Nudging Developers Toward Privacy: Evaluating the Impact of Personalized App Review Reports

Sai Teja Peddinti
Google

Omer Akgul
RSAC Research

Michelle L. Mazurek
University of Maryland

Nina Taft
Google

Abstract

Mobile application developers often struggle to create accurate privacy notices or implement robust privacy practices due to limited expertise or resources. While users share unsolicited privacy feedback in app reviews, and prior research has characterized this privacy feedback, uncovering developer reactions to this feedback remains unexplored. This study explores whether personalized privacy review reports—summarizing real user feedback for a developer's own app—can effectively nudge them toward planning privacy improvements. We surveyed 42 app developers, presenting them with reports containing privacy themes, temporal trends, peer benchmarks, and emotion distributions derived from their apps' reviews. Our findings indicate that these privacy report interventions proved highly effective, with 76% (32 of 42) of participants finding at least one section of the report useful. Furthermore, exposure to the report increased the participants' intent to pursue privacy-relevant actions – such as reorganizing the UI, enhancing privacy communications, or adding/removing features – with 69% (29 of 42) of participants indicating an increased intent to do so. Almost all developers expressed a desire to receive such privacy reports periodically or on demand. These results indicate that making this style of report broadly available across the industry could foster a more privacy-conscious mobile ecosystem.

1 Introduction

Mobile applications often request access to sensitive user data (e.g., location information) to enable core functionality or to facilitate analytics, personalization, testing, and advertising. Prior research has shown that users are uncomfortable when apps request potentially unnecessary permissions [26, 42, 48] or collect/use sensitive data in unexpected ways [36, 50]. While privacy notice surfaces (e.g., privacy policies, app descriptions, permission rationales, privacy labels) are intended to clarify privacy-relevant behaviors of an app, several studies have pointed out discrepancies between an app's notice and its actual behavior [2, 21, 56], impacting user trust.

While the responsibility for drafting accurate privacy notices and implementing sound privacy practices falls on app developers, the reality of the development landscape presents significant hurdles. Many mobile applications are developed by small teams or independent developers, who often lack privacy expertise or access to legal or technical privacy knowledge [3, 4, 5]. Consequently, these developers make unintentional mistakes [43, 44] or use default privacy-unfriendly configurations when integrating third party libraries [29], hindering the development of privacy-friendly apps.

Several efforts aim to assist developers in improving their apps' privacy postures through privacy policy analysis [23, 56], source code review [22, 55] or UI evaluation [37]. One prominent category of solutions focuses on extracting security and privacy concerns directly from user reviews [6, 18, 34, 35]. By leveraging recent advances in Natural Language Processing (NLP) and Large Language Models (LLMs), these approaches can synthesize vast quantities of free-form text at a scale impossible for traditional qualitative methods. While surveys and interviews are often constrained to a few tens of participants, automated review analysis can capture unsolicited privacy-related feedback shared by millions of users. In addition to designing and implementing scalable systems that can ingest and synthesize large app review datasets, prior work has extensively characterized this user-shared privacy feedback [6, 18, 34, 35]; notably, Akgul et al. [1] analyzed 12 million privacy reviews submitted over a decade to map the evolution of users' privacy perspectives on mobile apps.

Building upon this prior research, this paper investigates

USENIX Symposium on Usable Privacy and Security (SOUPS) 2026.
August 23–26, 2026, Hannover, Germany.

the practical utility of automated privacy review analysis from the developers' perspective. Specifically, we conduct a survey to determine if personalized privacy review reports—derived from their own apps' reviews—act as a catalyst in nudging developers toward privacy-enhancing actions. We further examine the specific categories of actions developers are inclined to take, and how influential the reports are in shaping those choices. Our work is motivated by evidence that developers implement necessary app updates [35] or reduce unnecessary permissions [39] once they are made aware of users' privacy concerns with their apps.

We focus on the following research questions in this study:

- **RQ1:** To what extent do developers perceive personalized privacy review reports as useful, and which specific components—such as temporal trends, peer benchmarks, or emotion distributions—provide the most significant utility?

- **RQ2:** How does exposure to the privacy report shift a developer's intent to implement privacy-related changes, and what specific actions are they most likely to adopt?

To address these research questions, we design a personalized survey centered on reports generated for each participant's own application. By building on mock scenarios prevalent in prior work [32, 33, 52, 54], this approach leverages developers' professional involvement in their own products to more accurately gauge their motivations and reactions. The structure of these privacy reports was curated by synthesizing insights from recent app review analysis literature, with the survey instrument refined through pilot studies. We conducted this study with 42 app developers, the majority of whom are seasoned professionals with more than five years of industry experience. Our results reveal a broad spectrum of initial privacy awareness and attitudes: while some participants were dismissive of user privacy concerns, others detailed manual and automated processes they follow to monitor user privacy sentiment. Our intervention proved highly effective, as 76% (32 of 42) of participants found at least one observation in the privacy report 'Very' or 'Extremely' useful, successfully elevating their privacy awareness (RQ1). Among the four report sections, privacy review volume and temporal trends emerged as the most valued, with 25 participants citing their high utility; however, the remaining sections – theme overviews, peer benchmarks, and emotion distributions – were also consistently rated as useful (RQ1). Further, the reports served as a clear catalyst for change: 69% (29 of 42) of participants indicated they are 'More' or 'Much more' likely to take at least one of seven privacy-relevant actions examined (e.g., changing feature implementation, adding/removing a feature, communicating privacy decisions) after reviewing their privacy report (RQ2). In addition, 69% (29 of 42) of participants indicated they were 'Some,' 'More,' and 'Very much' influenced by at least one section of the report. These findings underscore the significant potential influence of personalized app privacy reports in nudging app developers to take privacy improving actions within their apps.

2 Related Work

2.1 Developer studies

Prior work has extensively explored developers' perspectives, practices, and concerns, leveraging surveys and interviews [9, 19, 20, 28]. Balebako et al. [5] examined how app developers make privacy and security decisions and emphasized the need for simple and cost-effective privacy tools (e.g., privacy checklists). Later research focused on identifying developers' privacy and security behaviors, and their motivations and challenges when implementing privacy or security [4, 19, 46]. Some more recent work explored the impact of regulations on developers' and practitioners' behaviors, finding that they mostly rely on app markets to spot privacy issues, often struggle to implement and maintain privacy labels, or leverage third-party tools to maintain compliance [3, 19, 20, 24]. Prybylo et al. [41] conduct a mixed-methods study with members of different roles (e.g., scrum masters, product managers, quality assurance) in the Software Development Life Cycle (SDLC), and show the need for role-dependent solutions to address privacy challenges.

None of the prior work is personalized towards developers' own apps, except for [24], which asks developers about inconsistencies or changes to their app's privacy/data safety labels based on independent evaluations. In contrast, this work reports results from personalized developer surveys based on user-perceived privacy concerns with apps.

Developer Recruitment Approaches Almost all of this prior research has recruited freelance developers from recruitment platforms such as Prolific [1] [41], UpWork [2] [19], SurveyXact [3] [9], or via social media or online posting sites, such as LinkedIn, Twitter, Craigslist, or Reddit [5, 20, 28, 46]. Freelance developers are a great resource for learning about general practices or perceptions, as they are easier to recruit and scale to a few hundreds of developers.

On the other hand, email outreach is the preferred approach for focused studies where specific developers (e.g., who published apps in a certain app category or who followed a particular practice) are sought [3, 24]. The developers' email addresses are often obtained from Google Play app listings or gathered via online search. In this work, we recruited developers via email to complete personalized surveys based on privacy reviews of their apps, resulting in a diverse sample.

[1] https://www.prolific.com/
[2] https://www.upwork.com/
[3] https://rambollxact.com/surveyxact

Privacy Forums Some developer studies have analyzed data on forums such as Stack Overflow and Reddit to identify the types of help (e.g., ensuring compliance, implementing Privacy-by-Design, enforcing confidentiality) sought by developers when creating applications [45, 47], and how discourse happens in response to external events such as operating system changes or privacy laws [25, 38]. Prior work has also compared the utility of these developer forums to LLMs (such as ChatGPT), showing that the forums are generally more accurate [10]. While we do leverage LLMs in the current work, we do not focus on the developer help seeking behavior.

2.2 Vignettes, Mockups, and Personalized Surveys

Researchers have extensively leveraged mockups, vignettes, or personalized surveys to isolate contextual variables driving user privacy behaviors. Early studies utilized text-based vignettes to model user privacy preferences across diverse IoT contexts [31], while others employed diagramatic mockups to deconstruct user comfort with background resource accesses in mobile applications [49]. Methodologies eventually shifted toward interactive simulations; for instance, functional Android mockups demonstrated the utility of rationales during permission decisions [15], and a simulated web-based App store helped validate the usability of expandable-grid privacy labels [54]. Similar approaches have surfaced contextual justifications for data transparency in augmented reality [17], identified misalignment in user expectations regarding account deletion [27], and measured user privacy expectations of emerging IoT technology [52]. To bridge the gap between hypothetical and actual behavior, researchers increasingly turned to Experience Sampling Methods (ESM) and personalized surveys. While some utilized ESM to uncover reasons influencing users' permission decisions [7], others employed personalized surveys grounded in actual permission logs to contextualize those choices [51]. Notably, [53] contrasted vignettes with experience sampling to highlight the discrepancies between stated and situated privacy preferences.

Despite this rich literature on user preferences, there remains a significant gap in understanding the developer's perspective through similar situated methods. Early efforts in this space relied primarily on prototypes and role-playing scenarios to study the creation of authentication systems [32, 33]. More recently, Khandelwal et al. [24] conduct personalized surveys to understand developer perspectives and challenges when engaging with data safety sections. We build upon this trajectory by introducing a novel, feedback-driven approach; to the best of our knowledge, our work is the first to employ personalized survey instruments—grounded in real-user feedback—to evaluate the effectiveness of privacy reports as nudges for increasing developers' likelihood of pursuing privacy-relevant actions.

2.3 Review Analysis

A few works have explored user-written reviews to uncover their privacy and security perceptions of an app or a (IoT) product [18, 34, 35, 40]. Earlier research relied on simple supervised machine learning models, such as SVM, to detect security- or privacy-relevant reviews [6, 8, 30, 35], and correlated them with app updates [35] or ratings [6]. Later research leveraged deep learning methods or LLMs to identify and summarize privacy-relevant reviews [1, 18, 34]. We do not propose yet another solution for review analysis while conducting personalized surveys with developers, but instead create privacy reports based on privacy review detection and topic clustering by the work of Akgul et al. [1].

3 Method

Studying developer reactions to privacy reports can be achieved through real privacy reports—where we show developers reports for their own app; or mock reports—where we ask developers to assume they built a particular app and received a privacy report for that app. Developer feedback from privacy reports grounded in real-world user feedback specific to their own applications is likely to be more ecologically valid. However, ecological validity comes with trade-offs. Not all developers with a privacy report might be willing to participate in the study. Further, there is less consistency in what we ask each developer because different apps may not have the same privacy issues, impacting internal validity.

On the other hand, using mock reports could enable recruiting a larger sample (including freelance developers and developers from both Android and iOS platforms). Mock reports would also allow us to show the same report to all participants, increasing internal validity and the chance of finding small effects. However, this study design would reduce ecological validity by requiring developers to consider an app they are not familiar with, with specific privacy issues that may not resonate with their actual app(s) and experience. Matching app developers with privacy reports for functionally similar apps might mitigate this somewhat, but would still leave a gap between the privacy report and the developer's actual experience.

Hence, we formulate our study as a personalized one where we solicit feedback from developers after showing their own app's privacy report. The realistic setup of a personalized study provides a valuable opportunity to identify the actual utility of surfacing privacy review reports to developers. We use standardized questions and report structure to reduce confounds. We opt for a survey study (instead of interviews), as it enables us to easily scale our measurement.

3.1 App Privacy Reports

Design of the Personalized Privacy Report To ensure the personalized privacy reports were effective within the constraints of a survey, we designed them to be quickly scannable and highly informative. We began by consolidating insights from recent app review analysis literature [1, 18]. Prior work has demonstrated the utility of organizing privacy reviews into high-level thematic topics, visualizing temporal trends, and leveraging granular *app-type* data (e.g., "Rideshare & Taxis" within the "Maps & Navigation" app category) to facilitate comparative analysis across functionally similar apps.

Rather than overwhelming developers with all available data, we curated the content to maintain a manageable survey duration while maximizing relevance. An app's reviews may touch upon multiple topics, called *privacy themes*. We focused on surfacing only the top three privacy themes by review volume, so that the developer gets a quick overview of users' top privacy concerns with their app. Each theme is described by a short title (e.g., 'Unneeded Location Access'), as well as a multi-sentence description to clarify what the theme entails and some representative user reviews for each theme.

To keep the survey short, we provided detailed metrics for only one specific theme chosen from the top three. This theme was selected based on the participant's input, to ensure it addressed an area where they did not already have strong pre-existing convictions and hence were more likely to have a shift in intent after a single intervention. The detailed granular metrics shared for a theme include the volume and percentage of theme-specific reviews relative to the app's total privacy reviews, as well as the percentage change in the theme size over the preceding year. Additionally, the theme data included review emotion distribution across 9 emotion groups – based on Ekman's taxonomy [14] and grouped according to criteria by Demszky et al.'s [11]. Furthermore, prior research has demonstrated that comparison to functionally similar peers is a strong motivator for developers to adopt privacy-relevant changes [39]. Inspired by this, we utilized app-type information to generate benchmarks against functionally similar competitors. The peer benchmarks are intended to help a developer understand if the privacy theme is unique to their app or a problem for everyone else in the app's functional category. Our dual metric approach uses two lenses to help illustrate an app's standing. The *percentile ranking* metric illustrates how the app ranks in comparison to competitors. A higher percentile means more competitors have fewer privacy complaints than the app. For example, a percentile of 80% means that 80% of similar apps have fewer privacy reviews for the shown theme. Our second metric is a relative comparison capturing the *magnitude* of the gap. It compares an app's theme volume directly to the peer group's average. For example, 'relatively 100% more' means the app has double the privacy theme reviews compared to the average competitor.

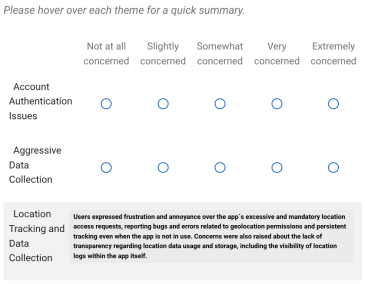
Consolidating these design choices, we organized each app privacy report into four distinct sections: (i) an overview of the top three privacy themes by volume, followed by a detailed look at a single theme, including its title, summary, and representative user reviews; (ii) the volume and percentage of privacy reviews within that theme, highlighting annual growth or decline; (iii) a peer benchmark showcasing a comparison of the app's privacy reviews in that theme against functionally similar apps; and (iv) the distribution of user emotions across the privacy reviews for that specific theme. A visualization of these sections as they appeared in the participant surveys is provided in Figure 1, and an anonymized sample privacy report is shared in Appendix, Section C. The report presentation was validated through pilot studies (Section 3.4).

Report Generation Because the deep learning models and app review datasets leveraged in prior work are not publicly available, we requested assistance from the authors of [1] for report generation. This was particularly vital given that the definition for app type or a functionally similar app is proprietary; it likely underpins the 'Similar Apps' section in the Play store, which appears when visiting a particular app. After detailing our specific requirements for the four report sections, the authors of [1] generated the reports and shared them with us via a formal data transfer agreement between our institutions. This yielded aggregated privacy review reports for ∼7000 applications, providing the foundational data required to populate our personalized surveys. This apps dataset satisfies several diversity constraints; the selected applications span a wide array of Play Store categories and exhibit significant variance in both installation numbers and total review counts. Each aggregated app privacy report we received identified the top three privacy themes by volume, and included the detailed metrics for all three themes. Based on each participant's input during the survey, detailed metrics for one theme are presented to the participant.

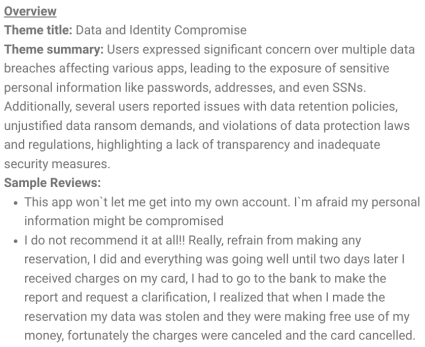
3.2 Participant Recruitment

We recruited participants in two waves through completely different channels. In the first wave, we crawled the Play store pages of the 7000 apps to obtain the contact email ID provided for each app, and sent out invitations to take our personalized survey. The email text states that the survey aims to understand developer reactions to personalized reports summarizing the privacy related comments in user reviews for their own app. We opted against advertising the study as 'general purpose' to prioritize recruiting privacy stakeholders for an app over broader, less relevant developers. Given the high volume of research invitations that developers receive, lead-in transparency was a necessary strategic choice to ensure sufficient response rate from the right personnel.

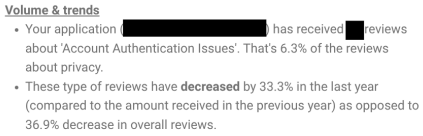
The recruitment email contains a unique (unguessable) survey link – which contains a key to help retrieve and populate the survey based on the app's privacy review report. We in-



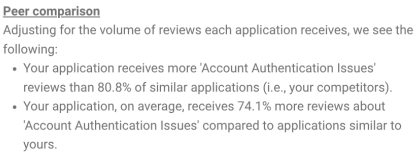
(a) Top three privacy themes



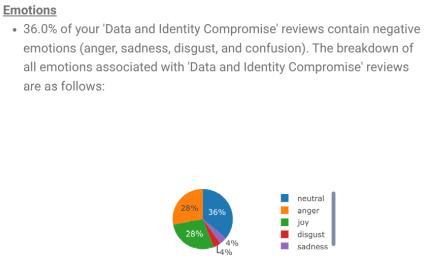
(b) Selected Theme overview



(c) Volume and temporal trends



(d) Peer benchmark metrics



(e) Emotion distribution

Figure 1: Visualization of App Privacy Report in the Participant Surveys

form the developer that only they can see this report, and that the full report will be available for download [4] at the end of the survey. In addition, upon successful completion of the survey, we offered a $30 Tango gift card (redeemable as other gift cards). The median time taken to complete the survey was 22 minutes, [5] and the monetary compensation for taking the survey is well above what is traditionally offered to developers on other survey platforms (e.g., Prolific). This first wave of outreach started in Sept 2024 and ran for a few months, during which we reached out to developers in batches. We only obtained 17 responses (less than 0.25% response rate) during this multi-month outreach.

For the second wave, we tapped into a developer interest list maintained by Google. Developers who are interested in participating in user studies and testing new unreleased Google Play product features proactively enroll with their email IDs and the products/apps they develop. From this list we identified a few hundred mobile app developers for whom we have privacy review reports. We obtained newer versions of the privacy reports for these developers' apps from the authors of [1] (as more than 6 months has passed since the first-wave reports were generated), and sent out personalized survey invitations to the enlisted email IDs. Due to program policies of the interest list, the developers in the second wave were not offered any monetary compensation for taking the survey. The second wave outreach resulted in 25 additional responses, resulting in a total of 42 responses that we investigate to understand developer perceptions toward personalized privacy review reports.

3.3 Survey Instrument

To evaluate whether and how privacy review reports are helpful in raising developers' awareness toward their apps' privacy practices, we structured the survey into three stages. In the first stage, we identify how developers currently assess users' perceptions of privacy within their application, how concerned they are about it, and how likely they are to take an action (e.g., add a feature, limit data collection, etc.) in response to their current understanding.

In the second stage, we show developers the privacy report for their app and solicit their reactions. Specifically, we inquire how useful they found different sections of the report, if they are willing to take an action based on what they learned in the report, and how influential was each observation in the privacy report in taking these actions.

In the third and final stage, we ask about their demographics and professional experience. We inquire about the number of

[4] Note that a developer already has access to all the reviews for their app, and the privacy report is just organizing the review content differently.

[5] This includes time taken to carefully review the personalized report for participants' own applications, not just to answer the survey.

years they have worked on mobile application development in general and specifically on the app referenced in the study, what roles have they served in the development of the app, whether they had a role that affects privacy-relevant aspects of the app, whether their company has a system for addressing privacy-relevant feedback, and how big their organization is (we separately inquire about the number of employees who primarily work on the app referenced in the study). The complete survey questionnaire is shared in Appendix A, and the survey's flow is described in Appendix B Figure 7.

3.4 Pilot Studies

We conducted two pilot studies with members of our own institution who have had prior app development experience and published apps on the Google Play Store. Since we didn't have personalized privacy reports for apps the pilot participants developed, we conducted mock studies where we picked a functionally similar app and asked participants to assume the survey and the privacy report are for their own app.

Based on the feedback obtained from these pilot studies, we updated the survey script. Specifically we made six improvements: (i) instead of reporting the actual number of privacy reviews received, which is unhelpful without the context of the total number of reviews, we switched to use percentages; (ii) clarified that the app benchmarking is with respect to other functionally similar apps, and is not a global benchmark; (iii) removed a question inquiring about how (un)concerned is the developer about each of the sections in the privacy report, as it was very similar to the usefulness question; (iv) since there are many sections in the privacy report, we added report snippets throughout the survey (which appear once hovered) to make it easy for the developer to understand which section is being referred to in each question; (v) Added 'wish to not answer' option for demographic questions inquiring about company/team sizes; and (vi) lastly we rephrased some questions for clarity.

3.5 Qualitative analysis

To contextualize trends within the quantitative survey results, we employed qualitative content analysis to systematically categorize the open-ended responses. The primary author reviewed all entries and utilized inductive coding to identify recurring themes, such as 'dismissive,' 'lacking process,' and 'active monitoring.' Each response was subsequently assigned to these mutually exclusive categories. While the high conceptual clarity and targeted nature of the responses rendered the initial categorization unambiguous, the resulting themes and code frequencies were critically reviewed and validated by all co-authors to ensure descriptive accuracy and interpretive consensus.

4 Ethics Statement

This research adheres to our institutions' ethical guidelines. We obtained Institutional Review Board (IRB) approval before launching the survey, and adhered to Google's policies when utilizing the developer interest list. Across both waves, all participants agreed to a thorough consent form that included information about the investigators, the risks, benefits, compensation (or lack thereof), and data confidentiality. All participants were informed about their voluntary participation and their right to withdraw at any time. No personally identifiable information was collected other than emails, which were used only to distribute gift cards after successful survey completion. Measures were in place to ensure the anonymity, confidentiality, and security of responses. We report only de-identified quotes and avoid including details that could reasonably re-identify participants. The contact information of the investigators and the IRB team was shared with the participants in case they had concerns about the study or the compensation, but no participant contacted the investigators or the IRB.

5 Limitations

Our findings are based on participant self-report data, and so are affected by participants' self-report bias, recall bias, and social desirability bias. The survey study invitation explicitly requests reactions to privacy-related user reviews, and this may introduce self-selection bias, as only developers who are concerned about the privacy of their app may participate. However, at baseline (Section 7.1) a third of our sample (14/42) reported being 'Not at all concerned' about users' privacy perceptions. Note that there might be different levels of bias between the two waves of recruitment, however this is hard to quantify. The presence of fake reviews could influence our privacy reports. We rely on existing Play protections to remove such fraudulent reviews [13, 16].

Our choice of personalized surveys over mock surveys severely restricted the list of developers we could reach out to. Due to our IRB requirement of not sending more than one email to each email ID, we couldn't send reminder emails to app developers to participate in the survey, which may have severely impacted the recruitment. While we offered monetary compensation for the first wave of recruitment, we couldn't do so for the second wave due to the policies around using the developer interest lists. However, looking at the response rates of the two waves, we don't see this as a limiting factor. Beyond monetary incentive, we also offered developers an option to download their full privacy review report at the end of the survey, and this might have acted as a strong incentive. All likely actions that developers indicated they might take after looking at the privacy report might not actually happen in practice, however, we consider these responses to be a good proxy of how they are thinking to change their

app based on the privacy reports.

6 Participants

We received 42 survey responses from app developers across the two recruitment waves. We did not observe differences in responses between the two (email and interest-list) recruitment cohorts. Consequently, we merged the two into a single dataset. Below we dive into the demographics describing our participants.

Size Range (# of Employees)	# of Participants (Organization)	# of Participants (App Team)
1	7	10
2–5	15	16
6–10	5	5
11–20	4	5
21–100	4	3
101–500	5	1
501–2000	1	0
Wish to not answer	1	2

Table 1: Participants' organization and app development team sizes.

Table 1 shows the sizes of the participants' organizations and their app development teams. We see that 27 (of 42) participants are from organizations with fewer than 10 employees, of which 7 indicate that they are the sole individual developing and maintaining the app. Only 6 participants are from organizations with 100 or more employees. The organizational size distribution of our participants aligns with a recent large-scale study of 600 app developers, which found that 70% of respondents originate from 'micro-sized' companies with fewer than 10 employees [12]. Table 1 also shows that the size of the core team developing an app is even smaller, compared to the size of the organization. In fact, one participant shared that their app development team has just 2–5 members, when the organization has 101–500 employees.

Years of experience	# of Participants (Developing apps)	# of Participants (Surveyed app)
1	2	4
2-5	6	12
5-10	11	14
10-15	19	12
15+	4	0

Table 2: Distribution of Participants' responses to years worked on mobile applications generally vs. the specific surveyed application.

Table 2 shows how long the recruited participants have worked on general mobile application development, and specifically in the development of the app referenced in their survey. Only 8 (of 42) participants have less than 5 years experience in mobile application development, indicating that most are 'seasoned/senior' professionals. More than half of participants (26 of 42) have worked for more than 5 years on the specific app referenced in their survey, showcasing that they are potentially knowledgeable about their app and are in a position to make appropriate privacy decisions based on the shared app review reports.

Role	Count
Software engineer	28
Quality assurance	24
Marketing	24
Project manager	23
UI/UX designer/researcher	22
Technical manager	21
Other	12

Table 3: Roles that participants served in the development of their survey app. Note that each participant can serve in multiple roles.

Table 3 lists the roles participants served in the development of their apps. Prior work has shown that many mobile apps are developed by small teams, with team members often assuming multiple roles in the software development life cycle (SDLC) of the app [3, 4, 5]. We see this in our dataset, where 32 (of 42) participants have selected more than one role. Of the 10 participants who assumed only one role: 4 are Software engineers, 2 are in Marketing, and 4 indicated 'Other' job roles, such as Compliance Operations, VP of Engineering, and Legal. Twelve participants selected all 7 role options (including 'Other'), and the reasoning they shared was "*I'm the founder and the solo person maintaining the app.*"

When asked specifically about their role in privacy-relevant aspects of the app, the majority (27 of 42) reported being the main privacy decision maker for their app (Table 4). In fact, all the participants reported making privacy decisions, implementing privacy features, or having other roles that impact privacy-relevant aspects. This further affirms that our participants are the right audience for the shared app privacy reports, and that our methodological choice to explicitly advertise the study as being privacy-related helped recruit the apps' privacy stakeholders.

We asked participants to elaborate on how their role impacted privacy-relevant aspects of the app; only a few shared responses. One participant in a Compliance Operations role wrote that their "*role is responsible to keep compliance between the market regulation, the store policies and the game needs*," and a participant in Marketing role pointed out that

Privacy Role	Count
I'm the main privacy decision maker	27
I help make privacy-relevant decisions.	17
I implement privacy relevant features.	14
I have another role that impacts privacy-relevant aspects.	5

Table 4: Distribution of privacy roles. Note that each participant can serve in multiple roles.

they "*alert [team] on new requirements and such.*" Other responses reflected working alone or in small teams: one shared they "*built [app] from scratch and have maintained it over the last 10 years . . . helped author our privacy policy and implemented our data safety disclosure, app permissions, and privacy-related user experiences.*" Another answered: "*We are a small team, and I am the final decision-maker. My influence is absolute.*"

In response to a question asking whether the application development team has a system for addressing users' privacy-relevant feedback, the majority (30 of 42) answered 'Yes,' 8 answered 'No,' and 4 opted not to answer. This suggests that these participants may already have some process in place to identify and evaluate privacy-relevant app reviews, which we explore further in Section 7.1.

To summarize, most of our participants come from small organizations with fewer than 10 employees, are seasoned professionals with more than 5 years of app development experience, serve multiple privacy-relevant roles within the organization, and are often the main privacy decision maker for their app.

7 Findings

In this section we uncover developers' perceptions towards privacy in general and their reactions after seeing the privacy reports for their apps.

To contextualize the findings, it is vital to note the scale and diversity of the applications supported by the surveyed participants. The 42 surveyed apps span 15 distinct Play Store categories – led by Games (6), Tools (5), Productivity (5), Music & Audio (4), and Communication (4). These 42 apps discuss 65 distinct privacy themes, of which 37 themes appeared in only one app. The most frequently occurring privacy themes include: 'Excessive Permissions' (6 apps), 'Aggressive Data Collection' (6 apps), and 'Data Theft' (5 apps). These applications command massive real-world adoption, with install counts ranging from 100K to over 100M. This includes a heavily weighted tier of popular applications: 7 apps with over 100M installs, 5 with 50M–100M installs, and 17 with 10M–50M installs. This broad distribution ensures that our findings reflect the insights from developers supporting popular applications across diverse Play store categories and having reviews that touch upon diverse privacy themes.

7.1 Existing Privacy Practices and Perceptions

We first asked our participants how they currently assess their users' perceptions of privacy within their application, and participants shared a wide spectrum of responses. A couple (2 of 42) were dismissive: "*There are no privacy concerns at all, the average use doesn't care*" or "*Most don't care.*" Others (13 of 42) indicated that they already do everything necessary: "*They have full information about this on the store listing*," "*People are generally happy with privacy. We don't ask too much permissions. We allow easily to export or delete their data*," or "*The privacy of our app users' data is taken with utmost priority, none of the data is used for any other purposes, everywhere in the app where we request access to specific user data, we have provided consent and usage information for clarity.*"

A few (5 of 42) acknowledged that they "*don't have an actual process to validate how our users handle privacy*," "*currently do not have a methodology in place*," or "*hope [their] users perceive privacy in [app] as adequate.*" The rest (22 of 42) explicitly listed actions they currently undertake or data sources they tap into. These range from manual checks (e.g., "*randomly check the feedback*" or "*scroll through negative reviews*"), utilizing multiple feedback channels (e.g., "*see what they post on socials or if they contact customer support*," "*user interviews, market place reviews, customer support tickets*," "*monitor mails, app reviews, and community forums to identify recurring privacy concerns or misunderstandings*"), to relying on external support (e.g., "*use third-party analyzers*," "*utilize external review monitoring tools to keep tabs on trends related to privacy . . . utilize internal AI-based summarization and notification for privacy based customer inquiries*.")

Based on their current understanding of users' privacy perceptions within their app, we asked participants how concerned (or not) they are, and their responses are shown in Figure 2. Only 13 (of 42) participants were very or extremely concerned about users' privacy perceptions of their app. On the other hand, roughly an equal number (14 of 42) of participants were not at all concerned. Mapping these responses to the prior question about privacy assessment practices, we see a clear trend — participants who are dismissive of user privacy concerns or think they already do everything necessary with respect to privacy also indicated they are not at all concerned. In contrast, participants who are very or extremely concerned are actively employing various techniques, tools, or data sources to understand users' privacy perceptions.

7.2 Likelihood of Taking a Privacy Action

To establish a baseline for participant intent, we asked how likely participants were to take specific privacy-relevant actions within their app based on their current understanding

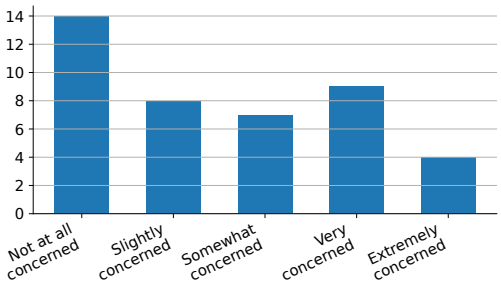


Figure 2: Participant responses to the question: 'Based on your current understanding of users' privacy perceptions within your application, how concerned or not concerned are you?'

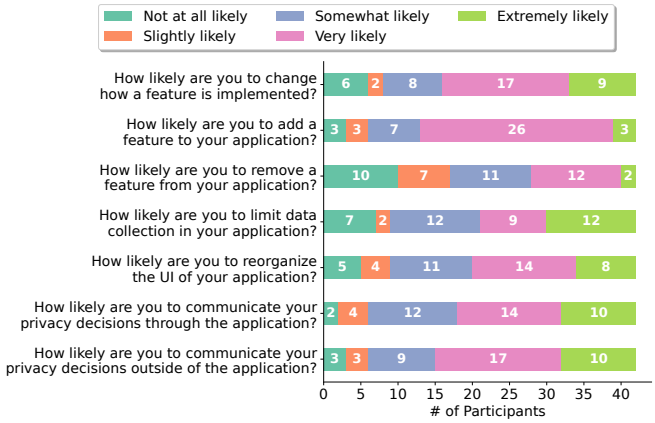


Figure 3: Responses to the question 'How likely are you to take the following actions in response to user privacy perceptions?' This was asked before showing the privacy review report to the participant.

of user perceptions. We evaluated seven potential interventions and asked participants to indicate their likelihood on a 5-point Likert-type scale, ranging from 'Not at all likely' to 'Extremely likely'. As illustrated in Figure 3, 'Adding a feature,' 'communicating privacy decisions outside of the application,' and 'changing a feature' emerged as the top three priorities. Notably, more than 50% of participants expressed a high likelihood ('Very' or 'Extremely' likely) of adding or changing features, reorganizing the UI, and improving privacy communications both within and outside the app. While 'removing a feature' and 'limiting data collection' garnered comparatively less interest, a substantial subset of participants (14 of 42) nonetheless expressed a high likelihood of pursuing these more restrictive interventions. Significantly, while specific actions received 'Not at all likely' ratings from some participants, no participant was dismissive of all privacy actions. This suggests that **all** participants, even those who appeared privacy unconcerned, are receptive to certain app modifications to enhance the app's privacy. While participants exhibited a higher readiness to implement 'low-friction' improvements—such as enhancing privacy communications—it is notable that one-third of the cohort expressed a willingness to pursue more substantial, 'product-altering' interventions, including the removal of specific features.

7.3 Usefulness of the Privacy Report

Following the assessment of participants' initial perceptions and their baseline likelihood of taking privacy actions, we presented them with a personalized privacy report for their app, as detailed in Section 3.1. Participants subsequently rated the utility of each report section on a 5-point Likert scale, ranging from 'Not at all useful' to 'Extremely useful.' For comparative context, we also assessed the usefulness scores of existing practices for the 22 participants who reported actions they currently undertake to understand users' perceptions (Section 7.1). This comparison allows us to benchmark the perceived value of the automated reports against established practices.

As illustrated in Figure 4, a majority of participants (25 of 42) rated the volume and temporal trends section of the report as 'Very' or 'Extremely' useful. This was followed by theme overviews and peer benchmarks (17 of 42 each), while emotion distributions was found highly useful by 16 participants. Collectively, 76% (32 of 42) of participants found at least one section in the privacy report 'Very' or 'Extremely' useful, demonstrating the substantial utility of the privacy reports.

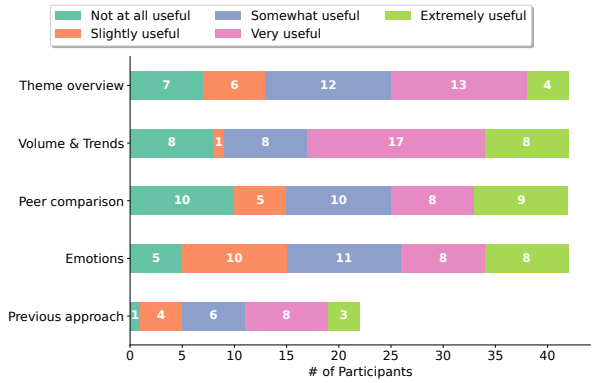


Figure 4: Responses to the question 'Please rate how useful or not useful you think each of the items are when understanding your privacy feedback'.

Interestingly, only 11 of 22 participants rated their existing practices as 'Very' or 'Extremely' useful. We hypothesize that this stems from the greater control participants maintain when navigating raw reviews via automated methods, as opposed to aggregated summaries and limited quotes we provided in our

reports. Analysis of qualitative responses outlining current practices (Section 7.1) supports this hypothesis; those who favored their existing approach often used third-party analyzers or external review monitoring tools. This preference suggests that while aggregate privacy review summaries are beneficial, many participants desired deeper visibility into user privacy experiences. The remaining 11 participants who rated their existing practices as not very useful primarily employ manual processes to randomly check negative reviews, go through customer support tickets, emails, or social media posts.

Finally, the two participants who gave responses dismissive of user privacy (e.g., stating "*Most don't care*") rated all the report sections as 'Not at all useful'. This underscores that a minority of developers remain indifferent even when presented with direct evidence of user privacy concerns.

7.4 Likelihood of Taking a Privacy Action After Seeing the Report

Following the exposure to the personalized privacy reports, we reassessed participants' likelihood of pursuing the same seven potential interventions as in Section 7.2. To explicitly measure the shift in intent from the baseline established in Section 7.2, we utilized a 5-point Likert scale ranging from 'Much less likely' to 'Much more likely.' Participant responses are illustrated in Figure 5, and we can clearly see the report serving as a catalyst for potential change: every potential intervention saw at least 10 participants reporting an increased likelihood ('More' or 'Much more' likely) of implementation. For certain interventions, such as reorganizing the UI and improving privacy communications both within and outside the application, this number rose to 20 participants. Notably, the reports successfully nudged developers toward actions that initially garnered lower interest (Section 7.2); for instance, the number of participants more likely to 'remove a feature' and 'limit data collection' reached 10 and 16, respectively. Overall, 69% (29 of 42) of participants indicated they are 'More' or 'Much more' likely to take at least one of the seven privacy-relevant actions after reviewing their privacy report. Notably, four of these participants exhibited a dramatic reversal in intent, moving from being 'Not at all' or 'Slightly' likely to act before seeing the report to 'More' or 'Much more' likely afterwards. Since 52% of participants (22 of 42) already had established feedback processes, these observed effects are attributable to our report's specific design rather than mere exposure to privacy feedback, underscoring the significant influence of our personalized reports in shifting developer priorities towards privacy-centric improvements.

Participants provided diverse justifications for their increased likelihood of taking action, primarily centered on improving user communication. Several realized that "*not explaining privacy decisions directly to users on the app leaves room for them to seek this information [elsewhere] which may result in inaccurate or straight out misinformation.*" To address this, developers expressed a need to "*make the UI explain things better*" and more "*effectively [communicate] our privacy engineering principles*". Beyond communication, the reports prompted specific technical shifts: some aimed to improve reliability by providing "*more precise information to the user when something goes wrong*," while others intended to reorganize the UI to ensure users are "*aware that they can turn accesses off when they want*". Further, the reports acted as a signal for increased functionality — such as adding a "*password protected documents*" feature in response to user requests — and for regulatory compliance, noting that features impacting privacy standards must be removed to "*comply with the stores' policies*."

A second key observation is that for all proposed actions, the highest concentration of participants remained in the 'Neither more nor less likely' rating. This plateau in intent likely stems from the strong baseline established in the initial assessment (Figure 3); many participants had already expressed a 'Very' or 'Extreme' likelihood of taking these actions before viewing the privacy report, leaving little room for measurable upward shift in opinion. Conversely, a small number of participants reported a decreased likelihood ('Less' or 'Much less' likely) of taking privacy-relevant actions after reviewing the report. While data to fully explain this reduced intent is limited, free-text responses from those who selected 'Less,' 'Much less,' or 'Neither more nor less' likely across all seven actions provide several insights. Some participants expressed technical helplessness, stating there was "*no possible action to take*" due to dependencies on third-party SDKs. Additional barriers included prior resolution (e.g., "*We have removed that for future games*"), concerns regarding data volume (e.g., "*useful, but ... very little data*"), and even the outright dismissal of user concerns, with one participant remarking, "*You can't just react because one of your 30000 [Daily Active Users] thinks he is being spied upon.*"

We explicitly asked participants to evaluate the extent to which specific sections in the report influenced their likelihood of taking action, using a 5-point Likert scale ranging from 'No influence at all' to 'Influenced very much.' Participant responses are illustrated in Figure 6. By aggregating responses of 'Some,' 'More,' and 'Very much' influence, we assessed the total number of participants affected by each section. Our results show that theme overviews, volume and trends, and peer benchmark comparisons were all highly influential, with each section impacting at least 26 of 42 participants. While the influence of emotion-based sections was comparatively lower, they still informed the decision-making of 19 participants. Overall, 69% (29 of 42) of participants indicated they were 'Some,' 'More,' and 'Very much' influenced by at least one section of the report. Notably, only eight participants indicated that all of the report sections had 'No influence at all' on their potential actions. A closer look showed that one participant was dismissive of privacy, but the rest seven already expressed 'Very' or 'Extreme' likelihood of

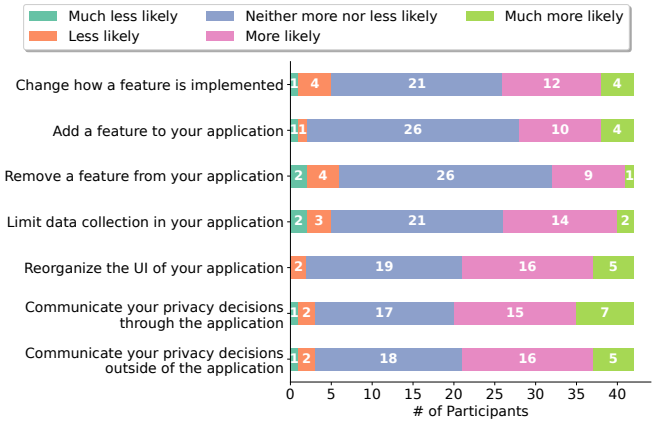


Figure 5: Responses to the question 'Now that you have seen the privacy report...How much more or less likely are you to take the following actions?' This was asked after showing the privacy review report to the participant.

taking at least one privacy action before viewing the privacy report. These findings suggest that aggregated privacy reports can serve as effective nudges for motivating developer action.

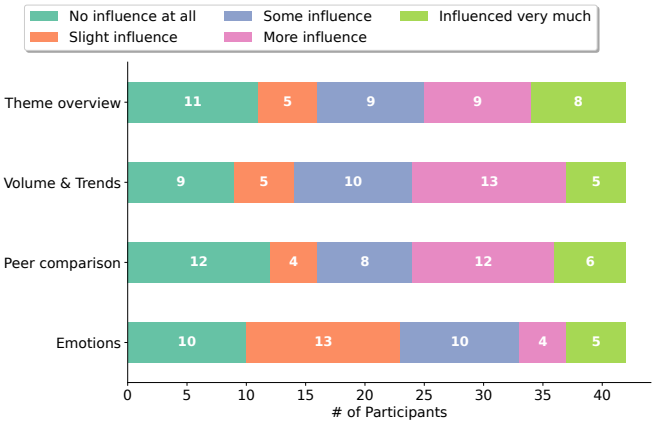


Figure 6: Influence of each section in the privacy report

7.5 Report regularity

To gauge the long-term utility of this type of privacy report, we asked participants to select their preferred frequency and method for accessing the reports. Participants could select multiple options, including periodic intervals (weekly, monthly, annually), event-driven notifications (upon sudden changes), or on-demand access. As shown in Table 5, nearly all (39 of 42) expressed interest in receiving these reports. Only three participants indicated they 'never' wanted to access such reports; notably, these individuals indicated they were 'Not at all concerned' about users' privacy perceptions during our baseline assessment (Section 7.1).

Participant preferences regarding report delivery were diverse. Specifically, 64% prefer receiving reports at regular periodic intervals (weekly, monthly, or annually), while 60% valued the ability to access the report on-demand. Additionally, 50% indicated an interest in event-driven notifications triggered by significant changes in the trends. Regardless of the delivery method or frequency, these preferences emphasize that developers value privacy analytics not just as static summaries, but as actionable, automated tools essential for monitoring and addressing the evolving user privacy needs.

Report Preference	Count
Look up the report on demand.	25
Receive a report when there is a sudden change	21
Monthly reports.	17
Weekly reports.	11
Annual reports.	8
Never want to see such reports.	3

Table 5: Preferences for Privacy Review Reporting Frequency

8 Discussion

Ours findings have important implications for future research and industry deployments. We discuss how data-driven, personalized privacy reports may be particularly effective, how greater customizability could increase utility, considerations for deployment at scale, and whether these reports benefit all developers equally.

Data Driven Privacy Advice The privacy communication we serve to participants is grounded in developer-specific evidence: app reviews. This is in contrast to static privacy advice and nudges developers (or end users) might receive [4]. By translating user feedback into a report of personalized themes and trends with contrast to the current state of the market (i.e., similar apps), our approach offers a *data-* and *competition-informed* form of privacy advice. Together with evidence from prior work [39], our results suggest this form of advice can be highly motivating in triggering privacy-protective change intent. However, a longitudinal analysis is required to determine how many developers translate this intent into the implementation of privacy enhancements. Akin to the work of Nguyen et al. [35], we believe a continuous assessment framework — correlating privacy related code changes with sentiment shifts in user reviews — is essential. Establishing an active feedback loop, where developers can observe positive user sentiment trickle in following a privacy improvement, would likely provide a powerful incentive for sustained engagement.

Interactive and Customizable Privacy Reports The findings in Section 7.5 and Table 5 reveal a distinct preference for periodic or interactive delivery of app privacy reports, underscoring a developer proclivity for monitoring evolving user needs. While aggregated privacy reports demonstrated substantial utility (Section 7.3), participants also expressed a strong desire for greater control when navigating user feedback, specifically the ability to explore individual reviews or drill down into the data, as supported by third-party analyzers or external monitoring tools (Section 7.3).

We further observe that while participants assign different usefulness ratings to different report features, no single feature dominated across all participants. This suggests that the resonance of specific features is highly contextual, likely depending on the unique challenges a developer faces with their particular application. Consequently, a multi-faceted report, such as ours, that captures a breadth of information—ranging from temporal trends to peer benchmarks—is more likely to provide value to a broad developer cohort. While priorities identified in our work establish a strong baseline on what to present to users, the most valuable features ultimately depend on an individual developer's goals and priorities. Consequently, future report interfaces should be interactive and customizable, providing "drill-down" capabilities that allow developers to investigate the granular data driving high-level trends or the ability to prioritize report features (e.g., peer comparison over emotions).

Further, as several participants rely on third-party analyzers, external monitoring tools, and alternative data sources — such as support emails or community forums — future platforms should offer integration support for these external inputs. Such an extensible architecture would ensure that review analysis tools remain relevant to a broad developer audience.

Scaling Privacy Nudges Since our results demonstrate that personalized privacy reports are effective at raising awareness (Section 7.3) and serve as a catalyst for potential privacy improvements (Section 7.4), industry adoption might be justified. Platform owners, such as Google Play Store or Apple App Store, are ideal candidates for surfacing these personalized reports, as they maintain the primary infrastructure for app reviews and developer interaction. Developers already use such consoles to publish applications and monitor performance metrics (e.g., ratings and reviews). Beyond first-party platforms, there is significant potential for independent third parties to generate and disseminate these reports. The technical feasibility of this approach is supported by the existence of review scrapers [6], published review datasets on platforms like Kaggle [7], and established auditing services like AppCensus [8] and Exodus Privacy [9]. App privacy review analysis could represent a valuable extension of these existing auditing services. Regardless of the delivery medium, our results confirm strong developer interest in accessing such insights both periodically and on demand (Section 7.5).

Does everyone benefit? Despite the utility of personalized privacy reports, some developers remain "unmoved" or "helpless" in the face of user concerns. "Technical helplessness" often arises when privacy issues are rooted in third-party SDKs over which the developer has limited control (Section 7.4 and [29]). This highlights a critical need for increased transparency and "privacy ratings" for SDKs, allowing developers to identify and select privacy-friendly alternatives. Further research is also needed to identify incentives for the minority of developers who remain indifferent even when presented with direct evidence of user dissatisfaction.

Finally, given that many developers are from 'micro-sized' companies with fewer than 10 employees, resource constraints are a significant barrier (Section 6). A cost-effective, automated tool providing ongoing privacy assessments is highly desirable for this demographic. We anticipate that surfacing personalized reports within app stores or via third-party assessment tools could bridge this resource gap and foster a more privacy-conscious mobile ecosystem.

9 Conclusion

This research demonstrates that personalized, aggregate privacy review reports are a powerful tool for surfacing users' privacy preferences/concerns and nudging app developers to make privacy improvements in their apps. We surveyed 42 app developers with varied initial privacy awareness, and 76% (32 of 42) of these participants found at least one section of the privacy reports very or extremely useful. The reports successfully "nudged" 69% of the surveyed developers to potentially take a privacy-relevant action that can improve their app's privacy posture. Specifically, report sections highlighting volume and temporal trends, peer benchmarks, and theme overviews were found most influential in driving the intent to change features or improve transparency. Developers expressed a strong interest in receiving these reports periodically or on-demand. Though participants found the aggregate privacy reports useful, we learned of their preferences to have interactive and customizable reports that support navigability. Integrating such privacy-relevant insights into developer workflows—whether through app stores or independent auditing platforms—could provide the necessary incentive and guidance for developers to prioritize user privacy throughout the app development process.

[6] https://pypi.org/project/google-play-scraper/ and https://pypi.org/project/app-store-scraper/

[7] https://www.kaggle.com/

[8] https://appcensus.io/

[9] https://exodus-privacy.eu.org/en/

Acknowledgments

We thank Kat Krol, Kevin Flores, and Nia Castelly for helping us with the second wave of participant recruitment. This project was supported in part by NSF grant CNS-1943215.

References

[1] Omer Akgul, Sai Teja Peddinti, Nina Taft, Michelle L. Mazurek, Hamza Harkous, Animesh Srivastava, and Benoit Seguin. A Decade of Privacy-Relevant Android App Reviews: Large Scale Trends. In *33rd USENIX Security Symposium*, August 2024.

[2] Mir Masood Ali, David G. Balash, Monica Kodwani, Chris Kanich, and Adam J. Aviv. Honesty is the Best Policy: On the Accuracy of Apple Privacy Labels Compared to Apps' Privacy Policies. *arXiv preprint*, 2024. https://arxiv.org/abs/2306.17063v2.

[3] Noura Alomar and Serge Egelman. Developers Say the Darnedest Things: Privacy Compliance Processes Followed by Developers of Child-Directed Apps. *Proceedings on Privacy Enhancing Technologies*, 2022.

[4] Rebecca Balebako and Lorrie Cranor. Improving App Privacy: Nudging App Developers to Protect User Privacy. *IEEE Security & Privacy*, 12(4), 2014.

[5] Rebecca Balebako, Abigail Marsh, Jialiu Lin, Jason I Hong, and Lorrie Faith Cranor. The Privacy and Security Behaviors of Smartphone App Developers. *DOI: http://dx. doi. org/10.1184*, 1, 2014.

[6] Andrew R Besmer, Jason Watson, and M Shane Banks. Investigating User Perceptions of Mobile App Privacy: An Analysis of User-Submitted App Reviews. *International Journal of Information Security and Privacy (IJISP)*, 14(4), 2020.

[7] Bram Bonné, Sai Teja Peddinti, Igor Bilogrevic, and Nina Taft. Exploring decision making with Android's runtime permission dialogs using in-context surveys. In *13th Symposium on Usable Privacy and Security*, July 2017.

[8] Lei Cen, Luo Si, Ninghui Li, and Hongxia Jin. User Comment Analysis for Android apps and CSPI Detection with Comment Expansion. In *PIR@ SIGIR*, 2014.

[9] Asmita Dalela, Saverio Giallorenzo, Oksana Kulyk, Jacopo Mauro, and Elda Paja. A Mixed-method Study on Security and Privacy Practices in Danish Companies. *arXiv preprint*, 2021. https://arxiv.org/abs/2104.04030v1.

[10] Zack Delile, Sean Radel, Joe Godinez, Garrett Engstrom, Theo Brucker, Kenzie Young, and Sepideh Ghanavati. Evaluating Privacy Questions From Stack Overflow: Can ChatGPT Compete? In *2023 IEEE 31st International Requirements Engineering Conference Workshops ESPRE*. IEEE, 2023.

[11] Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. GoEmotions: A dataset of fine-grained emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, July 2020.

[12] Department for Science, Innovation and Technology. App Developer Survey. GOV.UK, 2024. Published 5 December 2024.

[13] Harry Domanski. Google's deleted millions of fake reviews from the Play Store in the past week. https://www.techradar.com/news/googles-deleted-millions-of-fake-reviews-from-the-play-store-in-the-past-week, December 2018. Accessed: June 15, 2026.

[14] Paul Ekman. An Argument for Basic Emotions. *Cognition & emotion*, 6(3-4), 1992.

[15] Yusra Elbitar, Michael Schilling, Trung Tin Nguyen, Michael Backes, and Sven Bugiel. Explanation Beats Context: The Effect of Timing & Rationales on Users' Runtime Permission Decisions. In *30th USENIX Security Symposium*, 2021.

[16] Google Help. Flag an app or review on Google Play. https://support.google.com/googleplay/answer/2853570. Accessed: June 15, 2026.

[17] David Harborth and Alisa Frik. Evaluating and Redefining Smartphone Permissions with Contextualized Justifications for Mobile Augmented Reality Apps. In *17th Symposium on Usable Privacy and Security*, 2021.

[18] Hamza Harkous, Sai Teja Peddinti, Rishabh Khandelwal, Animesh Srivastava, and Nina Taft. Hark: A Deep Learning System for Navigating Privacy Feedback at Scale. In *2022 IEEE Symposium on Security and Privacy*. IEEE, 2022.

[19] Stefan Albert Horstmann, Samuel Domiks, Marco Gutfleisch, Mindy Tran, Yasemin Acar, Veelasha Moonsamy, and Alena Naiakshina. "Those things are written by lawyers, and programmers are reading that." Mapping the Communication Gap Between Software Developers and Privacy Experts. *Proceedings on Privacy Enhancing Technologies*, 2024.

[20] Leonardo Horn Iwaya, Muhammad Ali Babar, and Awais Rashid. Privacy Engineering in the Wild: Understanding the Practitioners' Mindset, Organizational Aspects, and Current Practices. *IEEE Transactions on Software Engineering*, 49(9), 2023.

[21] Akshath Jain, David Rodriguez, Jose M Del Alamo, and Norman Sadeh. ATLAS: Automatically Detecting Discrepancies Between Privacy Policies and Privacy Labels. In *2023 IEEE European Symposium on Security and Privacy Workshops*. IEEE, 2023.

[22] Vijayanta Jain, Sepideh Ghanavati, Sai Teja Peddinti, and Collin McMillan. Automated Generation of Accurate Privacy Captions From Android Source Code Using Large Language Models. *arXiv preprint*, 2026. https://arxiv.org/abs/2601.06276v1.

[23] Rishabh Khandelwal, Asmit Nayak, Paul Chung, and Kassem Fawaz. The Overview of Privacy Labels and their Compatibility with Privacy Policies. *arXiv preprint*, 2023. https://arxiv.org/abs/2303.08213v2.

[24] Rishabh Khandelwal, Asmit Nayak, Paul Chung, and Kassem Fawaz. Unpacking Privacy Labels: A Measurement and Developer Perspective on Google's Data Safety Section. In *33rd USENIX Security Symposium*, 2024.

[25] Tianshi Li, Elizabeth Louie, Laura Dabbish, and Jason I Hong. How Developers Talk About Personal Data and What It Means for User Privacy: A Case Study of a Developer Forum on Reddit. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3), 2021.

[26] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions. In *12th Symposium On Usable Privacy and Security*, 2016.

[27] Yijing Liu, Yan Jia, Qingyin Tan, Zheli Liu, and Luyi Xing. How Are Your Zombie Accounts? Understanding Users' Practices and Expectations on Mobile App Account Deletion. In *31st USENIX Security Symposium*, 2022.

[28] Laura MacLeod, Andreas Bergen, and Margaret-Anne Storey. Documenting and sharing software knowledge using screencasts. *Empirical Software Engineering*, 22(3), 2017.

[29] Abraham H Mhaidli, Yixin Zou, and Florian Schaub. "We Can't Live Without Them!" App Developers' Adoption of Ad Networks and Their Considerations of Consumer Risks. In *15th Symposium on Usable Privacy and Security*, 2019.

[30] Debjyoti Mukherjee, Alireza Ahmadi, Maryam Vahdat Pour, and Joel Reardon. An Empirical Study on User Reviews Targeting Mobile Apps' Security & Privacy. *arXiv preprint*, 2020. https://arxiv.org/abs/2010.06371v1.

[31] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. Privacy Expectations and Preferences in an IoT World. In *13th Symposium on Usable Privacy and Security*, 2017.

[32] Alena Naiakshina, Anastasia Danilova, Christian Tiefenau, Marco Herzog, Sergej Dechand, and Matthew Smith. Why Do Developers Get Password Storage Wrong?: A Qualitative Usability Study. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017.

[33] Alena Naiakshina, Anastasia Danilova, Christian Tiefenau, and Matthew Smith. Deception Task Design in Developer Password Studies: Exploring a Student Sample. In *14th Symposium on Usable Privacy and Security*, 2018.

[34] Preksha Nema, Pauline Anthonysamy, Nina Taft, and Sai Teja Peddinti. Analyzing User Perspectives on Mobile App Privacy at Scale. In *Proceedings of the 44th International Conference on Software Engineering*, 2022.

[35] Duc Cuong Nguyen, Erik Derr, Michael Backes, and Sven Bugiel. Short Text, Large Effect: Measuring the Impact of User Reviews on Android App Security & Privacy. In *2019 IEEE Symposium on Security and Privacy*. IEEE, 2019.

[36] Elleen Pan, Jingjing Ren, Martina Lindorfer, Christo Wilson, and David Choffnes. Panoptispy: Characterizing Audio and Video Exfiltration from Android Applications. *Proceedings on Privacy Enhancing Technologies*, 2018.

[37] Shidong Pan, Thong Hoang, Dawen Zhang, Zhenchang Xing, Xiwei Xu, Qinghua Lu, and Mark Staples. Toward the Cure of Privacy Policy Reading Phobia: Automated Generation of Privacy Nutrition Labels From Privacy Policies. *arXiv preprint*, 2023. https://arxiv.org/abs/2306.10923v1.

[38] Jonathan Parsons, Michael Schrider, Oyebanjo Ogunlela, and Sepideh Ghanavati. Understanding Developers Privacy Concerns Through Reddit Thread Analysis. In *Joint Proc. of REFSQ-2023 Workshops, Doctoral Symposium, Posters & Tools Track and Journal Early Feedback co-located with the 28th Int. Conf. on Requirements Engineering: Foundation for Software Quality*, 2023.

[39] Sai Teja Peddinti, Igor Bilogrevic, Nina Taft, Martin Pelikan, Úlfar Erlingsson, Pauline Anthonysamy, and Giles Hogben. Reducing Permission Requests in Mobile Apps. In *Proceedings of the Internet Measurement Conference*, 2019.

[40] Taufiq Islam Protick, Sai Teja Peddinti, Nina Taft, and Anupam Das. LLM-Powered Analysis of IoT User Reviews: Tracking and Ranking Security and Privacy Concerns. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2026.

[41] Maxwell Prybylo, Sara Haghighi, Sai Teja Peddinti, and Sepideh Ghanavati. Evaluating Privacy Perceptions, Experience, and Behavior of Software Development Teams. In *20th Symposium on Usable Privacy and Security*, 2024.

[42] Jingjing Ren, Martina Lindorfer, Daniel J. Dubois, Ashwin Rao, David Choffnes, and Narseo Vallina-Rodriguez. Bug Fixes, Improvements, ... and Privacy Leaks: A Longitudinal Study of PII Leaks Across Android App Versions. In *The 25th Annual Network and Distributed System Security Symposium*, 2018.

[43] Irina Shklovski, Scott D Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014.

[44] Matthew Smith. Usable Security—The source awakens. San Francisco, CA, January 2016. USENIX Association.

[45] Mohammad Tahaei, Julia Bernd, and Awais Rashid. Privacy, Permissions, and the Health App Ecosystem: A Stack Overflow Exploration. In *Proceedings of the 2022 European Symposium on Usable Security*, 2022.

[46] Mohammad Tahaei, Alisa Frik, and Kami Vaniea. Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2021.

[47] Mohammad Tahaei, Kami Vaniea, and Naomi Saphra. Understanding Privacy-Related Questions on Stack Overflow. In *Proceedings of the 2020 SIGCHI Conference on Human Factors in Computing Systems*, 2020.

[48] Lynn Tsai, Primal Wijesekera, Joel Reardon, Irwin Reyes, Serge Egelman, David Wagner, Nathan Good, and Jung-Wei Chen. Turtle Guard: Helping Android Users Apply Contextual Privacy Preferences. In *13th Symposium on Usable Privacy and Security*, 2017.

[49] Daniel Votipka, Seth M Rabin, Kristopher Micinski, Thomas Gilray, Michelle L Mazurek, and Jeffrey S Foster. User Comfort with Android Background Resource Accesses in Different Contexts. In *14th Symposium on Usable Privacy and Security*, 2018.

[50] Primal Wijesekera, Arjun Baokar, Ashkan Hosseini, Serge Egelman, David Wagner, and Konstantin Beznosov. Android Permissions Remystified: A Field Study on Contextual Integrity. In *24th USENIX Security Symposium*, 2015.

[51] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, and Konstantin Beznosov. The Feasibility of Dynamically Granted Permissions: Aligning Mobile Privacy with User Preferences. In *2017 IEEE Symposium on Security and Privacy*. IEEE, 2017.

[52] Maximiliane Windl, Omer Akgul, Nathan Malkin, and Lorrie Faith Cranor. Privacy Solution or Menace? Investigating Perceptions of Radio-Frequency Sensing. In *34th USENIX Security Symposium*, 2025.

[53] Shikun Zhang, Yuanyuan Feng, Lujo Bauer, Lorrie Cranor, Anupam Das, and Norman Sadeh. "Did you know this camera tracks your mood?": Understanding Privacy Expectations and Preferences in the Age of Video Analytics. *Proceedings on Privacy Enhancing Technologies*, 2021.

[54] Shikun Zhang, Lily Klucinec, Kyerra Norton, Norman Sadeh, and Lorrie Faith Cranor. Exploring Expandable-Grid Designs to Make iOS App Privacy Labels More Usable. In *20th Symposium on Usable Privacy and Security*, 2024.

[55] Sebastian Zimmeck, Rafael Goldstein, and David Baraka. PrivacyFlash Pro: Automating Privacy Policy Generation for Mobile Apps. In *NDSS*, volume 2, page 4, 2021.

[56] Sebastian Zimmeck, Peter Story, Daniel Smullen, Abhilasha Ravichander, Ziqi Wang, Joel Reidenberg, N Cameron Russell, and Norman Sadeh. MAPS: Scaling Privacy Compliance Analysis to a Million Apps. *Proceedings on Privacy Enhancing Technologies*, 2019.

A Survey Questionnaire

(All variables in the survey are indicated within angular brackets, e.g., <name>.)

This study consists of four stages:
- A short questionnaire on your perception of user privacy

feedback.
- An overview of the privacy feedback your application has been getting.
- Your reactions to this feedback.
- Demographics and professional experience questions.

We will occasionally show our findings of privacy-relevant reviews of your application. This feedback is real.

If you complete the survey successfully you will see an option to download the full report for your application.

Q1. How do you currently assess your users' perceptions of privacy within your application (<package_name>)?

Q2. Based on your current understanding of users' privacy perceptions within your application, how concerned or not concerned are you? *There are no right or wrong answers, we're only interested in what you think.*
- Not at all concerned
- Slightly concerned
- Somewhat concerned
- Very concerned
- Extremely concerned

Q3. Consider your current understanding of how users perceive the privacy of your application. How likely are you to take the following actions in response to user privacy perceptions? (5-point Likert scale of 'Not at all likely' to 'Extremely likely').
- How likely are you to change how a feature is implemented?
- How likely are you to add a feature to your application?
- How likely are you to remove a feature from your application?
- How likely are you to limit data collection in your application?
- How likely are you to reorganize the UI of your application (e.g., make privacy settings more accessible)?
- How likely are you to communicate your privacy decisions through the application (e.g., a warning, explanations in privacy settings)?
- How likely are you to communicate your privacy decisions outside of the application (e.g., the app description page, a mailing list, your website)?

Q4. We developed methods to detect and analyze privacy-relevant feedback from reviews left on Google Play. We found that users of your application (<package_name>) have left privacy-relevant reviews summarized under the following themes. How concerned are you with each theme? *Please hover over each theme for a quick summary.* (5-point Likert scale of 'Not at all concerned' to 'Extremely concerned').
- Theme1 title (summary displayed on hover)
- Theme2 title (summary displayed on hover)
- Theme3 title (summary displayed on hover)

Q5. Why were you '<concern_level>' with '<theme>'? *You may hover over the theme for a summary.* (We prefer themes where the participant indicated an intermediary concern level over themes with 'Not at all concerned' or 'Extremely concerned' responses. This same theme is chosen as the <selected_theme> for the rest of the survey.)

Next we'll show you a privacy report based on '<selected_theme>' reviews for your application (<package_name>). The rest of this study will focus on '<selected_theme>'. Please take a couple minutes to read each one of our observations.

Theme Overview
- Theme title
- Theme summary
- Sample Reviews: (quote1, quote2)

Volume & trends
Your application (<package_name>) has received N reviews about '<selected_theme>'. That's <percentage>% of the reviews about privacy. These type of reviews have <increased/decreased> by <change>% in the last year (compared to the amount received in the previous year) as opposed to <all_reviews_change>% in overall reviews.

Peer comparison
Adjusting for the volume of reviews each application receives, we see the following: Your application receives more '<selected_theme>' reviews than <percentile>% of similar applications (i.e., your competitors). Your application, on average, receives <relative_percentage>% reviews about '<selected_theme>' compared to applications similar to yours.

Emotions
<percentage>% of your '<selected_theme>' reviews contain negative emotions (anger, sadness, disgust, and confusion). The breakdown of all emotions associated with '<selected_theme>' reviews are as follows: <pie-chart>.

Now that we've shown you the privacy report. We'd like to ask you a couple questions about the report.

Q6. Please rate how useful or not useful you think each of the items are when understanding your privacy feedback. The first four items are from the report about '<selected_theme>'. The last item is how you approached understanding privacy feedback before this study. *Hover over each item to view the observation.* (5-point Likert scale of 'Not at all useful' to 'Extremely useful').
- Theme overview
- Volume & Trends

- Peer comparison
- Emotions
- Your previous approach

Q7. Why did you respond with '<usefulness _value>' to '<item>'? *You may hover over the observation to see details.* (We randomly select the observation in the privacy report that received the lowest usefulness score.)

Q8. Would you want access to privacy reports similar to the one shown in this study (hover here for a quick summary)? *Please select all that apply.*
- I would like to be able lookup the report on demand.
- I would like to receive report when there is a sudden change in privacy reviews.
- I would like weekly reports.
- I would like monthly reports.
- I would like annual reports.
- I never want to see such reports.

Q9. Now that you have seen the privacy report on '<selected_theme>' (hover here for a quick summary), consider if you are willing to take action based on what you have learned. Compared to your prior understanding of how users perceive the privacy of your application. How much more or less likely are you to take the following actions? (5-point Likert scale of 'Much less likely' to 'Much more likely').
- How much more or less likely are you to change how a feature is implemented?
- How much more or less likely are you to add a feature to your application?
- How much more or less likely are you to remove a feature from your application?
- How much more or less likely are you to limit data collection in your application?
- How much more or less likely are you to reorganize the UI of your application (e.g., make privacy settings more accessible)?
- How much more or less likely are you to communicate your privacy decisions through the application (e.g., a warning, explanations in privacy settings)?
- How much more or less likely are you to communicate your privacy decisions outside of the application (e.g., the app description page, a mailing list, your website)?

Q10. Please rate how much each of the observations influenced or didn't influence your likelihood of taking action. The first four items are from the report about '<selected_theme>'. The last item is how you approached understanding privacy feedback before this study. *Hover over each item to view the observation.* (5-point Likert scale of 'No influence at all' to 'Influenced very much').
- Theme overview
- Volume & Trends
- Peer comparison
- Emotions
- Your previous approach

Q11. Conditional We've noticed that you responded with "much less likely," "less likely," or "neither more or less likely" to all previous actions. Why are you "much less likely," "less likely," or "neither more or less likely" to take the actions listed (hover for list)?

Q12. Conditional What would need to be different with the privacy report (hover here for a quick summary) for you to be more likely to take action? (Shown if Q11 is shown.)

Q13. Conditional Why were you more likely to '<action>' compared to other options (hover for all options)? (Shown if the participant indicated more likelihood to take a certain action over others.)

Q14. Conditional Why were you '<likelihood_response>' to '<action>'? (Shown if Q11, Q12, and Q13 are not shown)

Demographic and professional experience questions

Q15. For how many years have you worked on mobile applications?

Q16. For how many years have you worked on the application we focused on in this study (<package_name>) specifically?

Q17. What role(s) have you served in the development of <package_name>? *Please select all that apply.*
- Software engineer
- Technical manager
- Project manager
- UI/UX designer/researcher
- Quality assurance
- Marketing
- Other

Q18. Did you ever have a role that affects the privacy-relevant aspects of the application (<package_name)? *Please select all that apply.*
- I'm the main privacy decision maker
- I help make privacy-relevant decisions.
- I implement privacy relevant features.
- I have another role that impacts privacy-relevant aspects.
- I wish to not answer.

Q19. How does/did your role impact privacy-relevant aspects?

Q20. Does the company that develops the application (<package_name>) have a system for addressing users' privacy relevant feedback?

- Yes
- No
- I wish to not answer

Q21. Approximately how many employees in your organization are primarily working on the application (<package_name>)? This may include employees that do development, research, maintenance, support, leadership, legal work, and marketing.
- Range from 1 to >10,000, including 'I wish to not answer'.

Q22. Approximately how many employees are in your organization in total? This includes all employees.
- Range from 1 to >10,000, including 'I wish to not answer'.

B Survey Flow

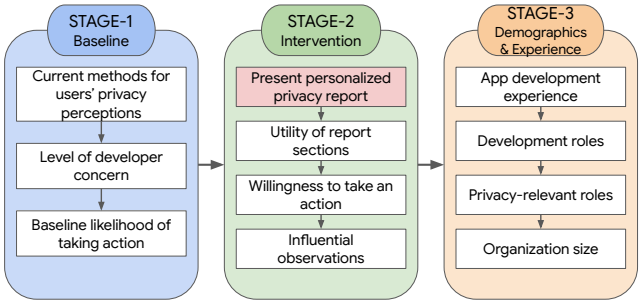


Figure 7: A flow diagram showing the three stages and the different categories of questions in the survey.

C Sample Privacy Report for a Gaming Application

Top Privacy Themes

- **Data Theft:** Users complain about the app's data security and privacy policy. Many users reported their data being leaked or misused after using the app.
- **Account Hacking:** Users reported issues such as account hacked, data breach, privacy concerns, payment issues, etc.
- **Data Deletion:** Users are frustrated with the app's data deletion policies. They complain that they are unable to delete their data, that their data is still being used even after they have deleted it, and that they are not given enough information about how their data is being used.

Selected Theme Overview: Data Deletion
Representative Reviews:

- "I don't just want to log out and delete the app, I want to permanently delete all my data and information so that you cannot find it. Tell me where I can delete all my data!"
- "Can you PLEASE HELP ME FIGURE out how to delete my 8 ball account! I know how to uninstall the app but I want to actually delete my account..."

Volume & Trends:

- Volume: 84
- Percentage of privacy reviews: 3.48%
- Growth/Decline in the last year: -31.58%

Peer Benchmark Metrics:

- Percentile ranking: 84%
- Relative magnitude to average: 163.2%

Emotion Distribution:

- Neutral (46.34%)
- Confusion (30.49%)
- Anger (12.20%)
- Joy (8.54%)
- Desire (2.44%)