Who's Watching You Zoom?
Investigating Privacy of Third-Party Zoom Apps

Saharsh Goenka
Arizona State University
Tempe, Arizona, USA
sgoenka1@asu.edu

Adit Prabhu
Arizona State University
Tempe, Arizona, USA
adprabh2@asu.edu

Payge Sakurai
Arizona State University
Tempe, Arizona, USA
psakurai@asu.edu

Mrinaal Ramachandran
Arizona State University
Tempe, Arizona, USA
maramach@asu.edu

Rakibul Hasan
Arizona State University
Tempe, Arizona, USA
rakibul.hasan@asu.edu

Abstract

Zoom serves millions of users daily with a growing marketplace of third-party apps that can reach those users. Despite being the leading remote collaboration platform, the privacy and security aspects of marketplace apps have not been investigated. This paper examines the evolution of the Zoom Marketplace over one year, identifying trends in apps, their patterns of permission requests for user data, and the transparency of their privacy policies. Our findings surface increasing over-collection of user data, obscurity in data collection and sharing purposes, and potential non-compliance with laws in the education and healthcare sectors. In light of these findings, we provide concrete recommendations for Zoom and directions for future research to improve the privacy posture of this emerging platform.

Keywords

Privacy, Zoom Marketplace, Third-party applications, Privacy policy analysis, Data permissions.

1 Introduction

Zoom has become the dominant remote conferencing platform with more than 300 million daily active users [60]. In the US, Zoom has also become the official tool in many public and private organizations, including education and healthcare institutes, as well as business entities. Recently, Zoom launched a marketplace for external applications (apps) offering additional services.[1] Similar to apps on other marketplaces (e.g., Android apps and Alexa skills), Zoom apps are most commonly developed by third parties. Their access to user data thus raises concerns about privacy, safety, and security. Researchers have extensively investigated other marketplaces and uncovered many privacy and security issues (see § 2). Zoom marketplace, however, differs from those platforms in several ways, as explained below.

First, Zoom allows apps to access user and meeting data through multiple application programming interfaces (APIs) (see details in § 2.1). This includes a server-to-server communication API that allows apps to query data about users and meetings (including recorded meeting content). Such data access models might lead to novel privacy threats even from apps the users never use after installation (§ 2.1). Second, being primarily a remote collaboration tool, Zoom (and marketplace apps) heavily rely on audio and visual data. Not only can many other types of sensitive information (e.g., biometric features, as well as physiological, psychological, and affective properties) be extracted from these data, but they can also be used to create deep fakes [44]. Further, third-party apps can get a comprehensive view of people's lives through their access to profile data, contact information, and details such as daily schedules and activities. Third, Zoom (and marketplace apps) has become ubiquitous in education and healthcare contexts, serving students (including minors) and healthcare consumers. Personal data created in these contexts is deemed more sensitive and is protected by specific laws in the US. Thus, investigating the privacy practices of marketplace apps is critical and urgent.

To shed light in this matter, we present analysis of the marketplace using a longitudinal dataset. We monitored the marketplace for one year (from December 2023 to December 2024) and created a comprehensive dataset containing 97,194 snapshots of all apps on the marketplace. We additionally collected and analyzed privacy policies from the marketplace or external websites of the developers. Our findings reveal growth trends in third-party applications and shifts in the popular categories. We observed potential misuse of app categories, with instances where apps encompass multiple categories, possibly to broaden their reach, often without providing relevant functionalities. Our examination of data permission requests documented a notable rise in data access requests over time, possibly beyond what is required. Notably, we find that newer apps blindly request for all available access permissions, including meeting content, which arguably contains the most sensitive data including video streams, screen shares, and chat messages. Furthermore, we analyzed privacy policies at multiple time points to discover trends in the disclosure of data collection and use. We uncovered transparency issues, such as vague data collection statements and omissions of data collection purpose. We also found that only a small number of privacy policies indicate their compliance with relevant laws (e.g., FERPA [45]). We discuss the privacy, safety,

[1] https://marketplace.zoom.us/

Proceedings on Privacy Enhancing Technologies 2025(4), 530–548
© 2025 Copyright held by the owner/author(s).
https://doi.org/10.56553/popets-2025-0143

and ethical implications of these findings. In sum, we make the following contributions:

(1) We conduct the first longitudinal investigation of the Zoom marketplace, documenting its evolution over a year from user privacy perspectives.
(2) We surface potential privacy issues—e.g., misuse of app categories, excessive data access requests, and incomplete or vague privacy policies—that could impact millions of users, notably, including students, children, and healthcare recipients.
(3) In light of these findings, we provide concrete recommendations for Zoom and outline future research directions.

We make the dataset[2] and the code base[3] for data collection and analysis public for reproducibility and support of future research.

2 Background and literature review

2.1 Background on Zoom marketplace and APIs

Zoom launched the Marketplace in October of 2018 for third-party developers to publish apps that will operate within the Zoom client for desktop/laptop and mobile platforms. Third-party apps can interact with the Zoom client and access user data in several ways. Contextual data can be accessed via the Zoom Apps SDK [6] that facilitates communication between the marketplace application and the Zoom client. In-meeting apps, embedded directly in the Zoom client, can use this JavaScript SDK to access information such as participant lists, meeting IDs, and real-time UI hooks [6]. Server-side data and events from the Zoom account, including calendar information, meeting reports, cloud recordings, and account data, are accessed via Zoom REST APIs [2]. OAuth-based apps, operating outside of meetings, leverage these APIs to access user-level, account-level, and meeting-level data [2]. Media Streams and meeting chat data are accessed via Zoom Meeting SDK [4], which uses meeting bots to connect to meetings as a participant and generate or process the media data streams. For use cases where developers want to build their own real-time audio and video applications outside of the Zoom client, the Zoom Video SDK provides direct access to Zoom's underlying media infrastructure without any Zoom UI constraints [5].

When a user first starts using an app, it receives a refresh token and an access token. The access token is valid for one hour and can be refreshed using the refresh token without the user's knowledge, allowing continued access for up to 90 days [59]. These tokens enable the app to query Zoom's servers for user data (e.g., profile information, contact details) as well as meetings created by the user, including past recorded sessions. This access occurs through server-to-server communication, which is invisible to the user and may persist even if the user never interacts with the app again. This data access model contrasts with that of mobile apps or voice assistant skills and may pose a greater level of privacy threats.

2.2 Past research on Zoom

Despite huge popularity of Zoom, surprisingly, research on privacy and security of Zoom has been scarce, while the same on third-party apps has been virtually non-existent. Achilleos *et al.* [8] analyzed video conferencing apps, including Zoom, on the Android platform and reported that they ask for many of the so-called dangerous permissions. Kagan and colleagues examined privacy and safety risks from video conferencing by analyzing publicly available video recordings of Zoom meetings to extract participants' data [27]. Sun *et al.* [43] developed computer vision and audio processing tools to remove sensitive data while remote collaboration through Zoom. Woo *et al.* [52] identified privacy risks due to pinning features during Zoom meetings. Liu and Bikzók studied the interdependent privacy issues from third-party apps on multiple platforms, including Zoom, and reported widespread use of permissions to access data about people other than app users [33]. Notably, Zoom has faced backlash over artificial intelligence (or AI) based features, such as emotion recognition [39], and its policy to use customer data to train AI models [14]. Although Zoom has backed off those plans, the marketplace hosts many third-party apps that provide similar features. Thus, a comprehensive and systematic investigation of the marketplace's evolution, apps' data collection practices, and privacy policy is critical.

2.3 Related work on other platforms

We review past research exploring the evolution of marketplaces, data request patterns, and privacy policies.

Longitudinal analysis of marketplaces. Wang *et al.* [49] investigated the evolution of the Google Play store and reported that many apps were requesting additional permissions without adding corresponding functionality, permission requests increased alongside app popularity, and the accessibility of privacy policies decreased over time. Similar longitudinal studies on iOS [32], WeChat [55], and GPT [56] platforms identified possible misuse of data permissions and privacy policy violations.

App permission evolution. The first comprehensive study on permission evolution in the Android platform was conducted by Wei *et al.* [51]. They examined 346 pre-installed and 237 third-party apps over three years and revealed that apps increased their permission requests over time, with the so-called "dangerous" permissions being the most frequent category. Also, an increasing percentage of apps (44.8%) violated the principle of least privilege by requesting permissions they did not use. Calciati and Gorla largely reproduced these findings with a larger number (n=14000) of apps [13]. Many other works (e.g., [19, 31, 38]) uncovered problematic data requests and access patterns by apps in various platforms.

Privacy policy analysis. Privacy policies document and inform data collection and sharing practices, and thus have attracted considerable research efforts. Past research has automated analysis and summarization of policy documents [17], detected inconsistencies by comparing stated policies with actual app behaviors [10, 25, 50], and identified non-compliance with regulatory measures [53]. Past research also revealed a high rate of missing or invalid policy documents [9, 18, 58] and found policy documents losing comprehensibility over time [47].

Thus, past research has made significant contributions in identifying and mitigating privacy issues on multiple platforms. Similar investigations on the emerging Zoom marketplace are urgently needed, given its recent explosive popularity.

[2] https://github.com/PERSUE-Lab-ASU/Zoom-Marketplace-Dataset
[3] https://github.com/PERSUE-Lab-ASU/Zoom-Privacy-Data-Collection-Framework

3 Methods

3.1 App marketplace data collection

From December 2023 to December 2024, we collected data about apps on the Zoom marketplace. First, we crawled the marketplace directory to compile an exhaustive list of URLs to individual app pages and then crawled those pages to collect app details and privacy policies. The crawling was executed at the beginning of each week (Sundays at midnight) to maintain data currency while minimizing disruption to the marketplace's operations. Each complete crawling session of the Zoom Directory, the individual pages of the app, and the associated privacy policies typically took approximately four hours on average, with a 10-second delay between consecutive requests to avoid exhausting the server.

For this purpose, we developed a specialized crawler based on the Puppeteer library [37] and a parser, and periodically updated them to handle technical issues and changes in the marketplace and app details page format. The first change was the introduction of app categories in the marketplace in March 2024, which subsequently underwent additional changes, such as different locations of category data on the page and how they were presented. Additionally, the privacy policy link on app pages was moved from the bottom to the top of the page. We addressed these issues by updating the parser to check the new location and using keywords to search for the link rather than solely relying on CSS tags to locate the element. We likewise updated the parser to address changes in how and where app scopes were listed. Some of the technical issues we faced were the unavailability of the server at times, slow loading of pages leading to timeout errors, and invalid or non-existent links to other pages and documents (particularly privacy policies).

We created another crawler and parser to handle a major change in the app category listing after May 2024. Previously, all categories under which an app was listed were included on the app details page. However, after May 2024, only the first category was listed, and additional categories were loaded and made visible after hovering over the category-listing area. Triggering this hovering action automatically could not be done reliably. Thus, we created a crawler that periodically visited all web pages that listed app names under specific categories (there were 32 categories in total). We also made a parser to extract app names and other details for post-processing.

Despite these technical challenges, we ensured reliability in the data collection process through extensive logging, robust error-handling mechanisms, and recovery steps for any lost data. The crawlers and parsers logged every request, as well as errors and exceptions they faced. The project lead would receive email notifications if they had to halt operation, and manually review logged messages and update the data collection framework as needed. To prevent data loss, the crawler saved all HTML pages so that even if the parser fails (e.g., due to a new change in page format), we could adapt the parser and recover data from the saved pages. We also implemented automated verification steps to enhance data accuracy and completeness. For example, after each cycle, the system compared the total number of apps listed in the marketplace directory with the number of apps for which data had been collected, ensuring that data was gathered for all available apps. There was also an edge case where an app could be created or deleted during the 4-hour data collection window. In such instances, the Data framework documented these changes in the email log sent to researchers, allowing for manual verification and recovery. The parser checked for any null values in essential data fields for each app, such as the app developer and the privacy policy. If a null value was detected, the app would be reported in the email log. We also encountered inconsistencies in how data values were displayed. For instance, "Health & Wellness" appeared as "Health & Wellness " with an extra space at the end when the categories were first added to the Zoom Marketplace. This issue has since been fixed, and we updated our dataset to ensure consistency with the new data.

These measures ensured data completeness and accuracy, with particular attention to issues such as failed page loads, missing or incomplete data, and data consistency across different phases of collection. This methodology allowed us to create a comprehensive dataset while maintaining high data quality standards and respecting the technical constraints of the platform.

3.2 Privacy policy analysis method

Privacy policy collection. Our privacy policy analysis methodology is built upon the marketplace data collection infrastructure. Using the Puppeteer library, we had already implemented for marketplace crawling; we extended our automated collection system to handle privacy policy documents. The system was configured to access the privacy policy URLs identified during the initial marketplace crawling phase, maintaining the same 10-second delay between requests to respect server limitations and implementing similar error-handling mechanisms as our main crawler. For each application in our dataset, we visited the previously stored URLs for the corresponding policy page and downloaded it (if the link was valid). For retrials and manual reviews, the framework kept logs of failure cases, e.g., due to non-existent links or any errors due to parsing or network connectivity. For example, if there were a failure in obtaining a privacy policy, the framework would automatically attempt to rerun the HTML download. If the immediate rerun also failed, the framework would add the app to the queue for another attempt in the second pass at the end of the data collection for all apps in the first pass. Apps that still had errors after the second pass were logged for manual verification and reported to researchers via email. We saved the raw HTML content for further analysis and maintained detailed logs of any failed attempts for manual verification and retry procedures.

Automated privacy policy analysis. To process the collected policy documents, we utilized PoliGraph [17], a specialized natural language processing tool that analyzes unstructured privacy policy texts to create knowledge graphs. PoliGraph identifies statements about data collection and sharing in privacy policies, and builds relationships among data, actors, and actions, such as what data is being collected, who is collecting it, and for what purposes. It then creates knowledge graphs containing nodes and links to represent these relationships.

Post processing knowledge graphs. Knowledge graphs are visually rich and facilitate manual reviews to grasp data flows, yet, our ultimate goal was to summarize data collection and sharing statistics. Thus, we developed a Python script to post-process the graphs; it enumerated graph specifications and parsed different relationships (such as generic data types like ‘contact information’ and specific data types such as ‘phone number’), identified unique data collector entities and purposes and aggregated all these results to compute high-level statistics.

Manual annotation. To ensure reliability, we complemented automatic analysis with manual annotation and verification steps at multiple stages (see the Findings section). Three authors were involved in these processes. For example, we manually reviewed the descriptions of 10% apps from each category (randomly selected) to assess the alignment between the functionalities they offer and the categories they were listed under. Two authors independently assessed the apps; any disagreements were resolved with the help of a third author (majority voting). The same approach was followed to identify permissions irrelevant to app features, examine changes in the description of an app after its categories had changed, review statements in privacy policies, and map data types mentioned in privacy policies to data permissions shown to users. Employing two independent annotators and another arbiter ensured the reliability of the annotated data.

Ethical considerations. We exercised strategies to minimize the impact on the servers from which we gathered data [20], ensuring that normal operations remain unaffected. For example, we implemented a gap of 10 seconds between two server calls. Additionally, we conducted our weekly data collection between 12 AM and 4 AM on Sundays, when server usage is presumably minimal. We also note that our research can potentially benefit Zoom by helping them identify malicious apps, which can outweigh the computation cost we incurred; we are already in the process of reaching them with our findings.

4 Findings

We report the current (as of December 2024) status of the Zoom marketplace and how privacy and security-relevant factors (e.g., data access permissions) have changed between December 2023 and December 2024. We investigated changes across different time intervals: monthly and half-yearly. For the latter, we compared three time points—December 2023, May 2024, and December 2024. Note that Zoom introduced categories in March 2024; therefore, results that rely on category data were reported by comparing data between May and December of 2024. We supplemented quantitative data with manual reviews of apps and their privacy policies to provide a deeper and nuanced understanding of privacy and security issues.

4.1 App trend analysis

4.1.1 Number of apps over time. There were 2,438 apps on the marketplace in December 2023. That number increased linearly each month and reached 2,893 by December 2024 (Figure 1). Although the trend was upward, a small number of apps were also removed (or renamed) from the marketplace each month. We used the URL for each app to determine if it was removed (the URL led to a non-existent page) or renamed (the page existed but with a different app name). We found that, between December 2023 and December 2024, 667 new apps were added, 212 apps were removed, and 49 apps were renamed. Looking across categories (introduced in March 2024), between May 2024 and December 2024, the largest number of apps were removed from the *Meeting* category (n=20), which is now no longer a category, followed by *Scheduling* (n=19) and *Education* (n=17).

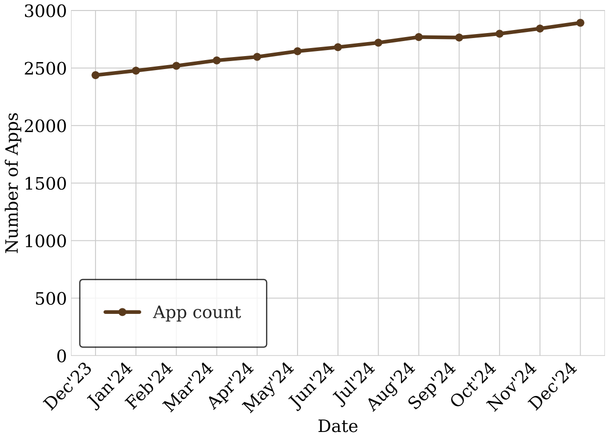

Figure 1: Monthly number of apps on the marketplace.

The marketplace had 32 app categories as of December 2024. The largest number of apps was under *Productivity* (n=556) while *Virtual Backgrounds & Scenes* had the fewest apps (n=11). The general trend of increasing the number of apps was also observed across the categories. As Figure 2 shows, almost all categories grew in the number of apps between May and December 2024, with a few (such as *Productivity* and *Scheduling*) experiencing relatively larger growth. Additionally, between May and December of 2024, six categories—*Events, Meetings, Phone, Rooms, Team Chats*, and *Webinars*—were removed, and one category (*Whiteboard*) was added. This is why these categories in Figure 2 are empty, even though the apps still exist in the marketplace under other categories. For example, #AskAway, initially listed under both *Games* and *Meetings*, is now only listed under the *Games* category.

4.1.2 Overlaps in app categories and correlation with app features. An app can be listed under multiple categories, reaching a larger potential user base. Yet, this feature can be misused to spam users as well as to ask for unnecessary permissions [40]. Investigating cross-category overlaps, we found that the number of categories per app dramatically changed between May and December 2024. In May 2024, almost 89.68% (n=2373) of apps were listed under a single category; only 205 apps had two, and 68 had three categories, respectively. The *Education* category had the largest overlaps with other categories: for 197 apps in *Learning & Development*, n=51 apps in *Scheduling*, and 34 apps in *Collaboration.*

In contrast, by December 2024, only 46% (n=1344) apps were listed in one category; 735 apps had two, 686 apps had three, and 128 had four categories. For example, apps such as Akute (*Health*) and Intellecta (*Education*) were listed under one category in May, but that changed to four categories by December 2024. Figure 3 visualizes cross-category overlaps. Generally, thematically similar categories had large overlaps, such as *Health* and *Health & Wellness*, *Education* and *Learning & Development*, and *Transcription & Translation* and *Note taking*. However, overlaps existed between seemingly

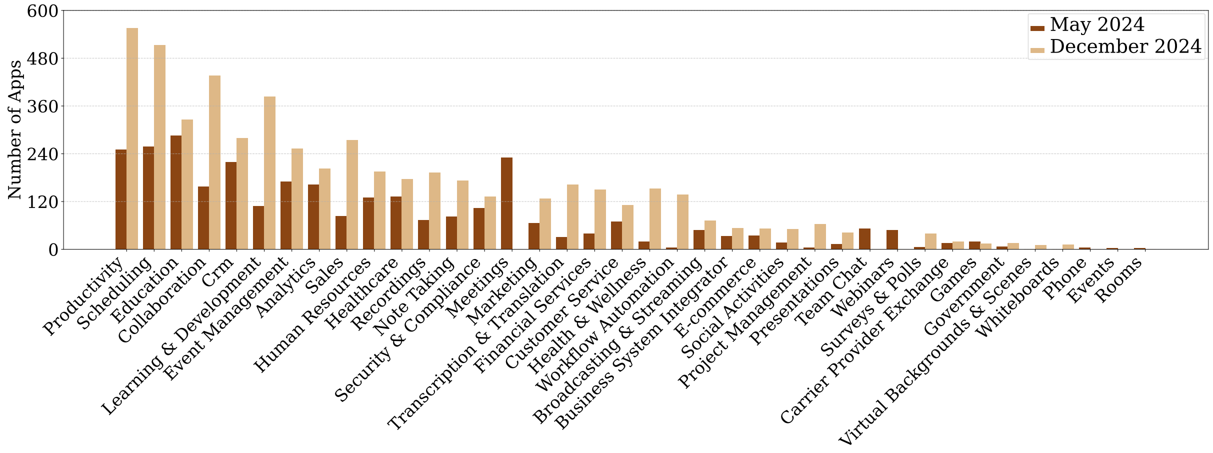

Figure 2: Change in the number of apps per category from May 2024 to December 2024.

unrelated categories, such as *Customer service* and *Learning & Development*. By manually reviewing the descriptions of 10% of apps in each category, we identified 22 potentially mis-categorized apps (see Table 9 in the Appendix). For example, WRKiiT Beta provides event management services but was cross-listed under *Health & Wellness* and *Learning & Development*. We also identified potential mismatches between the functionality an app provides and its category: for example, YouStudio and Kindred Minds provide remote class and AI-based leadership coaching services, respectively, but both were also listed under *Health & Wellness*.

Motivated by the above examples, we next investigated whether the inclusion of new categories in existing apps was accompanied by additional functionality relevant to those new categories. Since app pages include descriptions of app functionality, we examined whether apps included new categories between May and December 2024 and whether their descriptions changed within that time interval. We found that among the 2484 apps that were present in both May and December of 2024, 1356 (55%) apps added at least one new category, but only 184 revised their descriptions. Moreover, revisions for 165 of those 184 apps only included adding or removing white spaces. We compared the old and new descriptions of the remaining 19 apps and found that for 10 apps, the changes were minor (such as adding or removing punctuation or correcting misspellings) and unrelated to features. Only nine apps (Live Chat, Topicflow, Claap, Salesify, VA OAuth, LMS for Zoom, Timer, Zeplyn, and PixelMixer Assistant) revised descriptions to mention new or updated features. Looking at the categories for the 184 apps, we found that several of them appeared to have included categories without providing associated functionality; for example, Music Player - YouTube, Spotify & More from BlueSky Apps streams music, but was listed under *Healthcare* and *Event management* (in addition to *Broadcasting & Streaming*). We also manually reviewed 20 randomly selected apps (Table 11 in the Appendix) that made no change to their descriptions after including new categories. We identified two apps (Thalamus and Edit on the Spot) that were possibly misclassified. For example, Thalamus, an interview management program for Graduate Medical Education, was listed under *Healthcare* and *Health & Wellness* but did not appear to provide any health-related services. These results hint at potential spamming activities, where an app bundles unrelated categories and keywords to appear more frequently in search results and reach a larger user base [41].

Adding more categories was also accompanied by apparently unnecessary data permissions. For example, Wavoto provides Sales website templates and hosting services but was listed under *Learning & Development* and *Scheduling*, and requires view and manage access to meeting content and participants' profiles. The most popular category being added to existing apps was *Learning & Development* (n=237), followed by *Health & Wellness* (n=126). Since user data generated in education and health contexts is deemed more sensitive and is protected under additional laws such as FERPA [45] and HIPAA [46], this practice of overusing categorization and data permission requests raises privacy compliance concerns.

Takeaway #1: We find evidence of overusing or misusing categories by apps, possibly to reach more users or request more user data.

4.2 App permission analysis

4.2.1 Overall trend. Zoom has different permission categories that provide either view (read-only) or manage (edit) access to user data. This data can be associated with only the user who added an app to their Zoom client (User only) or with other people (User and others), such as meeting participants (see Table 1).

Figure 4 shows the cumulative distribution function of the total number of permissions (view and manage, combined) per app based on the latest data (collected in December 2024); the majority of the apps require between 6 to 10 permissions. Table 2 and Table 3 list the number of apps that require different view and manage permissions, respectively. For view permissions, almost all apps access the profiles and contact information of the primary users, which may include personal information about other people who are in the users' contact list (Table 1). This was closely followed

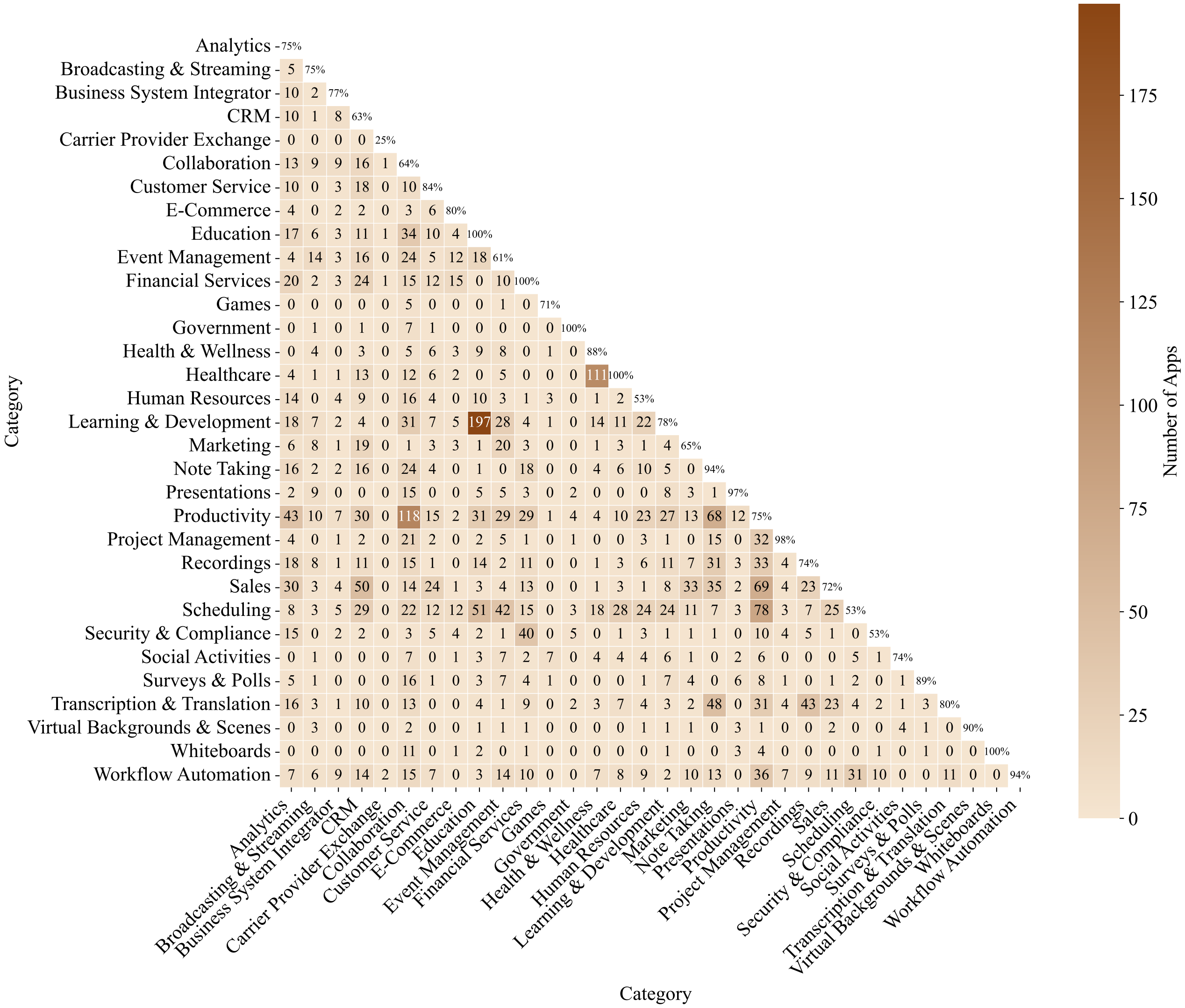

Figure 3: Overlaps in app categories (diagonal cells show the percentage of apps in a category shared with other categories.)

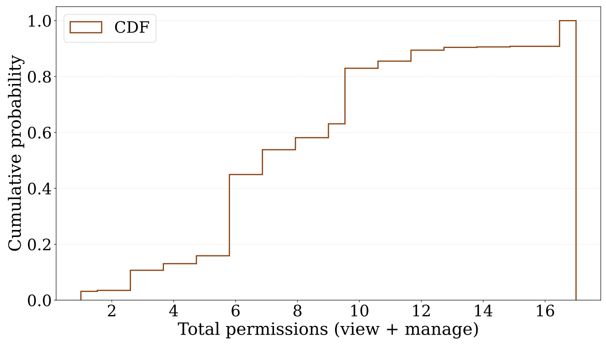

Figure 4: CDF Plot of Total Permissions per App.

by arguably less sensitive data about product usage and settings. In contrast, apps most frequently require managing permissions to meeting content, participants, and registration and scheduling information (Table 3).

4.2.2 Permissions analysis across app categories. Figure 5 shows box plots for the number of permissions required by apps in the top 20 categories that cover 97.23% (n=2,813) of all apps with the largest number of apps. Interestingly, the cross-category distributions of inter-quartile range look similar, indicating that apps, regardless of the types of functionalities they provide, request roughly the same number of permissions—between 6 and 10. The two exceptions are (*Note Taking* and *Transcriptions* and *Translations*) apps that require a slightly higher number of permissions. The overlap in categories (and hence functionality and data requirements) may partly explain this uniformity. Yet, our manual review of the permissions from 64

Permission Category	Data Accessed
Profile & Contact Information (User only)	user name, display name, picture, email address, phone number, job information, stated locale, account, user ID, contact lists
Product Usage (User and others)	when participants join/leave, whether participants sent messages and who they message, performance data
Settings (User only)	whether a passcode or a waiting room is required, permitted event capacity, screen sharing settings
Content (User and others)	audio, video, messages, transcriptions, feedback, responses to polls and Q&A, files, invitation details, meeting or chat name, and meeting agenda
Calendars (User only)	calendar of scheduled Zoom meetings and webinars
Registration Information	name and contact information, responses to registration questions
Participant Profile & Contact Information (User and others)	name, display name, email address, phone number, user ID
Functional (User and others)	Zoom user ID, session IDs, meeting role, and information about your meeting, webinar, or chat
Device Information (User only)	speakers, microphone, and camera, OS version, hard disk ID, PC name, MAC address, IP address and general location at the country level derived from it
Account Information (User only)	administrator name, account email address, billing information, and account plan information

Table 1: Permission categories and associated user data. (User only) implies data about the primary user, as opposed to also about other meeting participants.

View Permission December 2024	Count	Percentage
Profile & Contact Information	2661	92.0%
Product Usage	2584	89.3%
Settings	2540	87.8%
Content	1757	60.7%
Calendars	1587	54.9%
Registration Information	1528	52.8%
Participant Profile & Contact Information	1494	51.6%
Functional	452	15.6%
Device Information	416	14.4%
Account Information	369	12.8%

Table 2: View Permission Counts and Percentages

Manage Permission (Dec 2024)	Count	Percentage
Content	2080	71.9%
Participants	1998	69.1%
Registration & Scheduling	1938	67.0%
Settings	568	19.6%
Profile & Contact Information	487	16.8%
Account Information	278	9.6%
Devices	228	7.9%

Table 3: Manage Permission Counts and Percentages.

apps (two apps from each category) identified 8 apps that ask for data seemingly unrelated to their functionalities (see Table 10) in the Appendix). For example, Calendly for Zoom, which automatically creates video conference details and saves them to Calendly event, requests access to meeting content, such as audio, video, and messages, generated by all participants.

Takeaway #2: Regardless of categories, most apps ask for 6–10 permissions; some apps ask for data that is (apparently) unrelated to their functionalities.

4.2.3 Meeting content permissions. Meeting content (that includes audio, video, shared screens and documents, and chat messages) is arguably the most sensitive data type; a host of other information about participants, including biometrics, emotional states, psychological traits, and disability status [15, 27, 34], as well as various confidential data can be learned (e.g., from a financial document during screen shares) from them. Moreover, audio-visual data from meetings can be used to create deepfakes [44] that can have devastating consequences. Thus, we investigate content data use more closely by looking at each category, as shown in Figure 6. It is safe to say that apps in some categories (e.g., *Collaboration*) have legitimate needs to process meeting content, but for some other categories (e.g., *Scheduling*), this need is unclear. To dig deeper, we manually reviewed 10 randomly selected apps (Calendly, Skeding, Leadline, Salesforce, Close, Calero-Saas, Tote, AlignTogether.live, Niuco, and Pentugram) from two such categories, *Scheduling* and *E-commerce*. We identified instances of potentially over-permission requests. For example, three scheduling apps (Calendly, Skeding, and Leadline Connected Calendar for Zoom) provide services to automate meeting creation and invitation, and there is no apparent need, based on the app descriptions and features, for them to view or manage content generated during the meeting. Likewise, we identified two E-commerce applications (Calero-SaaS Expense Management and Niuco) that provide services to manage Zoom licenses, and it's unclear why they require access to meeting content.

Takeaway #3: Despite much heightened privacy concerns about meeting content, their access is required by many apps without a clear need.

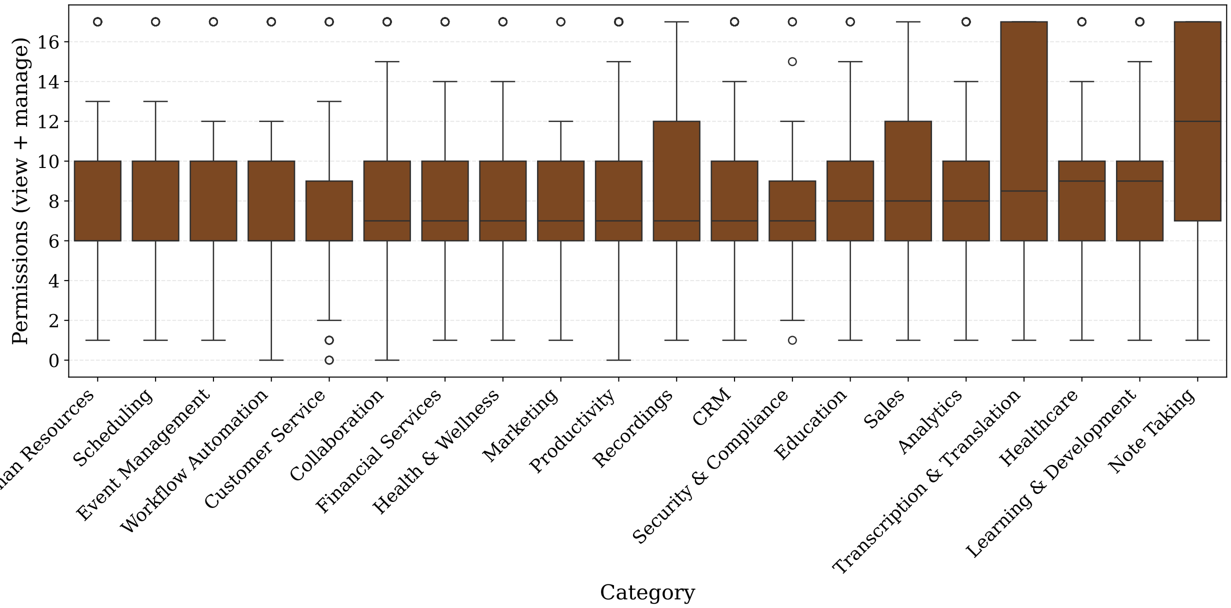

Figure 5: Permission count for top 20 categories (sorted by the median number of permissions).

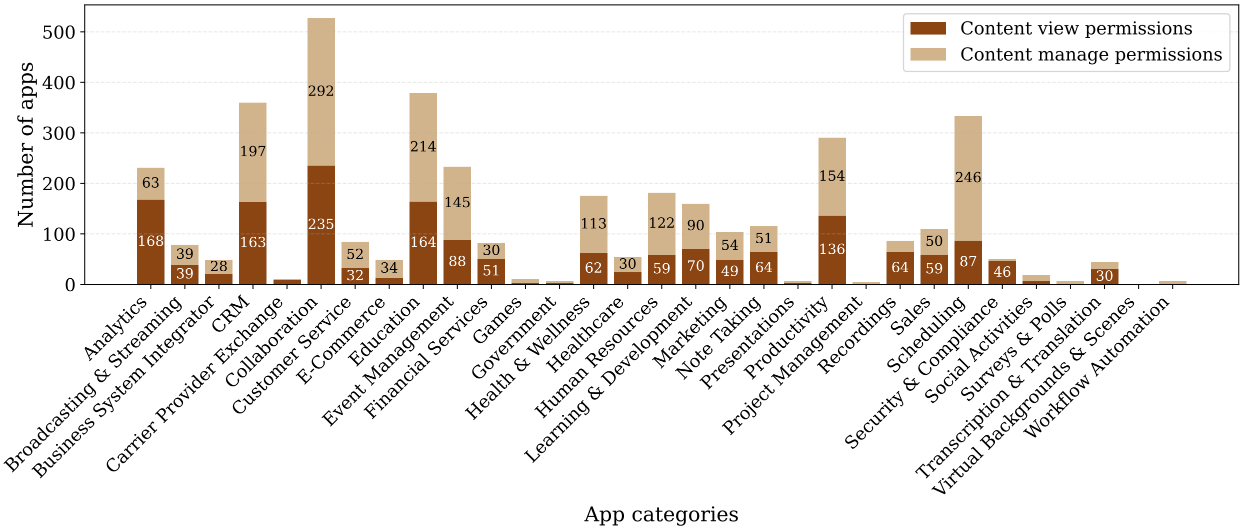

Figure 6: Number of apps in different categories that request to view or manage meeting content.

4.2.4 Changes in permissions requirements over time. Table 4 shows the number of apps that increased or decreased permissions between May and December of 2024. As the table shows, permission requests by existing apps remained relatively stable over time; few apps changed view or manage permissions between May and December 2024. In particular, only two apps (Theta Lake eComms Archive and Biznest-AI Discovery Sidekick) removed permissions, and the total number of apps that added one or more permissions is below 100 for both view and manage permissions.

Apps created after May 2024, however, tended to ask for more permissions than older apps. For example, only 4% (n=25) of apps in May 2024 asked for all view permissions, whereas 30.56% (125 out of 409) of apps added between May and December 2024 asked for all view permissions. Likewise, except for 24 apps, no other app that existed in May 2024 required any manage permissions, but 30.56% (the same 125 apps) of apps created afterward asked for all seven manage permissions. This tendency of requiring all permissions was most prevalent among *Note-taking* apps where 40.46% (70

# Permissions change	# Apps (View)	# Apps (Manage)
-5	1	0
-1	1	0
1	13	12
2	5	2
3	26	20
4	7	2
5	1	0
6	3	0
7	11	51
8	1	0
9	28	0

Table 4: View and manage permissions changes in existing Apps from May 2024 to Dec 2024

apps) asked for all permissions, followed by *Transcription & Translation* (28.22% or 46 apps)—these two categories also had the highest median number of permission requirements (Figure 5). A manual review of 10 randomly selected notetaking apps (Embra AI Notetaker, Grain, AI Product Assistant by BuildBetter.ai, Rafiki.ai for Zoom Meetings, Zocks Meeting, Otter.ai, AI Notetaker by Fathom, Jump, Gong for Zoom, and Notimo). For example, Gong for Zoom Meetings, a recording bot, only requires access to meeting content and provides transcription and meeting analysis services; yet, another app (e.g., Embra AI Notetaker) providing similar features requires all permissions.

Takeaway #4: Newly released apps require many more permissions compared to existing apps.

4.3 Privacy policy analysis

We identified and analyzed 1,831 valid privacy policies in the latest dataset (after December 2024). We could not analyze policies for the remaining 1,079 apps either because of invalid or non-existent links to privacy policies (n=978) or non-English privacy policies.

4.3.1 Data collection and sharing practices. In total, the collected privacy policies contain 43,467 statements about data collection for 7,238 unique data items (e.g., *Phone number* and *Geolocation*). Figure 8 provides high-level trends: e.g., the distribution of the number of data items per policy is highly skewed, where a majority (n=1,230) mention fewer than 20 data items (including 230 only mentioning one data item), a significantly large number of policies (n=118) specify 20–60 data items, and finally a small number (n=38) of them specify more than 100 data items.

Table 5 lists the 20 most frequently mentioned data items. Worryingly, at the top is *UNSPECIFIED_DATA*, meaning that the statements were vague and did not refer to any concrete data item. We manually reviewed 20 example apps (Table 12) with such statements in the privacy policy and identified two apps that included statements revealing data collection practices without specifying what data (e.g., "We may allow third-party advertising partners to set tracking tools to collect information regarding your activities" (PocketSuite) and " we may use Google Analytics and other analytics tools such as Fabric.io to collect and process data" (Ovatu)).

Next come Cookie and Pixel tags, which facilitate online tracking and surveillance. Various types of unique identifiers (e.g., IP addresses, email addresses, voiceprint, and personal identifiers) and quasi-identifiers (e.g., geolocation, postal address, and pseudonymized information) are frequently collected for advertising and unspecified purposes. Table 5 also contains broad and potentially vague categories, such as *Information about you*, *Internet activity*, and *Non-personal information*. They are accompanied by statements such as "We use the information we collect or receive [· · ·]" (Staircare) and "[· · ·] In general, we use the information we collect to provide and administer the Services [· · ·]"(SalesHood).

Privacy policies may refer to data at a different level than what users see in the permission dialog. For example, "contact information" in the policy can refer to either phone number, physical or email address, or all of these. To understand if all the data items mentioned in privacy policies are visible to users (through permission prompts), we manually reviewed the text corresponding to the top 100 data items mentioned in privacy policies and mapped them to permissions that are visible to users (Table 8). We also manually reviewed 10 randomly selected privacy policies (read.ai, Datadog, Insight LMS, Asana, FrontRace, 5mins.ai, Rooster, Jump, Glyue, and AvatarLink). We found that most data items can be mapped to permissions, but there are exceptions since developers can collect data through other means. For example, the privacy policies of several apps (e.g., read.ai Insight LMS, 5mins.ai) state that they may use the data collected to derive new information related to demographics, employment, and behavioral characteristics. Moreover, some developers state that they might sell the collected or derived data to brokers and other third parties (e.g., Warmly, the developer of *Nametags*). We also found that developers may collect data about their users from other sources, including business partners, social media, other customers, and data brokers (e.g., people.ai and read.ai), and these data items do not correspond to permission prompts and may go unnoticed by the users.

Takeaway #5: Collections of data that facilitate tracking, identifying, or profiling users are commonplace. Developers may also collect data from other sources or sell user data to other parties.

Data collection purposes. About 29.32% (n=12,744) of data statements about data collection did not have any specific purpose stated in the privacy policy. Table 6 shows their purposes for the remaining statements. Looking at the app level, almost all (95.68%) apps mentioned at least one purpose behind collecting data; the remaining apps did not specify any purpose for any of the data items they collected.

Table 7 shows the most frequently mentioned entities that collect or receive user data; unspecified actors are the most common recipients after the first party (i.e., developers). Most (93.83%) apps we analyzed stated in their privacy policy that they share at least one data item with third parties. Across all policy documents, there

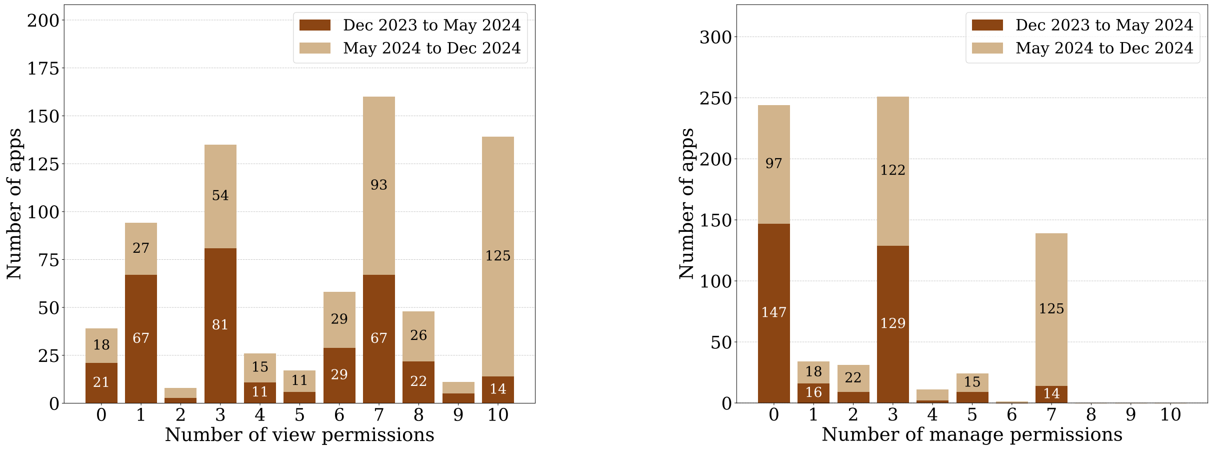

Figure 7: Distribution of view and manage permissions requested by Zoom apps created before and after May 2024.

Data Type	Total	Purpose					
		Services	Analytics	Advertising	Security	Legal	Unspecified
UNSPECIFIED_DATA	4535	1361	724	657	603	560	2678
Cookie / Pixel Tag	2078	1246	1134	744	461	404	497
Personal Information	1972	959	697	611	396	382	868
Email Address	1293	902	722	608	516	471	321
Person Name	969	677	535	448	400	341	240
IP Address	835	505	472	353	283	247	247
Geolocation	677	517	392	307	407	296	91
Contact Information	600	499	336	290	312	241	63
Information About You	598	369	296	242	208	190	192
Aggregate / Deidentified / Pseudonymized Information	515	203	177	114	80	77	243
File	392	114	106	17	15	9	273
Phone Number	370	327	272	248	212	195	21
Internet Activity	357	205	180	169	129	97	104
Postal Address	327	280	236	205	188	185	34
Credit / Debit Card Number	289	143	98	77	94	61	137
Personal Identifier	264	178	149	110	93	48	71
Information We Collect	253	192	155	138	121	114	54
Non-Personal Information	249	144	136	116	86	91	52
Usage Information	219	214	124	121	204	119	3
Browser Type	201	125	101	64	65	65	50
Identifier	198	118	89	81	116	113	30
Voiceprint	183	180	177	91	91	90	3

Table 5: Top 20 data items and the purposes for their collection. Note that each item can have multiple purposes.

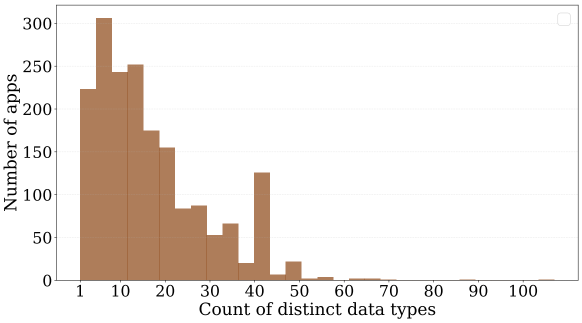

Figure 8: Histogram of the number of distinct data items mentioned in privacy policies.

Purpose	Count	Percentage
Services	4494	30.99%
Analytics	3371	23.24%
Advertising	2719	18.75%
Security	2210	15.25%
Legal	1810	12.47%

Table 6: Purposes for Data Collection.

Recipient	Count	Recipient	Count
App developer	1787	UNSPECIFIED	1331
Service Provider	360	Advertiser	304
Google	237	Analytic Provider	187
Social Media	150	Business Partner	138
Zoom	94	Invitee	93
Integration	90	Meeting Host	88

Table 7: Frequency of data recipients

were 18,862 such data-sharing statements, but only 42.26% (n=7972) were accompanied by any purpose.

> **Takeaway #6:** About 1-in-3 data collection lacks any stated purpose. Common data receivers include advertisers and business partners.

4.3.2 Change in privacy policies over time. Similar to permission requests (§ 4.2.4), we investigated if privacy policies had changed over time and if this change correlated with changes in permission requests. For this, we compared the data at two points (May and Dec of 2024) that had 1,119 common privacy policies. The majority (63%) of them were updated within this time; the rest remained unchanged. The number of unique data items in these 1,119 policies increased from 4,985 to 5,217; this increase seems commensurate with the slight rise in permission requests (Table 4).

A closer look revealed that 32% (n=358) of policies added at least one new data collection statement, while 30% (n=331) removed at least one such statement. The three most frequently added data items were voice prints (n=261), audio transcripts (n=129), and meeting content (n=129). On the other hand, top removed data items were account information (n=66), feedback provided to Zoom about product ownership (n=64), and website virtual chat data (n=64), which might indicate a reduced commitment to incorporating user input into product development and privacy practices.

Transparency in data collection also did not change much: in May, 70% of 26,295 data collection statements specified a purpose; in December, it increased to 71% (total 28,000 statements). The number of statements on sharing the collected data with third parties that specified the purpose of sharing remained constant at 42%.

> **Takeaway #7:** Privacy policy revisions resulted in the inclusion of arguably more sensitive data (e.g., voice print) and exclusion of less sensitive data (e.g., user feedback). Purpose specification and disclosures for sharing did not change.

4.3.3 Data items in privacy policies across app categories. Figure 9 shows the number of data items mentioned in the policies for apps in the top 20 categories (in terms of the number of apps). The distributions of data collection disclosures in privacy policies were less uniform across categories compared to the distributions of permission requests (Figure 5). Of particular note, the three categories with the highest number of median permissions: *Note taking*, *Learning & Development*, and *Healthcare* were less transparent in their privacy policies than other categories. This disconnect between data access and disclosure was supported by the result that the number of permissions an app requires was uncorrelated with disclosures and data collection statements in its privacy policy ($r = 0.025$, $p > .05$).

Health and education apps. We pay special attention to apps providing health and education services as the data created in those contexts is subject to additional regulatory protections such as FERPA [45], HIPAA [46], and COPPA [3], which are the US federal regulations that dictate the collection and use of educational records, health records, and data about children, respectively. Zoom, noting the increasing popularity of third-party apps among students (including minors) in their app review guideline [7], provided guidelines for app compliance with these laws. Developers can comply with those regulations either directly or by entering into a business contract with other covered entities (such as a publicly funded school).

To determine compliance with HIPAA, we first searched for related keywords (e.g., "HIPAA") in the descriptions and privacy policies of health apps. Among the 142 apps in *Healthcare* and *Health and Wellness* categories, 74%(n=105) did not mention HIPAA, 32 apps explicitly mentioned that they are HIPAA compliant (two of them are developed by Zoom). The remaining five apps delegated the responsibility for HIPAA compliance to their clients (i.e., healthcare providers and users). For example, IntakeQ, an intake form management service, specifically states that the use and disclosure of protected health information was "governed by your Provider's terms and conditions and privacy practices."

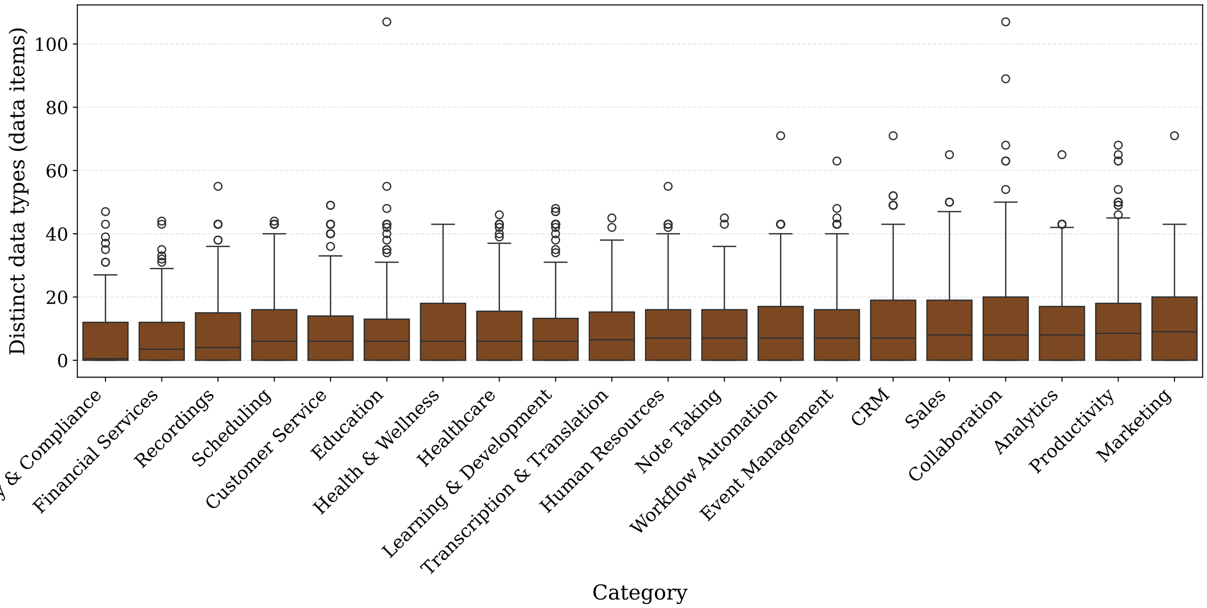

Figure 9: Distribution of the number of unique data items in the privacy policies of apps in the top 20 categories.

Similar analysis of 346 apps in the *Education* and *Learning & development* categories again revealed three compliance patterns: 306 apps did not mention either FERPA or COPPA, 18 apps affirmed compliance with both COPPA and FERPA, and the remaining 22 apps were compliant with only COPPA. Notably, FERPA-compliant apps were more likely to be directly integrated with educational institutions and student information systems as opposed to being used as standalone apps.

Takeaway #8: Disclosures about data collection in privacy policies were uncorrelated with permission requests and heterogeneous across categories. For education and health apps, compliance with relevant laws cannot be determined from privacy policies.

4.4 Limitations

Our methodologies have several limitations. First, we relied on automation to scale our analyses, e.g., using PoliGraph to analyze privacy policies. While it is state-of-the-art in this area and outperforms earlier tools by large margins [17], it still may not detect everything correctly. We focused on the US market and analyzed English privacy policies; thus, our findings may not generalize to other countries. We also could not establish the impact of potential privacy violations, as data about app downloads and user reviews are not public; yet, we can safely assume that the effect is significant since Zoom has been the most popular remote communication tool for the last few years. Finally, while we examined compliance with regulations based on privacy policies, determining which entities (e.g., an app provider) is covered under those laws requires extensive analysis from legal perspectives, and identifying violations is challenging due to exploitable loopholes [12, 30, 36]. However, we note that negative consequences from improper data collection practices remain the same regardless of whether the collector is covered under privacy regulations.

5 Discussions

This paper provides a comprehensive picture of how the Zoom marketplace evolved in a year, identifying patterns in app characteristics that may impact user privacy, security, and safety. We detected a trend of increasing use of a larger number of data access requests by apps. However, unlike other marketplaces, where existing apps add more permissions over time [18, 51], in the Zoom marketplace, we observed that existing apps rarely increased the number of permissions; rather, newer apps required more data access than older apps. Our automated privacy policy analysis identified a large number of statements about data collection, use, and sharing that include broad or vague classes of data, which may preclude holding the responsible entity accountable in case of privacy invasions. We also noticed a trend of listing existing apps under an increasing number of categories

We complemented those results with findings from manual analysis of subsets of apps. For example, comparing app categories with feature descriptions, we observed concerning practices such as mis- or over-categorization of apps, likely to attract more users. Additionally, we identified permission requests for user data likely unnecessary for the app features. Moreover, manual reviews of privacy policies revealed a lack of transparency about data collection and sharing. Finally, we found that most apps shared data with third parties (including selling or renting data, as the privacy policy of read.ai states), including advertisers and other companies known for extensive online tracking and surveillance activities.

Implications of the findings. Our findings generally agree with past research on other platforms: many apps demonstrate privacy-invasive behaviors. However, the privacy and safety harms from such practices can be more severe in this case, given that the most popular apps serve in education and healthcare domains, and the user base includes minors. When services like Zoom are institutionally procured, contracts typically restrict data use and sharing; however, past research has shown that a long chain of vendors and sub-vendors makes it extremely difficult, if not impossible, to track data use or hold entities accountable in case of privacy violations [30]. Use of these apps with personal accounts by educators for teaching purposes has also become ubiquitous [29]; as there is no institutional contract restricting data use, and most apps may not directly comply with FERPA or COPPA, such uses raise severe privacy concerns.

This situation is further complicated by the ability to infer new data using machine learning models. For example, Insight LMS, a learning management system app, states in its privacy policy that it may use the collected data to infer other information and profile users. Past research has shown that interaction data from such LMS tools can be used to predict demographic attributes, such as gender and age group [21]. Apps that access audio or video data can infer much more: recent machine learning models can be used to predict many demographics, behavioral patterns, and affective status, as well as physical and cognitive disabilities [15, 34]—most of which are protected under FERPA [45] and HIPAA [46]. In addition to raising legal compliance issues, such practices also raise ethical concerns, since these data can lead to discrimination and even incrimination of people (e.g., those who are gender non-conforming).

Lastly, Zoom apps may create interdependent privacy issues that impact people other than the app users [11], as many permissions allow accessing data about online contacts and collaborators of the (primary) app users. Capabilities like contact sharing dramatically expand the number of people who can be brought under surveillance by malicious entities. Zoom lists the data an app collects about other people on the app details page, which could raise awareness among users [28] and possibly encourage them to respect others' privacy [24]. Zoom also requires obtaining consent from the data subjects. However, Zoom users often have power asymmetry (e.g., between an instructor and students, between a hospital and patients, and so on) that may turn informed consent into a coercive one [54].

Recommendations for Zoom. To mitigate privacy and safety risks for app users, Zoom could implement stronger enforcement of its developers' guidelines. For example, in addition to encouraging the data minimization principle, Zoom could adapt methodologies proposed for other platforms (e.g., [48, 57]) to detect apps with excessive permission requests and take actions to discontinue such practices. Likewise, automated comparisons between app feature descriptions and categories, perhaps followed by manual audits, can be used to prevent mis- or over-categorization of apps. Arguably, one of the biggest threats to users' privacy comes from apps (or their developers) being able to invoke Zoom server API functions to access data about users and meetings. Such communications are invisible to the users and may happen even if the user has never used the app after installing it. To reduce the likelihood of privacy violations, either intentional or accidental, Zoom could monitor API calls by apps and possibly compare them with app uses within the Zoom client. For apps without a client component, Zoom could periodically notify users about API access and check if the users are continuing to use the associated app. Finally, several steps could potentially reduce privacy violations of people other than the users. When notifying about data permissions by an app, Zoom could contextualize messages and highlight power asymmetry and other socio-political issues to encourage more careful and selective revealing of data about other people. Additionally, before third-party apps access audio or video data, Zoom could implement technical means to selectively obfuscate content (e.g., [22, 23, 26, 42, 43]) to improve privacy without degrading functionality. Finally, instead of granting all data requests at install time, Zoom could facilitate a dynamic and need-based permission model.

Future research directions. There is a lot more to do in this important but relatively unexplored app ecosystem. While we analyzed permissions and privacy policies, future work could study app behaviors, use of permissions, and consistency of data collection with privacy policies. Additionally, network traffic analysis could at least partially reveal who receives the data that apps access.

Our legal compliance was based on a direct evaluation of privacy policies to check compliance with US-based laws. Future work could investigate compliance based on app behaviors and expand to other privacy laws at the Federal (e.g., GDPR [1]) and State level (e.g., CCPA [35]), as well as regulations specific to marketplaces (e.g., DSA [16]).

Another strand of research could help identify apps that better respect user privacy. For example, by grouping apps based on features and corresponding data requests, apps could be ranked based on the amount of data they require to provide the same set of features. Such a ranking could then be used to recommend apps on the marketplace. Other user-centric research could focus on understanding users' understanding of the threat model and developing strategies to raise awareness of privacy threats to app users and other people.

6 Conclusions

We conducted the first longitudinal investigation of the Zoom marketplace and surfaced many issues related to user privacy and safety. Yet, perhaps more than anything, we demonstrated the necessity for much more research to explore those issues in-depth, given the popularity of the apps and their reach to vulnerable populations. We hope our research will guide future investigations, and the associated code base and data that we publish will facilitate their execution.

Acknowledgments

We thank the reviewers and editor for their valuable feedback on this paper. Rakibul Hasan was partially supported by the Air Force Office of Scientific Research (AFOSR) under Award #FA9550-24-1-0227. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Air Force Office of Scientific Research.

References

[1] 2024. General Data Protection Regulation. https://gdpr-info.eu/.

[2] 2025. API reference. https://developers.zoom.us/docs/api/.

[3] 2025. Children's Online Privacy Protection Rule. https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa.

[4] 2025. Meeting SDK. https://developers.zoom.us/docs/meeting-sdk/.

[5] 2025. Video SDK. https://developers.zoom.us/docs/video-sdk/.

[6] 2025. Zoom Apps JS SDK References. https://appssdk.zoom.us/classes/ZoomSdk.ZoomSdk.html.

[7] 2025. Zoom Marketplace app review guidelines and principles. https://developers.zoom.us/docs/distribute/app-review-guidelines/.

[8] Georgios Achilleos, Konstantinos Limniotis, and Nicholas Kolokotronis. 2023. Exploring Personal Data Processing in Video Conferencing Apps. *Electronics* 12, 5 (Jan. 2023), 1247. https://doi.org/10.3390/electronics12051247 Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

[9] Hamad Alamri, Carsten Maple, Saad Mohamad, and Gregory Epiphaniou. 2022. Do the Right Thing: A Privacy Policy Adherence Analysis of over Two Million Apps in Apple iOS App Store. *Sensors (Basel, Switzerland)* 22 (2022). https://api.semanticscholar.org/CorpusID:253790360

[10] Ryan Amos, Gunes Acar, Eli Lucherini, Mihir Kshirsagar, Arvind Narayanan, and Jonathan Mayer. 2021. Privacy Policies over Time: Curation and Analysis of a Million-Document Dataset. In *Proceedings of the Web Conference 2021 (WWW '21)*. Association for Computing Machinery, New York, NY, USA, 2165–2176. https://doi.org/10.1145/3442381.3450048

[11] Gergely Biczók and Pern Hui Chia. 2013. Interdependent privacy: Let me share your data. In *International conference on financial cryptography and data security*. 338–353.

[12] Michael Brown and Carrie Klein. 2020. Whose data? Which rights? Whose power? A policy discourse analysis of student privacy policy documents. *The Journal of Higher Education* 91, 7 (2020), 1149–1178. https://doi.org/10.1080/00221546.2020.1770045 Publisher: Routledge.

[13] Paolo Calciati and Alessandra Gorla. 2017. How Do Apps Evolve in Their Permission Requests? A Preliminary Study. In *2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR)*. 37–41. https://doi.org/10.1109/MSR.2017.64

[14] Jen Caltrider. 2023. Zoom Now Says They Can Use Your Data For Training AI — What Does That Actually Mean? https://foundation.mozilla.org/en/privacynotincluded/articles/zoom-now-says-they-can-use-your-data-for-training-ai-what-does-that-mean/.

[15] Che-Sheng Chu, Di-Yuan Wang, Chih-Kuang Liang, Ming-Yueh Chou, Ying-Hsin Hsu, Yu-Chun Wang, Mei-Chen Liao, Wei-Ta Chu, and Yu-Te Lin. 2023. Automated Video Analysis of Audio-Visual Approaches to Predict and Detect Mild Cognitive Impairment and Dementia in Older Adults. *Journal of Alzheimer's Disease* 92, 3 (April 2023), 875–886. https://doi.org/10.3233/JAD-220999 Publisher: SAGE Publications.

[16] European Commission. 2025. The Digital Services Act. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en

[17] Hao Cui, Rahmadi Trimananda, Athina Markopoulou, and Scott Jordan. 2023. PoliGraph: Automated privacy policy analysis using knowledge graphs. In *Proceedings of the 32nd USENIX security symposium (USENIX security 23)*.

[18] Jide Edu, Xavier Ferrer-Aran, Jose Such, and Guillermo Suarez-Tangil. 2022. Measuring Alexa Skill Privacy Practices across Three Years. In *Proceedings of the ACM Web Conference 2022 (WWW '22)*. Association for Computing Machinery, New York, NY, USA, 670–680. https://doi.org/10.1145/3485447.3512289

[19] Jide S. Edu, Xavier Ferrer-Aran, Jose Such, and Guillermo Suarez-Tangil. 2023. SkillVet: Automated Traceability Analysis of Amazon Alexa Skills. *IEEE Transactions on Dependable and Secure Computing* 20, 1 (Jan. 2023), 161–175. https://doi.org/10.1109/TDSC.2021.3129116 Conference Name: IEEE Transactions on Dependable and Secure Computing.

[20] Florian Hantke, Sebastian Roth, Rafael Mrowczynski, Christine Utz, and Ben Stock. 2024. Where Are the Red Lines? Towards Ethical Server-Side Scans in Security and Privacy Research. In *2024 IEEE Symposium on Security and Privacy (SP)*. 4405–4423. https://doi.org/10.1109/SP54263.2024.00104

[21] Rakibul Hasan and Mario Fritz. 2022. Understanding Utility and Privacy of Demographic Data in Education Technology by Causal Analysis and Adversarial-Censoring. *Proceedings on Privacy Enhancing Technologies* 2022, 2 (April 2022), 245–262. https://doi.org/10.2478/popets-2022-0044

[22] Rakibul Hasan, Eman Hassan, Yifang Li, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. 2018. Viewer Experience of Obscuring Scene Elements in Photos to Enhance Privacy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173621

[23] Rakibul Hasan, Yifang Li, Eman Hassan, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. 2019. Can Privacy Be Satisfying?: On Improving Viewer Satisfaction for Privacy-Enhanced Photos Using Aesthetic Transforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300597

[24] Rakibul Hasan, Rebecca Weil, Rudolf Siegel, and Katharina Krombholz. 2023. A Psychometric Scale to Measure Individuals' Value of Other People's Privacy (VOPP). In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–14. https://doi.org/10.1145/3544548.3581496

[25] Saad Sajid Hashmi, Nazar Waheed, Gioacchino Tangari, Muhammad Ikram, and Stephen Smith. 2021. Longitudinal Compliance Analysis of Android Applications with Privacy Policies. arXiv:2106.10035 [cs.CR] https://arxiv.org/abs/2106.10035

[26] E T Hassan, R Hasan, P Shaffer, D Crandall, and A Kapadia. 2017. Cartooning for enhanced privacy in lifelogging and streaming videos. In *2017 IEEE conference on computer vision and pattern recognition workshops (CVPRW)*. 1333–1342. https://doi.org/10.1109/CVPRW.2017.175 ISSN: 2160-7516.

[27] Dima Kagan, Galit Fuhrmann Alpert, and Michael Fire. 2024. Zooming Into Video Conferencing Privacy. *IEEE Transactions on Computational Social Systems* 11, 1 (Feb. 2024), 933–944. https://doi.org/10.1109/TCSS.2022.3231987

[28] Bernadette Kamleitner and Vince Mitchell. 2019. Your Data Is My Data: A Framework for Addressing Interdependent Privacy Infringements. *Journal of Public Policy & Marketing* 38, 4 (Oct. 2019), 433–450. https://doi.org/10.1177/0743915619858924

[29] Easton Kelso, Ananta Soneji, Syed Zami-Ul-Haque Navid, Yan Soshitaishvili, Sazzadur Rahaman, and Rakibul Hasan. 2025. Investigating the Security & Privacy Risks from Unsanctioned Technology Use by Educators. *arXiv preprint arXiv:2502.16739* (2025).

[30] Easton Kelso, Ananta Soneji, Sazzadur Rahaman, Yan Shoshitaishvili, and Rakibul Hasan. 2024. Trust, Because You Can't Verify: Privacy and Security Hurdles in Education Technology Acquisition Practices. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security (CCS '24)*. Association for Computing Machinery, New York, NY, USA, 1656–1670. https://doi.org/10.1145/3658644.3690353

[31] Konrad Kollnig, Anastasia Shuba, Reuben Binns, Max Van Kleek, and Nigel Shadbolt. 2022. Are iPhones Really Better for Privacy? Comparative Study of iOS and Android Apps. *Proceedings on Privacy Enhancing Technologies* 2022, 2 (April 2022), 6–24. https://doi.org/10.2478/popets-2022-0033 arXiv:2109.13722 [cs].

[32] Fuqi Lin, Haoyu Wang, Liu Wang, and Xuanzhe Liu. 2021. A Longitudinal Study of Removed Apps in iOS App Store. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*. Association for Computing Machinery, New York, NY, USA, 1435–1446. https://doi.org/10.1145/3442381.3449990

[33] Shuaishuai Liu and Gergely Biczók. 2025. IDPFilter: Mitigating interdependent privacy issues in third-party apps. *Computers & Security* 151 (April 2025), 104321. https://doi.org/10.1016/j.cose.2025.104321

[34] Yash Mehta, Navonil Majumder, Alexander Gelbukh, and Erik Cambria. 2020. Recent trends in deep learning based personality detection. *Artificial Intelligence Review* 53, 4 (April 2020), 2313–2339. https://doi.org/10.1007/s10462-019-09770-z

[35] State of California Department of Justice. 2025. California Consumer Privacy Act. https://oag.ca.gov/privacy/ccpa

[36] Britt Paris, Rebecca Reynolds, and Catherine McGowan. 2022. Sins of omission: Critical informatics perspectives on privacy in E-Learning systems in higher education. 73, 5 (April 2022), 708–725. https://doi.org/10.1002/asi.24575 Number of pages: 18 Place: USA Publisher: John Wiley & Sons, Inc. tex.issue_date: May 2022.

[37] Puppeteer. 2025. Puppeteer: Headless Chrome Node.js API. https://github.com/puppeteer/puppeteer Accessed: 2023-12-26.

[38] Irwin Reyes, Primal Wijesekera, Joel Reardon, Amit Elazari Bar On, Abbas Razaghpanah, Narseo Vallina-Rodriguez, and Serge Egelman. 2018. "Won't Somebody Think of the Children?" Examining COPPA Compliance at Scale. https://dspace.networks.imdea.org/handle/20.500.12761/551 Accepted: 2021-07-13T09:33:32Z.

[39] Ryan Daws. 2022. Zoom receives backlash for emotion-detecting AI. https://www.artificialintelligence-news.com/news/zoom-receives-backlash-for-emotion-detecting-ai/.

[40] Suranga Seneviratne, Aruna Seneviratne, Mohamed Ali Kaafar, Anirban Mahanti, and Prasant Mohapatra. 2015. Early Detection of Spam Mobile Apps. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 949–959. https://doi.org/10.1145/2736277.2741084

[41] Suranga Seneviratne, Aruna Seneviratne, Mohamed Ali Kaafar, Anirban Mahanti, and Prasant Mohapatra. 2017. Spam Mobile Apps: Characteristics, Detection, and in the Wild Analysis. *ACM Trans. Web* 11, 1 (April 2017), 4:1–4:29. https://doi.org/10.1145/3007901

[42] Mohd Farhan Israk Soumik, W. K. M. Mithsara, Abdur R. Shahid, and Ahmed Imteaj. 2025. Exploring Audio Editing Features as User-Centric Privacy Defenses Against Large Language Model(LLM) Based Emotion Inference Attacks. https://doi.org/10.48550/arXiv.2501.18727 arXiv:2501.18727 [cs].

[43] Yuanyi Sun, Sencun Zhu, and Yu Chen. 2022. ZoomP3: Privacy-Preserving Publishing of Online Video Conference Recordings. *Proceedings on Privacy Enhancing Technologies* (2022). https://petsymposium.org/popets/2022/popets-2022-0089.php

[44] Rebecca Umbach, Nicola Henry, Gemma Faye Beard, and Colleen M. Berryessa. 2024. Non-Consensual Synthetic Intimate Imagery: Prevalence, Attitudes, and Knowledge in 10 Countries. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–20. https://doi.org/10.1145/3613904.3642382

[45] US Department of Education. 2025. Family Educational Rights and Privacy Act. https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html Accessed: 2025-02-26.

[46] U.S. Department of Health and Human Services. 2025. Summary of the HIPAA Privacy Rule. https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html Accessed: 2025-02-26.

[47] Isabel Wagner. 2022. Privacy Policies Across the Ages: Content and Readability of Privacy Policies 1996–2021. arXiv:2201.08739 [cs.CR] https://arxiv.org/abs/2201.08739

[48] Liuhuo Wan, Chuan Yan, Mark Huasong Meng, Kailong Wang, and Haoyu Wang. 2025. Analyzing Excessive Permission Requests in Google Workspace Add-Ons. In *Engineering of Complex Computer Systems*, Guangdong Bai, Fuyuki Ishikawa, Yamine Ait-Ameur, and George A. Papadopoulos (Eds.). Springer Nature Switzerland, Cham, 323–345. https://doi.org/10.1007/978-3-031-66456-4_18

[49] Haoyu Wang, Hao Li, and Yao Guo. 2019. Understanding the Evolution of Mobile App Ecosystems: A Longitudinal Measurement Study of Google Play. In *The World Wide Web Conference (WWW '19)*. Association for Computing Machinery, New York, NY, USA, 1988–1999. https://doi.org/10.1145/3308558.3313611

[50] Yin Wang, Ming Fan, Junfeng Liu, Junjie Tao, Wuxia Jin, Qi Xiong, Yuhao Liu, Qinghua Zheng, and Ting Liu. 2023. Do as You Say: Consistency Detection of Data Practice in Program Code and Privacy Policy in Mini-App. arXiv:2302.13860 [cs.CR] https://arxiv.org/abs/2302.13860

[51] Xuetao Wei, Lorenzo Gomez, Iulian Neamtiu, and Michalis Faloutsos. 2012. Permission evolution in the Android ecosystem. In *Proceedings of the 28th Annual Computer Security Applications Conference (ACSAC '12)*. Association for Computing Machinery, New York, NY, USA, 31–40. https://doi.org/10.1145/2420950.2420956

[52] SEUNGWON WOO, Wonho Song, and Min Suk Kang. 2024. I know you pin me: Privacy risks in user pinning of zoom video conferencing. In *14th international workshop on socio-technical aspects in security*. 14th International Workshop on Socio-Technical Aspects in Security.

[53] Anhao Xiang, Weiping Pei, and Chuan Yue. 2023. PolicyChecker: Analyzing the GDPR Completeness of Mobile Apps' Privacy Policies. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23)*. Association for Computing Machinery, New York, NY, USA, 3373–3387. https://doi.org/10.1145/3576915.3623067

[54] Tianyi Yang and Rakibul Hasan. 2024. Discovering Privacy Harms from Education Technology by Analyzing User Reviews. In *Proceedings of the 23rd Workshop on Privacy in the Electronic Society (WPES '24)*. Association for Computing Machinery, New York, NY, USA, 186–192. https://doi.org/10.1145/3689943.3695050

[55] Yue Zhang, Bayan Turkistani, Allen Yuqing Yang, Chaoshun Zuo, and Zhiqiang Lin. 2021. A Measurement Study of Wechat Mini-Apps. *Proc. ACM Meas. Anal. Comput. Syst.* 5, 2, Article 14 (June 2021), 25 pages. https://doi.org/10.1145/3460081

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

[57] Lu Zhou, Chengyongxiao Wei, Tong Zhu, Guoxing Chen, Xiaokuan Zhang, Suguo Du, Hui Cao, and Haojin Zhu. 2023. {POLICYCOMP}: Counterpart Comparison of Privacy Policies Uncovers Overbroad Personal Data Collection Practices. 1073–1090. https://www.usenix.org/conference/usenixsecurity23/presentation/zhou-lu

[58] Sebastian Zimmeck, Ziqi Wang, Lieyong Zou, Roger Iyengar, Bin Liu, Florian Schaub, Shomir Wilson, Norman M Sadeh, Steven M Bellovin, and Joel R Reidenberg. 2017. Automated Analysis of Privacy Requirements for Mobile Apps.. In *NDSS*, Vol. 2. 1–4.

[59] Zoom. 2025. Refresh an access token. https://developers.zoom.us/docs/integrations/oauth/#refresh-an-access-token

[60] Inc. Zoom Video Communications. 2021. Annual Report Fiscal 2021. https://investors.zoom.us/static-files/a17fd391-13ae-429b-8cb3-bfd95b61b007 Accessed: 2025-02-26.

A Appendix

A.1 Additional results

Data	Count	Permission	Data	Count	Permission
UNSPECIFIED_DATA	4535	-	whiteboard	173	Content
cookie / pixel tag	2078	Product Usage	meeting content	173	Content
personal information	1972	Profile and Contact Information	message send to everyone	173	Content
email address	1293	Profile and Contact Information	configuration information	172	Settings
person name	969	Profile and Contact Information	audio setting	172	Settings
ip address	835	Device Information	information on device about face	172	Profile and Contact Information
geolocation	677	Device Information	information from zoom email services	172	Profile and Contact Information
contact information	600	Profile and Contact Information	message send to meeting group chat	172	Content
information about you	598	Profile and Contact Information	message send to they	172	Content
aggregate / deidentified / pseudonymized information	515	-	information about	157	-
file	392	Content	device identifier	154	Device Information
phone number	370	Profile and Contact Information	browsing / search history	144	Product Usage
internet activity	357	Product Usage	time	133	Device Information, Calendar
postal address	327	Profile and Contact Information	commercial information	129	-
credit / debit card number	289	Account Information	password	126	Account Information
personal identifier	264	Profile and Contact Information	sensitive personal information	119	Profile and Contact Information
information we collect	253	Product Usage	company name	117	Profile and Contact Information, Account Information
non-personal information	249	-	title	115	Profile and Contact Information
usage information	219	Product Usage	metadata	114	Product Usage
browser type	201	Device Information	registration information	111	Registration Information
identifier	198	Profile and Contact Information	payment information	105	Account Information
voiceprint	183	Profile and Contact Information	audio recording	99	Content
information about purchase	176	Account Information	information browser send	98	-
information from partner	174	Account Information	calendar	96	Calendars
audio transcript	173	Content	voice	95	Content

Data	Count	Permission	Data	Count	Permission
education information	94	Profile and Contact Information	datum about you	73	Profile and Contact Information
number	94	-	information collect	73	-
inference	94	-	Google	72	-
username	93	Account Information	device information	72	Device Information
date	92	Device Information, Calendars	specific information	70	-
industry	88	Profile and Contact Information	audio information	70	Content
Facebook	87	-	operating system	70	Device Information
information regard meeting invitation	87	Calendars	content	67	Content
email metadata use for basic email delivery	87	Functional, Content	datum collect	66	-
pixel	87	-	information we receive	66	-
information about account owner	87	Account Information	customer record information	65	-
screen sharing setting	86	Settings	payment datum	65	Account Information
mimeid	86	Account Information	account information	65	Account Information
information we receive from third party partner	86	-	information you provide	61	Registration Information
meeting asset	86	Content	information we collect about you	61	Profile and Contact Information
content host on account	86	Content	feedback	59	-
header	86	Content	personal datum in structured use machine readable format	59	-
develop app	86	-	job title	58	Profile and Contact Information
reference photo	86	Profile and Contact Information	date time stamp	57	Device Information
facial geometry	86	Profile and Contact Information	analytic provider	55	-
profile information	82	Profile and Contact Information	date of birth	55	Profile and Contact Information
information you provide to we	80	Registration Information	system log	53	Device Information
option	76	Settings	device type	52	Device Information
professional information	75	Profile and Contact Information	customer data	52	Profile and Contact Information
collect datum	75	Profile and Contact Information	internet	52	-

Table 8: The most frequent 100 data items mentioned in privacy policies, along with the total number of times they are mentioned, and which Zoom permission they map to (‘-” indicates no mapping was possible or uncertain).

Category	App Names
Analytics	Everlaw, Studyplus for School, MyComplianceOffice Archive, LeadAngel Calendar
Business System Integrator	Board Director Zoom Meeting Scheduler
CRM	Default, SharpSpring From Constant Contact, immedio, Accubate, LobbyCentral Scheduling, Wiwink meetings, Letterdrop, Day.ai Meeting Bot, Spiro Notetaker, Rally Observer Room Bot
Collaboration	PLDT Provider Exchange Portal, app.conote.ai, Product insights by PingMi
E-Commerce	Fanclb.com
Education	Asana Chat
Financial Services	Showmaster

Table 9: Potentially miscategorized apps listed by category.

App Name	Category	Permission(s)	Why unnecessary
ChipBrain	Financial Services	Profile & Contact Information, Settings	Unnecessary for an app that only provides conversation analytics.
HatQuest	Games	Participants, Registration & Scheduling	In-meeting trivia game does not need control over participants or scheduling.
Epic	Healthcare	Account Information, Profile & Contact Information	Unjustified to manage user and account info for basic telehealth video integration.
Donations by Pledge	Social Activities	Participants	Donation management does not need participant-level meeting control.
Campaigns by FaithApp	Surveys & Polls	Participants	Running a poll doesn't require controlling meeting participants.
Nametags by Warmly	Virtual Backgrounds & Scenes	Participants, Registration & Scheduling	Name overlays don't need scheduling or participant management.
Vote Now	Workflow Automation	Participants, Registration & Scheduling	Polling automation doesn't require participant or scheduling control.
Calendly for Zoom	Productivity	Content (Audio, Video, Chat)	Calendly only needs to create meeting links — access to full meeting content is excessive and unjustified.

Table 10: Apps that overrequest `Manage` permissions with justifications

App name	App name
DermEngine	Cerbo by MD HQ
Zapier integration for Zoom Events	Pocket HRMS
EclipseCAT	Builderall Booking
Smartsheet Notifications	Calendars By Geniam
Zoom Phone for Zendesk by Faye	Tu Consulta Digital
Laxis - Meeting Notes and Insight	STAR Balance
Continually	Edit on the Spot
ALLO	Thalamus
Conquer Video Conferencing	Engageable
Akute Health Telehealth	Fuse Classroom

Table 11: Apps that added categories but did not change feature descriptions.

App name	App name
PocketSuite	Ovatu
30mins.com	Abi
Acadly	Airtable
Bold_Video	Cafetalk
Calendr_for_Zoom	Call_AI_by_MindTickle
CrmOne_LLC	Garba
Hammer	HRCloud
JazzHR	Momentum
nClass	Neatcal
PeerKonnect	Peoplelogic

Table 12: Apps that had statements about UNSPECIFIED_DATA collection in the privacy policy.