurrency] coins** called to that group chat" (P312)
Group management	Moderation	"chatbot [...] can **ban users who do not behave properly according to the rules**." (P9)
	Instructions	"I always check for chatbots **providing instructions about group rules**." (P98)
	Group statistics	"The serverstats bot is often used on private servers to **monitor how many messages each user sends**." (P135)
	User verification	"they offer to **do user verification before confirming my request to join the group**." (P216)
	Welcome message	"the bot appears right away, **welcoming new members**." (P115)
	Spam prevention	"I check bots when logging in and see if they **prevent spam**." (P18)
Productivity	Automated tasks	"I noticed **a chatbot that provided reminders for our monthly family gatherings**." (P17)
Entertainment	Music	"they feature the ability to, for example, **play music for everyone to hear at the same time**." (P298)
	Game	"usually bots that provide some kind of **text-based game**." (P32)

Curiosity. Some participants check for chatbots out of curiosity, even without a specific intention to use them. For example, P141 shared, "I sometimes check for their presence when I first join a group just out of curiosity."

B.3 Platform Differences in User Understanding

To examine whether users' understanding of chatbots' data access capabilities varies across platforms, we compare participants' average understanding scores, excluding LINE and Viber due to insufficient data. We find that Discord users consistently outperform users of other platforms across all three questions.

For the first question on message types, 58% of Discord users answer correctly, compared to 45% for Telegram and 36% for WhatsApp. A Tukey HSD test confirms that the difference between Discord and WhatsApp is statistically significant ($p = 0.0011$), though the difference between Discord and Telegram is not ($p = 0.3575$). On the second question regarding message metadata, Discord users selected 3.36 out of five options correctly on average, slightly higher than Telegram (3.33) and WhatsApp (2.95). The only statistically significant difference was between Discord and WhatsApp ($p = 0.0372$). For the third question on group metadata, Discord participants again performed better, correctly selecting 3.29 out of four options on average, compared to 2.93 for Telegram and 2.60 for WhatsApp. Again, only the difference between Discord and WhatsApp reached statistical significance ($p < 0.0001$).

Several factors may contribute to these results. We hypothesize that one key reason is Discord's policy of officially supporting chatbots, along with its strong chatbot culture. Telegram, another platform where chatbots are common, also shows slightly higher scores. In contrast, WhatsApp does not officially support group chatbots. Instead, users rely on unofficial userbots that operate underground, which may lead to lower awareness of chatbot behavior among WhatsApp users. However, further research is needed to better understand the causes of these differences.

B.4 Analysis of User's Privacy Concerns Including Discord Participants

In Sec. 4.4, we noted that, due to a potentially misleading statement regarding the permissions of Discord chatbots, we excluded responses from participants who used Discord. In this subsection, we examine the effect of including Discord users in our analysis and demonstrate that our main conclusions remain consistent regardless of their inclusion.

Privacy Perception. Our survey results indicate that a large proportion of users express privacy concerns about chatbots in group chats. When including Discord respondents, 68.2% somewhat or strongly agree that chatbots accessing user data raises privacy concerns (compared to 73.2% when Discord participants are excluded). Similarly, 71.7% somewhat or strongly agree that they are concerned about the collection and misuse of personal information shared in group chats. (compared to 73.5% when Discord participants are excluded).

Thus, regardless of whether Discord users are included, the key insight remains consistent: users are generally cautious about chatbot access to their data and concerned about the potential risks of sharing sensitive information.

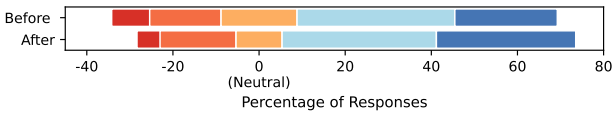


Figure 9: Privacy concerns before and after learning chatbot permissions, measured on a 5-point Likert scale from "strongly disagree" (left) to "strongly agree" (right).

We also compare the level of users' privacy concerns before and after informing them about the permissions granted to chatbots. As shown in Fig. 9, participants expressed greater concern after learning about these permissions. Before receiving this information, 60.4% of respondents somewhat or strongly agreed that chatbots raise privacy concerns. After being informed, this number rose to 68.2% (compared to an increase from 64.3% to 73.2% when Discord users are excluded). A Wilcoxon signed-rank test confirms that this change is statistically significant ($p < 0.001$) when Discord users are included.

These results support the same conclusion: users become significantly more concerned about privacy once they understand the extent of chatbots' data access.

Responses to Privacy Concerns. We examine how users respond to the privacy risks posed by chatbots. After being informed about the potential threats, participants were asked whether they would continue using chatbots and why. 42.8% of participants stated they would continue using them, which is similar to the 42.3% observed when Discord users are excluded. We also asked whether users would change their behavior in response to these privacy concerns. 65.5% of participants said they would adjust their behavior, such as avoiding controversial topics, due to the presence of chatbots in group chats. This is comparable to the 68.4% reported when Discord users are excluded. These results lead to the same conclusion: while many users may still choose to use chatbots, they are likely to change their behavior in response to privacy concerns, regardless of whether Discord participants are included.

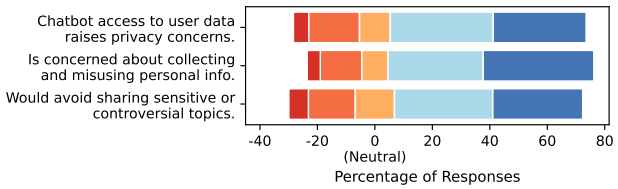


Figure 10: User's privacy concerns about chatbots, measured on a 5-point Likert scale from "strongly disagree" (left) to "strongly agree" (right).

B.5 Thematic Analysis of Users' Privacy Concerns About Chatbots

In the following, we present a more detailed thematic analysis of users' privacy concerns related to chatbots. A summary of the identified themes is provided in Table 1.

Lack of Transparency. A commonly seen concern for users are that chatbots often lack transparency on what they collect and how they handle the collected data, making users confused. As P95 put, "I do not understand how they protect our data or use it to be more specific besides helping us, is that where all the information ends up?" Moreover, many chatbots lack privacy policies, and even if there are privacy policies, they are often overlooked. As P115 put, "Often, when adopting the use of chatbots, we don't read the fine print, similar to how we often overlook the "terms of use" of software." The cross-border nature of chatbots also complicate the legal issues. As P49 mention, "Since privacy laws differ strongly across the globe, I'd be concerned about who's privy to my information and what it could be used for."

Excessive or Sensitive Data Collection. A natural privacy concern is about chatbots collecting too much data or collecting sensitive data, especially personally identifiable information, without consent. Many users worry that chatbots gather more information than necessary for their function. Some feel uncomfortable knowing that chatbots in group chats may collect data even when they are not actively being used. As P16 put it, "I feel like they should only be allowed to collect data that we gave them permission to collect, and not collect data at the time when they weren't called on to be used." Sometimes a group contains sensitive information that chatbots may inadvertently read. As P54 put it, "Since group chats often involve multiple participants, there's a risk that sensitive or private information might be unintentionally processed or stored by the chatbot without clear consent."

Data Misuses: Selling to Third Parties. A significant concern is the potential misuse of collected information. We have categorized data misuse into four types of data misuses and started with the first one: selling to third parties. One major fear is that chatbot providers might sell user data to third-party companies without consent. Some users suspect that their private conversations could be gathered and sold to data brokers. For example, P216 mentioned, "I have issues if the data collected is more than I agree to and later sold out to data brokers."

Data Misuses: Scams. There is also concern that chatbots could be used to collect sensitive data for fraudulent activities such as phishing or identity theft. P195 feared that chatbots could be "stealing your passwords or committing identity theft." P271 mentioned a specific scenario: "I am scared that I may share sensitive information and have it leaked [...] and later get scammed."

Data Misuses: Ads and Ad-motivated Tracking. Some users are uneasy about chatbots tracking their conversations for targeted advertising. They fear that their data might be used to create detailed profiles for marketing purposes. As P196 put it, "Only concern for me is whether or not my information is collected and used for advertisement purposes as that feels more intrusive."

Data Misuses: Censorship and Legal Enforcement. Some participants also worry about how their data might be used by governments or other authorities. Some fear that chatbots could monitor conversations for sensitive content that may lead to legal consequences. P49 highlighted this concern: "information accessed by chat bots could be used to identify and target people trying to get abortions, belong to a marginalized group that's discriminated against, and so on." P158* similarly mentioned "having data that can be subpoenaed by authorities to prosecute people."

Data Breaches. Some participants mentioned the risk of data breaches when using chatbots. If a chatbot is hacked or misconfigured, sensitive data could be exposed to unintended parties. For example, P300 highlighted the broader risk of hacking, saying, "If a chatbot is not properly secured, it could be a target for hacking, which could expose private conversations to unauthorized users." P12 also mentioned that "there is a risk of data breaches, where collected information could be accessed by unauthorized parties." These concerns reflect a common fear that chatbot operators might not have sufficient security measures to protect user information.

B.6 User Expectation of Chatbot Data Access

We present visualizations of the data on user expectations for chatbot data access, as discussed in Sec. 4.5. Fig. 11 illustrates user preferences regarding the types of messages chatbots should be allowed to access. Fig. 12 shows user preferences for chatbot access to different message metadata. Fig. 13 shows preferences related to access to group metadata.

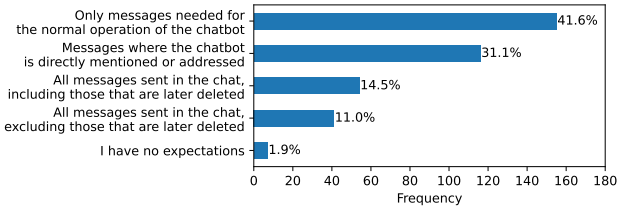


Figure 11: User preferences for what chatbots should have access to.

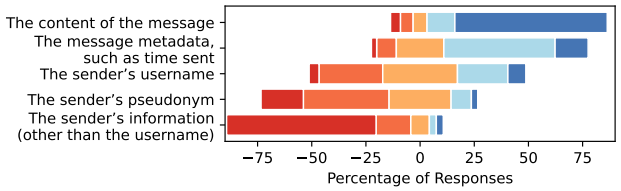


Figure 12: User preferences for chatbot access to different message metadata, ranked from most preferred (dark blue, rank 1) to least preferred (dark red, rank 5).

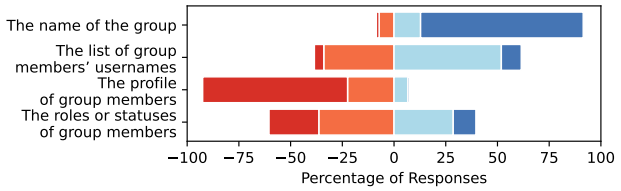


Figure 13: User preferences for chatbot access to different group metadata, ranked from most preferred (dark blue, rank 1) to least preferred (dark red, rank 4).

B.7 Other Misconceptions in User Study

During our analysis, we identified several misconceptions about instant messaging services and chatbots that fell outside our main research questions. Here, we highlight two notable examples.

Belief That Chatbot Operators Cannot Read Messages. Some participants believed that only the chatbot, not its human operators, could read the messages. For example, P272 remarked, "Chatbots [are] not real human beings[,] so there is nothing they can do with my personal or sensitive information I share [...]." This confusion between automated systems and their human operators may lead users to overshare sensitive information.

Viewing Service Providers as the Only Threat. Some participants viewed service providers as the only privacy threat. P141 stated, "I don't trust WhatsApp [...] I fear WhatsApp may use the chatbot to access our info and sell it to advertisers." P138 also said, "The messages... are end-to-end [encrypted], so the chatbot has access but WhatsApp doesn't. This assure[s] me that I don't have to worry about privacy." While service providers can pose risks, this narrow view overlooks the potential privacy threats from the chatbots.

B.8 Details of Statistical Inferences

In our analysis of the user study results, we performed several statistical inferences to support our arguments. All calculations were conducted using Scipy 1.11.3 [39]. This subsection summarizes these inferences.

User's Self-Reported and Actual Knowledge of Chatbot Permissions. In Sec. 4.3.2, we use Spearman's rank correlation to examine the relationship between users' self-reported knowledge and their actual understanding of chatbot permissions, including access to message content, message metadata, and group metadata. The null hypothesis states that there is no monotonic relationship between self-reported knowledge and actual understanding. The alternative hypothesis states that a monotonic relationship exists. The results are presented in Table 6.

Table 6: Spearman's rank correlation between participants' self-reported knowledgeability and their actual understanding of chatbot permissions.

Question	ρ (Spearman)	p-value	Significance
Message content	−0.0577	0.2654	n.s.
Message metadata	0.0389	0.4527	n.s.
Group metadata	0.1563	0.0024	*

n.s. = not significant at $p < 0.05$, * = significant at $p < 0.05$.

Changes in Privacy Concerns. In Sec. 4.4.1, we analyze how participants' privacy concerns changed after learning about chatbot permissions. We use the Wilcoxon signed-rank test to evaluate the difference. The null hypothesis states that the median difference in privacy concern scores before and after learning about chatbot permissions is zero. The alternative hypothesis states that the median difference is not zero. The results are shown in Table 7.

Table 7: Wilcoxon signed-rank test on change in privacy concern scores before and after participants learned about chatbot permissions.

Test Statistic	p-value	Median Diff.	95% CI
2261.5	0.00046	0.000	[−2.000, 3.000]

C User Study Protocol

C.1 Informed Consent Form

This research project aims to understand users' awareness regarding chatbots. This survey is anonymous and contains no personally identifiable information. The results will be used for statistical analysis and may be published in domestic and international journals, conferences, and educational materials.

Those who meet the following criteria are suitable to participate in this study:

- Regular users of instant messaging services, including WhatsApp, Telegram, WeChat, Viber, LINE, and Discord.

- Have seen or used chatbots on instant messaging services.
- Must be at least 18 years old and meet the legal age of adulthood in your locality to participate in this study.

Those who meet the following criteria cannot participate in this study:

- Do not use instant messaging services regularly.
- Under the age of 18 or below the legal age of adulthood in your locality.

You are free to decide whether or not to complete this survey. You can stop and leave the survey at any time without consequence. However, once you have submitted, the survey cannot identify you because it is anonymous, and we cannot delete your responses. Participants will receive £3 upon completion, which will be distributed through Prolific.

If you have any questions about the survey or are interested in the research results, this study will be completed by November 30, 2024. You are welcome to contact us using the contact information below to request a summary of the research results.

This study used the Qualtrics platform to collect the questionnaire results, and all the completed data will be automatically stored in the Qualtrics platform, and only the principal investigator and the research assistants will have the passwords of the account, and the passwords will not be leaked to anyone outside the project. In line with the current international trend of open science, we will remove inappropriate content from the raw data and publish it in an open access repository (e.g., Open Science Framework).

This research has been reviewed and approved by the Research Ethics Committee of our institution. If you have any concerns about your rights as a participant, or if you believe you have been harmed as a result of your participation in the research, please contact the Research Ethics Committee of our institution directly. The phone numbers are: (redacted for anonymous submission). After reviewing the form and content of the survey, no commercial benefit will be derived from the raw data, and no potential risks were anticipated. Therefore, no compensation or insurance is provided.

Do you accept the terms outlined in the consent form and agree to participate in this study?

- No
- Yes

What is your Prolific ID? Please note that this response should auto-fill with the correct ID.

- Prolific ID field

C.2 Services Usage

(1) Please select one instant messaging service from the following list that you use frequently (more than twice a week) and have seen chatbots on. If more than one applies, choose the one you use the most. The rest of the survey will be based on your selection here.
- WhatsApp
- Telegram
- WeChat
- Viber
- LINE
- Discord
- I do not use any of these services.

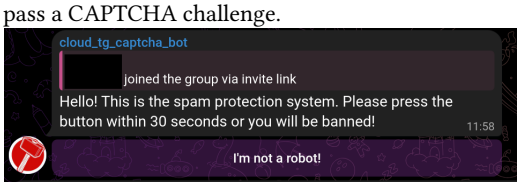
(2) How often do you use the instant messaging service you selected in the previous question?
- Daily
- 4-6 times per week
- 2-3 times per week

C.3 Chatbots Usage

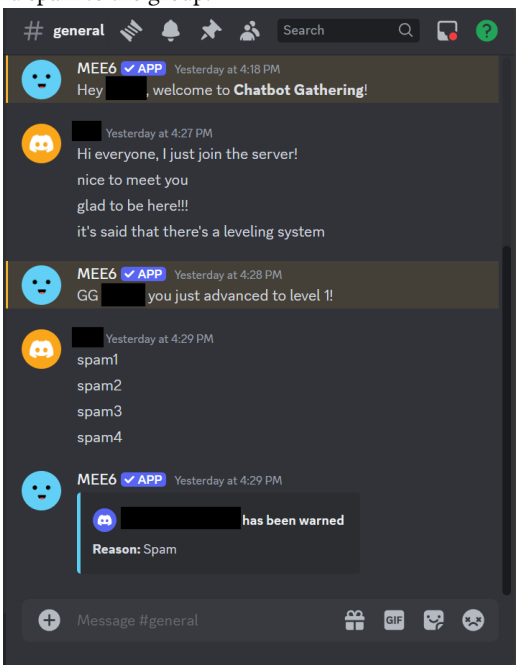
Chatbots in group chats on instant messaging services are automated programs designed to interact with users through text messages and simulate conversation with human users. Examples of popular platforms that support chatbots in their group chat functionalities include Slack, Discord, and Telegram. These platforms allow the integration of chatbots that can be customized to perform a wide range of actions, from managing workflows to engaging users in interactive conversations.

One of the primary functions of chatbots in group chats is to facilitate communication by offering quick responses to common queries. This can be particularly useful in customer service or support groups where chatbots provide instant answers to frequently asked questions, helping to streamline interactions and save time.

For example, in the following image, a chatbot asks a new user to pass a CAPTCHA challenge.



Another example is a moderation chatbot that bans users who send spam to the group.



In this section, you will be asked several questions about your experience using chatbots in group chats.

(3) Which of the following are considered chatbots in group chats and are therefore within the scope of this study? (Select all that apply)
This is a comprehension check. You must answer this question correctly before you can proceed.
- A chatbot in Slack that helps manage project workflows by assigning tasks to team members.
- A bot in a video game that plays against you using artificial intelligence.
- A chatbot in a customer support group on Telegram that provides instant answers to frequently asked questions.
- A virtual assistant on your phone (e.g., Siri, Google Assistant).
- A chatbot in Discord that facilitates interactive conversations among group members.
- A large language model like ChatGPT used in a one-on-one conversation to generate text.
- ChatGPT integrated into a group chat on Slack to provide automated responses to questions in real-time.

(4) How often do you notice chatbots in group chats, including seeing them send messages or actively searching for their existence?
- Daily
- 4-6 times per week
- 2-3 times per week
- Once or less than once per week
- Never
- Not Sure

(5) Please name 1-2 chatbots you encountered in group chats and describe them in 1-2 sentences. [Text field]

(6) How often do you interact with chatbots in group chats, including actively tagging/mentioning chatbots in your messages, chatbots replying to your messages, or chatbots reading your messages?
- Daily
- 4-6 times per week
- 2-3 times per week
- Once or less than once per week
- Never

(7) What types of group chats have you seen chatbots in? Select all that apply.
- Work-related groups
- Educational groups (classes, study groups)
- Friend or family groups
- Public forums or community groups
- Other (please specify)

(8) What purposes do you use chatbots for in group chats? Select all that apply.
- Business
- Productivity
- Developer Tools
- Books
- Education
- Entertainment & Games
- Finance
- Food & Drink
- Graphics & Design
- Health & Fitness
- News & Magazines
- Music
- Photo & Video
- Social Networking
- Travel

C.4 Awareness of Chatbots in Groups

Groups in instant messaging can generally be categorized into two types, though the boundary between them isn't always clear-cut: private and public groups.

Private groups in instant messaging are typically small, made up of close friends, colleagues, or family, and are private, often requiring an invitation to join. These groups focus on more personal conversations between members who know each other well. Conversations are expected to remain confidential and accessible only to group members and should not be leaked.

Public groups, on the other hand, are open to anyone interested in the topic of the group, such as developers working on open source projects or local community discussions. These groups tend to be larger, have less restrictive access, and cover broader topics, making them useful for networking and sharing information on a larger scale. Because of their open nature, conversations in these groups are not expected to remain confidential.

Each type of group serves different purposes: private groups for personal, intimate conversations and public groups for broader, more open discussions.

In this section, we will ask you some questions about your experience using chatbots in group chats.

The following questions start with **private groups**.

(9) Please select the third option: "Sometimes".
- Always
- Usually
- Sometimes
- Rarely
- Never

(10) Have you ever noticed any chatbots in private group chats you have participated in?
- Yes
- No
- Not sure

(11) How often do you check for the presence of chatbots whenever you join private group chats?
- Always
- Usually
- Sometimes
- Rarely
- Never

(12) Please elaborate on your answer to the previous question. [Text field]

(13) When do you check for the presence of chatbots in a private group? Select all that apply.
- When I first join a group
- When a new bot is added into the group
- When a bot is removed from the group
- When I notice unusual activity in the group

- I check every now and then to see if anything has changed

We will now move on to **public groups**. The questions are exactly the same, but instead we will focus on public groups. A public group, by the previous definition, is an open forum for anyone interested in the group's topic, ideal for networking and sharing information broadly, often with little expectation of confidentiality.

(14) Have you ever noticed any chatbots in public group chats you have participated in?
- Yes
- No
- Not sure

(15) How often do you check for the presence of chatbots whenever you join public group chats?
- Always
- Usually
- Sometimes
- Rarely
- Never

(16) Please elaborate on your answer to the previous question. [Text field]

(17) When do you check for the presence of chatbots in a public group? Select all that apply.
- When I first join a group
- When a new bot is added into the group
- When a bot is removed from the group
- When I notice unusual activity in the group
- I check every now and then to see if anything has changed

C.5 Chatbots' Data Access

(18) To what extent are you familiar with what information chatbots on [selected platform] can access in group chats?
- Not knowledgeable at all
- Slightly knowledgeable
- Moderately knowledgeable
- Very knowledgeable
- Extremely knowledgeable

(19) What type of messages do you believe a chatbot on [selected platform] has access to?
- All messages sent in the chat, including those that are later deleted.
- All messages sent in the chat, excluding those that are later deleted.
- Only messages needed for the normal operation of the chatbot.
- Messages where the chatbot is directly mentioned or addressed.
- I am not sure about the types of messages a chatbot receives.

(20) Following up on the previous question, what information do you believe is included with the message a chatbot has access to on [selected platform]? Select all that apply.
- The content of the message itself.
- The message metadata, such as time sent.
- The sender's username. A user has the same username in different groups.
- The sender's pseudonym. A user has different pseudonyms in different groups.
- The sender's metadata (other than the username), such as profile picture, gender, and email addresses.
- I am not sure about this.

(21) What type of information do you believe a chatbot on [selected platform] can access about the group? Select all that apply.
- The name of the group.
- The list of group members' usernames.
- The profile pictures of group members.
- The roles or statuses of group members (e.g., moderator).
- I am not sure what group information a chatbot can access.

(22) Chatbots from group chats with access to users' data as checked above would raise my privacy concerns.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(23) For this question, choose the fifth option, "Chatbots are never helpful."
- Chatbots are always helpful.
- Chatbots are usually helpful.
- Chatbots are sometimes helpful.
- Chatbots are rarely helpful.
- Chatbots are never helpful.

(Display the description of the selected platform.)

On WhatsApp, chatbots are able to access all messages in a group chat, regardless of whether they are mentioned. This access extends to both the content and metadata of messages, as well as the sender's username. Furthermore, chatbots can view the name of the group, along with the usernames, profiles, and roles of its members. Essentially, chatbots on WhatsApp have access to the same information as a normal group member.

On Telegram, chatbots are typically set up to access only messages that specifically mention them. This access extends to both the content and metadata of messages, as well as the sender's username. Furthermore, chatbots can view the name of the group, along with the usernames, profiles, and roles of its members.

On WeChat, chatbots are able to access all messages in a group chat, regardless of whether they are mentioned. This access extends to both the content and metadata of messages, as well as the sender's username. Furthermore, chatbots can view the name of the group, along with the usernames, profiles, and roles of its members. Essentially, chatbots on WeChat have access to the same information as a normal group member.

On Viber, chatbots are able to access all messages in a group chat, regardless of whether they are mentioned. This access extends to both the content and metadata of messages, as well as the sender's username. Furthermore, chatbots can view the name of the group, along with the usernames, profiles, and roles of its members. Essentially, chatbots on Viber have access to the same information as a normal group member.

On LINE, chatbots are able to access all messages in a group chat, regardless of whether they are mentioned. This access extends to

both the content and metadata of messages, as well as the sender's username. Furthermore, chatbots can view the name of the group, along with the usernames, profiles, and roles of its members. Essentially, chatbots on LINE have access to the same information as a normal group member.

On Discord, chatbots are able to access all messages in a group chat, regardless of whether they are mentioned. This access extends to both the content and metadata of messages, as well as the sender's username. The ability of chatbots to see usernames, profiles, and roles of group members can be adjusted depending on configuration settings.

To what extent do you agree or disagree with the following statements?

(24) Chatbots from group chats with access to users' data as listed above would raise my privacy concerns.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(25) I would be concerned that any personal information I share in the group chat could be collected by the chatbot and be used inappropriately.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(26) I would stop sharing some sensitive information or talking about controversial topics in group chat because of privacy concerns about chatbots.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(27) I would still use the chatbot after knowing what information it had access to.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(28) If you choose "somewhat agree" or "strongly agree" in the previous question, what are the reasons for continuing using the chatbot? [Text field]

(29) I think chatbots accessing more information than necessary could cause privacy issues in public group chats.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(30) I think chatbots accessing more information than necessary could cause privacy issues in private group chats.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(31) I understand how chatbot providers protect the data collected from group chats.
- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

(32) Do you have specific privacy concerns about chatbots in group chats? Please describe. [Text field]

C.6 Expectation

This section is about your preferences regarding the scope of access you believe should be granted to chatbots within group chats. Imagine that there is a chatbot that will be added to the group. You will be asked to specify the types of messages and information a chatbot should be able to access.

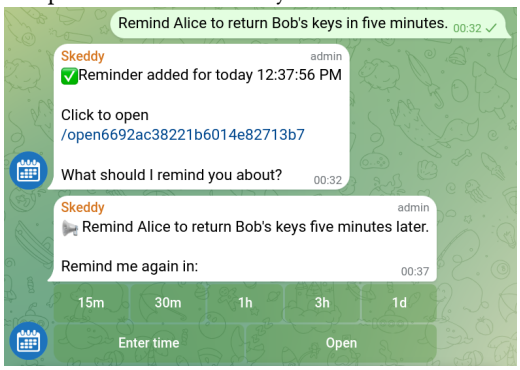
(33) Ideally, what type of messages do you expect a chatbot in a group chat has access to?
- All messages sent in the chat, including those that are later deleted.
- All messages sent in the chat, excluding those that are later deleted.
- Only messages needed for the normal operation of the chatbot.
- Messages where the chatbot is directly mentioned or addressed.
- I have no expectations.
- Other (please specify)

(34) Ideally, what specific information within a message would you expect a chatbot to have access to? Sort all options according to your level of expectation, with 1 being the most expected to be accessible, and 5 being least expected to be accessible.
- The content of the message itself.
- The message metadata, such as time sent.
- The sender's username. A user has the same username in different groups.
- The sender's pseudonym. A user has different pseudonyms in different groups.
- The sender's metadata (other than the username), such as profile picture, gender, and email addresses.

(35) Ideally, what information about the group itself would you expect a chatbot to be able to have access to? Sort all options according to your level of expectation, with 1 being the most expected to be accessible, and 4 being least expected to be accessible.
- The name of the group.
- The list of group members.
- The profile pictures of group members.
- The roles or statuses of group members (e.g., admin, moderator).

(36) Ideally, the numbers should be in ascending order, i.e., from 1 to 5. Please sort them accordingly.
- This is one.
- This is four.
- This is two.
- This is five.
- This is three.

Suppose there is a chatbot called Reminder. The user can add items to the reminder using a specific format that includes the time and the message. It then sends messages reminding group members at the specified time. What do you think it should have access to?

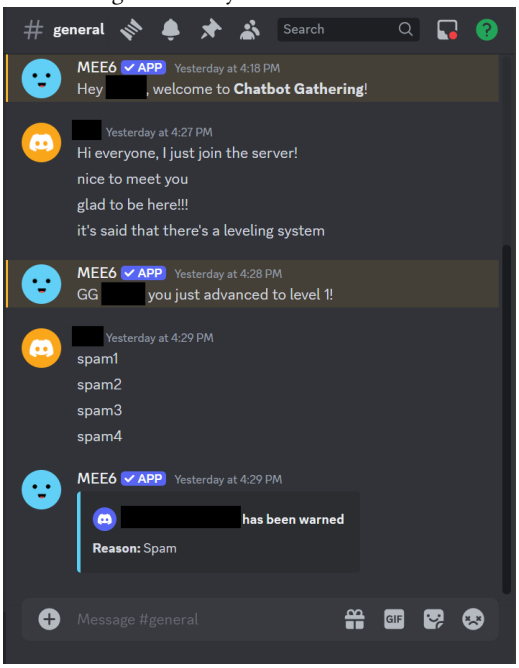


(37) Ideally, what type of messages do you expect Reminder has access to?
- All messages sent in the chat, including those that are later deleted.
- All messages sent in the chat, excluding those that are later deleted.
- Only messages in a specified format, e.g., "remind me to reply to the email in ten minutes."
- Messages where the chatbot is directly mentioned or addressed.
- I have no expectations.
- Other (please specify)

(38) Ideally, what specific information within a message do you expect Reminder to have access to? Sort all options according to your level of expectation, with 1 being the most expected to be accessible, and 5 being least expected to be accessible.
- The content of the message itself.
- The message metadata, such as time sent.
- The sender's username. A user has the same username in different groups.
- The sender's pseudonym. A user has different pseudonyms in different groups.
- The sender's metadata (other than the username), such as profile picture, gender, and email addresses.

(39) Ideally, what information about the group itself do you expect Reminder to be able to have access to? Sort all options according to your level of expectation, with 1 being the most expected to be accessible, and 4 being least expected to be accessible.
- The name of the group.
- The list of group members.
- The profile pictures of group members.
- The roles or statuses of group members (e.g., admin, moderator).

Suppose there is another chatbot called Moderation Helper. It is a collection of tools that help to moderate the group. For example, it sends welcome messages to new members, silences or bans members who send spam, and keeps track of members' activity for future auditing. What do you think it should have access to?



(39) Ideally, what type of messages do you expect Moderation Helper has access to?
- All messages sent in the chat, including those that are later deleted.
- Only messages that are present while the chatbot is active.
- Only messages required for group moderation, such as harassment messages, violent images, advertising messages, etc.
- Messages where the chatbot is directly mentioned or addressed.
- I have no expectations.
- Other (please specify)

(40) Ideally, what specific information within a message would you expect Moderation Helper to have access to? Sort all options according to your level of expectation, with 1 being the most expected to be accessible, and 5 being least expected to be accessible.
- The content of the message itself.
- The message metadata, such as time sent.
- The sender's username. A user has the same username in different groups.
- The sender's pseudonym. A user has different pseudonyms in different groups.
- The sender's metadata (other than the username), such as profile picture, gender, and email addresses.

(41) Ideally, the following options should be in alphabetical order. Please sort them accordingly.
- Artificial Intelligence
- Data Analysis
- Chatbots
- Blockchain
- Encryption

(42) Ideally, what information about the group itself would you expect Moderation Helper to be able to have access to? Sort all options according to your level of expectation, with 1 being the most expected to be accessible, and 4 being least expected to be accessible.
- The name of the group.
- The list of group members.
- The profile pictures of group members.
- The roles or statuses of group members (e.g., admin, moderator).

C.7 Demographics

All of the questions in this section are optional. There are no consequences if you choose not to answer.

(42) Age Group
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older

(43) Gender
- Male
- Female
- Non-binary
- Prefer to self-describe

(44) Highest level of education completed
- Less than high school
- High school graduate
- Some college but no degree
- Associate's degree in college (2-year)
- Bachelor's degree in college (4-year)
- Master's degree
- Professional degree (JD, MD, etc.)
- Doctorate

(45) Technical work experience
- No
- Yes

(46) Race and ethnicity
- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Mixed race
- Pacific Islander
- White or Caucasian
- Prefer to self-describe

D Codebooks

D.1 Noticing Chatbots in Group Chats

Participants were asked how often they check for the presence of chatbots when joining private (Q12) and public (Q16) group chats.

Theme: How do participants notice chatbots?

- **Code `N-A`**: Participants proactively check the group member list for chatbots.
- **Code `N-P`**: Participants involuntarily notice chatbots when they appear in the member list, send messages, or are interacting with other group members.
- **Code `N-NA`**: The responses are not suitable for categorization here.
- **Code `IR`**: The responses provided are not relevant to the question being asked.

Theme: Motivations behind proactive checking (`N-A`).

- **Code `A-W`**: Participants check chatbots because they want or need to use them.
- **Code `A-R`**: Participants check chatbots to understand administrative features or rules in group.
- **Code `A-S`**: Participants check chatbots due to concerns about potential security or privacy risks.
- **Code `A-NA`**: No specific reason mentioned.

Theme: Checking after involuntarily noticing chatbots (`N-P`).

- **Code `P-C`**: Participants choose to inspect the chatbot after noticing its presence.

Theme: Reasons for not checking after noticing (`N-P`).

- **Code `P-L`**: Participants feel unconcerned about the chatbot.
- **Code `P-NT`**: Interacting with chatbots is not a priority for participants.
- **Code `P-NN`**: Chatbots are easy to notice, so participants feel no need to check.
- **Code `P-T`**: Participants trust group administrators or members not to introduce harmful chatbots.
- **Code `P-U`**: Chatbots are so prevalent that participants often assume their existence without verifying it.
- **Code `P-NA`**: No specific reason mentioned.

Other.

- **Code `NA`**: The responses are not suitable for categorization here.
- **Code `IR`**: The responses provided are not relevant to the question being asked.

D.2 Continued Use Despite Privacy Risks

Participants were asked why they would continue using a chatbot even after learning about the types of information it can access.

Theme: Privacy risk is acknowledged but outweighed.

- **Code `U`**: Participants perceive chatbot utility to outweigh potential privacy concerns.
- **Code `C`**: Participants would change behaviors and become more cautious when sending messages in the group.

Theme: Privacy risks are ignored or rationalized.

- **Code `T`**: Participants trust the platform or chatbot developers.
- **Code `LC`**: Participants are inherently less concerned about privacy.
- **Code `PH`**: Participants feel powerless to mitigate privacy risks.
- **Code `FA`**: Participants feel that they are forced to accept privacy risks.
- **Code `NC`**: Participants generally do not share sensitive information online.
- **Code `TP`**: Participants believe that the privacy risks can be mitigated by transparency.

Other.

- **Code `IR`**: The responses provided are not relevant to the question being asked.

D.3 Additional Privacy Concerns

Participants were asked whether they had any additional privacy concerns regarding chatbots in group chats.

Theme: Participants express specific privacy concerns.

- **Code `TP`**: Participants request greater transparency from chatbots, including clearer data policies and self-introduction messages.
- **Code `DC`**: Participants are concerned that chatbots collect excessive amounts of user data.
- **Code `DM`**: Participants worry that their data may be misused.
- **Code `DB`**: Participants worry about potential data breaches involving chatbots.

Theme: No privacy concerns expressed.

- **Code `T-O`**: Participants trust the group members who added the chatbot.
- **Code `T-C`**: Participants trust the chatbot developers.
- **Code `T-S`**: Participants trust the platform to act as a gatekeeper and prevent harm.
- **Code `NO`**: No reason mentioned.
- **Code `IR`**: The responses provided are not relevant to the question being asked.