An empirical study of privacy labels on the Apple iOS mobile app store

Gian Luca Scoccia
DISIM, University of L'Aquila
L'Aquila, Italy
gianluca.scoccia@univaq.it

Marco Autili
DISIM, University of L'Aquila
L'Aquila, Italy
marco.autili@univaq.it

Giovanni Stilo
DISIM, University of L'Aquila
L'Aquila, Italy
giovanni.stilo@univaq.it

Paola Inverardi
DISIM, University of L'Aquila
L'Aquila, Italy
paola.inverardi@univaq.it

ABSTRACT

Privacy labels provide an easy and recognizable overview of data collection practices adopted by mobile apps developers. Specifically, on the Apple App Store, privacy labels are displayed on each mobile app's page and summarize what data is collected by the app, how it is used, and for what purposes it is needed. Starting from the release of iOS version 14.3 developers are required to provide privacy labels for their applications. We conducted a large-scale empirical study, collecting and analyzing the privacy labels of 17, 312 apps published on the App Store, to understand and characterize how sensitive data is collected and shared. The results of our analysis highlight important criticalities about the collection and sharing of personal data for tracking purposes. In particular, on average free applications collect more sensitive data, the majority of data is collected in an unanonimyzed form, and a wide range of sensitive information are collected for tracking purposes. The analysis provides also evidence to support the decision-making of users, platform maintainers, and regulators. Furthermore, we repeated the data collection and analysis after seven months, following the introduction of additional run-time tracking controls by Apple. Comparing the two datasets, we observed that the newly introduced measures resulted in a statistically significant decrease in the number of apps that collect data for tracking purposes. At the same time, we observed a growth in overall data collection.

KEYWORDS

iOS, Apps, Privacy

ACM Reference Format:
Gian Luca Scoccia, Marco Autili, Giovanni Stilo, and Paola Inverardi. 2022. An empirical study of privacy labels on the Apple iOS mobile app store. In *IEEE/ACM 9th International Conference on Mobile Software Engineering and Systems (MOBILESoft '22), May 17–24, 2022, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3524613.3527813

MOBILESoft '22, May 17–24, 2022, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9301-0/22/05…$15.00
https://doi.org/10.1145/3524613.3527813

1 INTRODUCTION

In conjunction with the release of iOS version 14.3, made available on the 14th of December 2020, Apple introduced *privacy labels* on the Apple App Store. Privacy labels are published on each mobile app's page and provide an easy and recognizable overview of data collection practices employed by the app, displaying what data is collected, how it is used, and for what purposes it is needed [27]. Developers are required to provide privacy labels for their applications, prior to submitting new apps or app updates to the App Store. An example of an app's privacy labels is shown in Figure 1.

Specifically, privacy labels display previously unavailable information about the kind of data collected by the application (e.g., location, contact info) and the form in which it is collected (i.e., anonymized or with references to the user identity). By opening a detail window, the purpose (e.g., analytics, third-party advertising) for which the information is collected is also visible. Notably, privacy labels place special emphasis on data used for tracking purposes, i.e., data shared with other companies for targeted advertising or advertising measurement purposes [29], displaying them separately from other collected data.

With the release of iOS 14.5 on the 26th of April 2021, Apple introduced further privacy-oriented changes to the iOS platform. After this release, app developers are required to explicitly request permission to track the user beyond the app in use [29]. Preliminary reports highlight how this need for authorization has been well received by users [4]. However, at the time of writing, there is a lack of evidence on how it impacted applications' data collection practices and on how developers have adapted to it.

While the research community has widely investigated the data collection practices of Android applications [40, 41, 54], less is known about the practices of iOS apps. To fill this gap, leveraging the information made newly available by privacy labels, we conducted a large-scale empirical study to understand and characterize how sensitive data is collected and shared by iOS applications. For this purpose, we collected the privacy labels of 17, 312 popular apps published across the 24 categories of the Apple App Store. From the privacy label of each application, we extracted information about the sensitive data used by the app, in which form it is collected, and the purpose of usage. We conducted a second data collection round after seven months, to assess how developers' data collection practices have changed after the introduction of run-time tracking controls.

Our analysis highlights important results about the collection and sharing of personal data for tracking purposes by App Store apps. Specifically, we found that on average free applications collect and share more sensitive data, with 48.79% of the analyzed free applications that collect at least one piece of data for tracking purposes. We found that the majority of data is collected in an unanonimyzed form, exposing end-users to more potential privacy risks [5, 24, 54]. A wide range of sensitive information are collected for user tracking, including user location and data created for diagnostic purposes (e.g., crash logs). Furthermore, regulations that restrict data collection for especially sensitive app categories appear to be effective, as app categories subject to stricter rules do exhibit a lower amount of tracking. Finally, the additional run-time tracking controls introduced by Apple appear to have discouraged aggressive data sharing practices by app developers, as we observed a statistically significant decrease in the number of apps that collect data for tracking purposes. However, at the same time, we observed a growth in overall data collection.

The target audience of our study is composed of end-users, mobile platform maintainers, and regulators. Specifically, we inform the more privacy conscious end-users by providing them with an overview on the state of sensitive data collection by iOS apps and by presenting them with recommendations that can be of guidance while choosing the apps to install on their devices. Platform maintainers can take advantage of our analysis as we supply them with evidence on the impact of the measures introduced by Apple, that can assist them in their future decision-making. Similarly, we provide regulators with evidence on the effectiveness of current privacy regulations.

To allow for independent verification and replication of the performed study, we make publicly available a replication package containing the collected data and all the code developed for data preparation and analysis[1].

The remainder of this paper is organized as follows: Section 2 provides detailed backround information about privacy labels, whereas Section 3 illustrates related work. Section 4 describes the design of our study. Section 5 presents the results of our analyses, which are are discussed in Section 6. Section 7 discusses the threats to the validity of our study. Lastly, Section 8 closes the paper.

2 BACKGROUND

Privacy labels [27] (also referred to as *privacy details*) assist users in understanding app's privacy practices prior to installing the app on the user's device. Specifically, privacy labels are displayed on each app's page on the Apple App Store and summarize what data is collected by the app, how it is used, and for what purposes it is needed. Privacy labels have been introduced on the Apple App Store on the 14th of December 2020, in conjunction with the release of iOS version 14.3. Currently, developers are required to provide information about the app's sensitive data handling practices before submitting new apps or app updates to the App Store. Noticeably, privacy labels have been introduced for all software products present on the App Store (including desktop software) but, in this paper, we will focus exclusively on mobile apps.

[1] https://bit.ly/3i1OEtN

An example of an app's privacy labels visualization is shown in Figure 1. Apple defines fourteen data types, each identified by a unique name and an icon. Developers can use data types to summarize the app's sensitive data-handling practices from a privacy perspective. Each data type groups together one or more related data, e.g.,, the *User ID* and *Device ID* data belong to the *Identifiers* data type. The data handled by the app are organized in the privacy labels page's section into three categories. Each category is displayed in a distinct card according to the way it is used: e.g., *"data used to track you"*, *"data linked to you"* and *"data not linked to you"*. Hereafter, for brevity, we will refer to these categories as *tracking data*, *linked data* and, *unlinked data*, respectively.

Tracking data refers to data used for tracking purposes. According to the official Apple website, tracking is defined as *"the act of linking user or device data collected from your app with user or device data collected from other companies [. . .] for targeted advertising or advertising measurement purposes."* [29]. Linked data refers to data that contains references to the user's identity, i.e., information that includes references to the user's account, device, or details (e.g., the user's phone number). Unlinked data refers to data that does not contain such references or has been stripped of them, and thus the data is fully anonymized.

In addition to this grouping of data types, additional information about the app's privacy practices can be accessed by opening a "details" pop-up. Here, one of six possible purposes is listed for each collected data: *"Analytics"*, *"App Functionality"*, *"Product Personalization"*, *"Third-Party Advertising"*, *"Developer's Advertising or Marketing"*, and *"Other Purposes"*. Moreover, for each collected data type the actual datums collected are listed, e.g., the email address for the *Contact Info* data type.

With the release of iOS 14.5 on the 26th of April 2021, app developers explicitly need to ask for the users' permission before tracking them or accessing their device's advertising identifier. For that purpose, developers use the ad-hoc `appTrackingTransparency` framework [29] for notifying the users at run-time when a tracking request is performed. In turn, users can grant or deny the permission to be tracked through a pop-up message. Notes that, starting with iOS 14.5, the usage of data for tracking purposes is no longer solely reflected in the displayed privacy labels of the application's store page, but it also has a direct impact on the application usage, hence potentially affecting negatively the user experience.

3 RELATED WORK

In this section, we discuss work related to our study by covering three main topics, i.e., literature about privacy labels, mobile permission systems, and mobile app privacy.

3.1 Privacy labels

The concept of "privacy labels" has been originally introduced by Kelley et al. [31] that, drawing from nutrition and energy labels, designed a privacy-centered label that presents the ways organizations collect, use, and share personal information. Specifically, the proposed label utilizes a grid structure to display the types of information collected (e.g., contact information, cookies), how the information is used (e.g., marketing, profiling), and with whom the

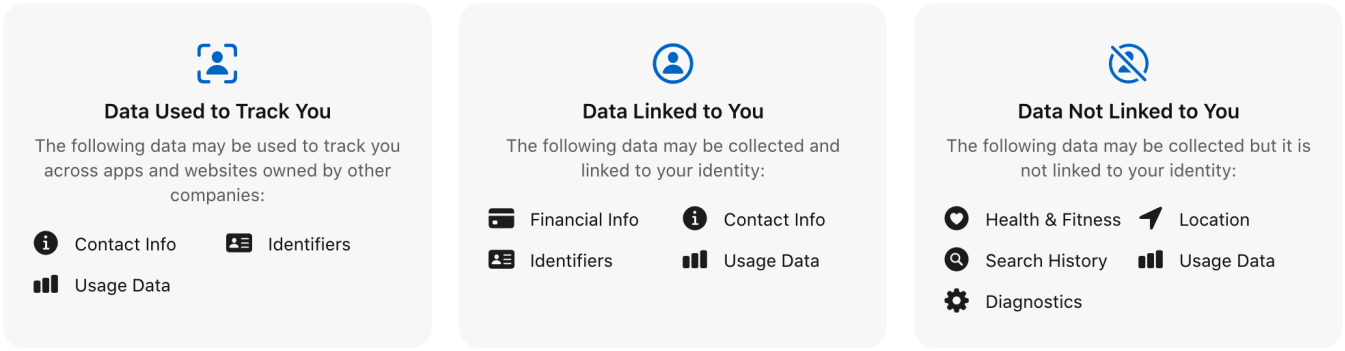

Figure 1: An example of Apple's App Store privacy labels

information is shared (e.g., other companies). More recently, similar labels have been proposed for Internet of things devices [17, 47].

To our knowledge, the use of privacy labels in the Apple App Store represents their first large-scale application in the mobile domain. However, differences do exist between the original proposal of Kelley et al. [31] and their usage in the App Store. For instance, the former employs colors on a scale from light to dark to highlight the severity of certain privacy practices and makes use of different symbols to convey different data collection options available (e.g., an `OPT` symbols signifies that the user can opt-out from the data collection). Both these peculiarities are absent from their implementation on the App Store. Apple's implementation of privacy labels has attracted considerable interest from journalists and specialized media [11, 52]. However, our study is the first research work focusing on privacy labels on the Apple App Store, leveraging the information contained in them to perform a large-scale analysis of data collection practices of iOS applications.

3.2 Mobile app privacy

The amount of sensitive data handled by mobile apps during their operation exposes users to severe risks for their privacy. Multiple studies have focused on understanding how mobile apps treat users' privacy and the factors that affect data collection practices. Razaghpanah and colleagues [40] performed a large scale study aiming at understanding, on the one hand, the actors involved in the mobile advertising and tracking ecosystem and, on the other hand, the commonly employed data flow. Their results show that these services can track users leveraging a wide range of device identifiers without providing any visual clues to the users inside the apps. They analyze network traffic showing that once the users' data is gained, sharing it with subsidiaries and third-party affiliates is the norm. Grace et al. [25] investigated potential privacy and security risks posed by embedded advertisement libraries commonly used in commercial Android apps. Their analysis highlights that these libraries often engage in risky behaviour for end-users privacy, ranging from uploading sensitive information to remote servers to executing untrusted code downloaded from Internet sources. Ren and colleagues [43] analyze how mobile privacy is evolving over time and whether it is getting better or worse. To do so, they study the privacy leaks from historical and current versions of 512 popular Android apps, covering app releases over eight years. Analyzing the network traffic generated by the apps, they found that collecting personally identifiable information has increased over time, involving a large number of third parties that can link user activities and locations across apps.

More closely related to our work, Kollnig et al. conducted a comparative study of the Google Play store and the Apple App store considering several dimensions pertaining to user privacy [33]. Their findings suggest that third-party tracking and the sharing of unique user identifiers are widespread in apps from both ecosystems. Moreover, they identified diffused potential violations of current privacy laws, as personal information is often collected without prior user consent. Our analysis is complementary to theirs, as by analyzing privacy labels we can investigate dimensions not considered in their study, e.g., purpose of data collection and the impact of in-app notifications.

3.3 Mobile permission systems

Current mobile operating systems primarily rely on *permission systems* to inform users when sensitive data are accessed and to regulate the access to sensible APIs of the platform [42]. Researchers have mostly focused on the Android permission system, uncovering issues on a variety of aspects, including user's comprehension [3, 20, 22, 32, 46], usability [35, 39, 44], misuses [30, 45, 49], and exploits [7, 9, 18, 41].

Fewer works have focused on the iOS platform. Agarwal et al. [1] propose *ProtectMyPrivacy*, a privacy-protecting architecture for iOS that notifies users when sensitive data are accessed by apps and provides a mechanism to either grant or deny such access. Anonymized data are used in place of sensitive data when the access is denied. Analyzing the data collected by the applications, they found that applications frequently access a variety of sensitive data, including almost half of the analyzed applications accessing the unique device identifier. Han and colleagues [26] examine applications that run on both Android and iOS to assess the difference in the usage of security-sensitive APIs. They found that iOS applications consistently access more sensitive APIs than their Android counterparts, likely due to the fact that accessing these APIs on the iOS platform is invisible to the user. Deng et al. [15] developed Iris, an hybrid analysis approach to detect unauthorized use of sensitive APIs in iOS applications. Using Iris, they found 146 applications

that stealthily access sensitive user information, violating Apple's terms of service.

Permission systems share some similarities with privacy labels and the app tracking controls, as these also inform users about access to sensitive information prior to app installation or during its execution [22, 39, 44]. However, the scope of permission systems is broader and, in addition to their informative role, they include access-control capabilities. As a consequence, an application can disable some or its entire functionality should access to some restricted part of the platform be denied [46]. On the contrary, withholding functionalities after permission to be tracked is explicitly disallowed for the `appTrackingTransparency` framework [29].

4 STUDY DESIGN

This section describes how we designed our study. In order to perform an objective study we followed the guidelines on empirical software engineering outlined in [48] and [53].

4.1 Goal and research questions

Intuitively, the *goal* of our study is to characterize how sensitive data are used by iOS apps on the Apple App store and derive insights that can better drive the choices of end-users, app store managers, and regulators. We also aim to measure the effects of recently introduced run-time privacy protections as a more fundamental contribution. This latter contribution can guide Apple, and other mobile platforms maintainers, in their decision-making. The *context* of this study is composed of 16, 425 free and 887 paid apps published in the Apple App Store. Table 1 provides a definition of the goal using the Goal-Question-Metric technique [8, 10].

Table 1: Goal definition

Object of study	Privacy labels
Purpose	Characterize
Quality focus	Sensitive data usage
Perspective	End-users, app store managers, and regulators
Context	Real-world iOS apps

To achieve this goal, we aim to provide an answer to the following research questions:

- **RQ1** Which forms of sensitive data are more frequently collected by App Store apps?
- **RQ2** Which is the usage distribution of sensitive data across App Store categories?
- **RQ3** Which sensitive data types are more frequently collected by App Store apps?
- **RQ4** Which intents more frequently characterise collecting sensitive data?
- **RQ5** What was the impact of having introduced in-app notifications on collecting tracking data?

RQ1 measures in which form (linked or unlinked) sensitive data is more frequently collected, and whether that data is used to track end users. RQ2 instead point attention to the distribution of sensitive data across different app categories. RQ3 focuses on data types (e.g., *Location*, *Contact Info*) and establishes their usage frequency, whereas RQ4 investigates the intent for which sensitive data are collected. The rationale of RQ5 is twofold: on the one hand, it provides an updated snapshot of sensitive data usage by the apps on the App Store; on the other hand, it aims to estimate the impact that run-time notifications can have on developers' data collection practices. Indeed, if from one side users might become more aware of tracking data collected by applications, on the other side, run-time popups have been found to be distracting and attention-consuming [2, 3, 6, 37]. Hence, app developers might wish to minimize the usage of tracking data accordingly. Moreover, as run-time pop-ups are increasingly used by mobile and web platforms [21, 38, 42], understanding how developers are adapting is a timely topic.

4.2 Hypotheses

To answer RQ5, we want to assess whether the introduction of run-time popup messages (that notify users of tracking requests) has caused a reduction in the use of personal data by the App Store apps. To this end, we formulate the null and alternative hypotheses as follows.

Let S and S' be snapshots of the applications' privacy labels collected before and after the introduction of in-app notifications, respectively. Let $count_t(a)$ be the total number of *tracking* data used by an app $a \in S$, and let $count'_t(a)$ be its analog in S'. The null and alternative hypotheses are defined as, respectively:

$$H_0^t : count'_t(a) = count_t(a) \qquad H_1^t : count'_t(a) < count_t(a) \tag{1}$$

The null hypothesis H_0^t states that distribution of the usage of tracking data does not differ significantly between the two snapshots. Instead, H_1^t states that a statistically significant reduction in the usage is observed between the two snapshots. Similarly, the null and alternative hypotheses for *linked* and *unlinked* data can be defined as, respectively:

$$H_0^l : count'_l(a) = count_l(a) \qquad H_1^l : count'_l(a) < count_l(a) \tag{2}$$

$$H_0^u : count'_u(a) = count_u(a) \qquad H_1^u : count'_u(a) < count_u(a) \tag{3}$$

To verify these hyphotheses, we rely on the metrics $count_t$, $count_l$, and $count_u$, computed for all applications at each snapshot.

4.3 Data collection

Our data collection process is summarized in Figure 2. We started our data collection on the 5th of March 2021. At the time, about two months had passed after the introduction of privacy labels (as previously discussed in Section 2). We therefore expected a considerable amount of apps having been updated to provide privacy labels on their App Store page. As the first step for our data collection, we considered (i) the list of the top one thousand free apps, and (ii) the list of the top one thousand top-grossing apps in the United States market for each of the 24 categories of the Apple App Store (according to the AppMagic[2] mobile market intelligence tool). Due to the fact that some categories contain less than one thousand entries, after merging the two lists and removing duplicates (as apps can appear in both lists and in more than one category), we identified

[2] https://appmagic.rocks/

30,753 distinct applications. For each application, we collected its app store page details employing an open-source tool[3]. From this initial app set A, we further discarded apps that had not received an update after the 15th of December 2020, since only those apps updated after this date are required to provide privacy labels on their store page. A total of 16,425 free and 887 paid apps survived this filtering step, and they constitute our initial data snapshot S. For each app in S, we collected the privacy labels and its sensitive data usage details using a custom-built tool. By parsing the privacy labels of each application, we were able to extract: *(i)* the sensitive data used by the app (e.g., location, contact info); *(ii)* the way it is used (i.e., linked, unlinked or tracking); *(iii)* the purpose of usage (e.g., analytics, third-party advertising).

To collect an updated snapshot of sensitive data usage, we repeated the collection of the privacy labels on the 11th of November 2021, employing the original app set A. By that time, more than six months had passed since the introduction of the App Tracking Transparency framework, and we consider this an ample time for developers to adapt their apps' sensitive data usage practices in response to the newly introduced run-time privacy controls (refer to Section 2). Of the apps in the original list, 15,617 free and 853 paid apps survived all the data collection and filtering steps. The reduced number was mostly due to apps that were removed from the app store (this can happen if the app infringes some app store rules or if the developer decides to pull the app from the market) or geographically restricted (thus making their data unattainable). The privacy labels collected during this second round of data collection constitutes our updated snapshot S'. The collected snapshots S and S' are made publicly available in the online replication package[4].

4.4 Analysis

To provide an answer to RQ1, we resort to descriptive statistics of the usage rates of tracking, linked and unlinked data of the apps in S. Similarly, to answer RQ2, RQ3, and RQ4 we rely on counts, usage rates, and visualizations of the collected data, to assess potential differences across app categories, data types, and purposes, respectively.

[3]https://github.com/facundoolano/app-store-scraper
[4]https://bit.ly/3i1OEtN

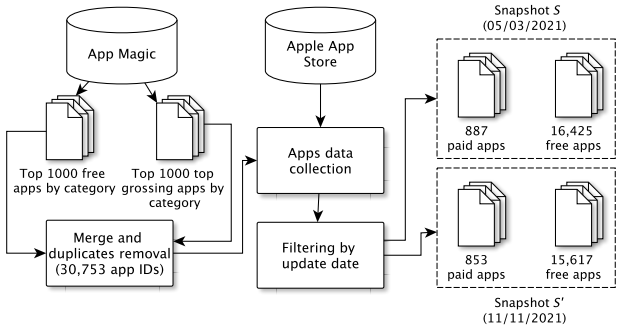

Figure 2: Data collection process

Table 2: Descriptive statistics of sensitive data usages for dataset S (# = count, μ = mean, $\mu_{\frac{1}{2}}$ = median, σ = standard deviation)

	Free apps				Paid apps			
	#	$\mu_{\frac{1}{2}}$	μ	σ	#	$\mu_{\frac{1}{2}}$	μ	σ
Tracking	22,277	0	1.36	1.83	175	0	0.2	0.69
Linked	44,625	2	2.72	2.86	380	0	0.43	1.33
Unlinked	30,594	1	1.86	1.82	798	0	0.9	1.38
Total	**97,496**	**6**	**5.94**	**4.02**	**1,353**	**0**	**1.53**	**2.44**

To answer RQ5, in addition to the visualizations of the collected data, we employ hypothesis testing [34]. For testing our hypotheses, defined in Section 4.2, we adopt the Wilcoxon signed-rank test [12], a non-parametric test used to determine whether two related paired samples come from the same distribution. We selected this test as it does not assume the normality of the data being tested, an assumption that we know does not hold for our case as we are dealing with discrete data. As the test requires paired data, prior to testing our hypotheses, we filter out those apps that appear in S but not in S'.

For assessing the magnitude of potential differences found for each of our hypotheses, we calculate effect size using the r measure based on the Z-score statistic [50], according to which we have a small effect size for $0.1 \leq r < 0.3$, a medium one for values $0.3 \leq r < 0.5$, and a large effect sizes for $r \geq 0.5$.

5 RESULTS

This section presents the results of our analysis, organized accordingly to the related research question.

5.1 RQ1: Which forms of sensitive data are more frequently collected by App Store apps?

Following the analysis of the privacy labels for apps in S, we found that among free applications 8,013 (48.79%) are using tracking data, 10,010 (60.94%) are using linked data, and 11,164 (67.97%) are using unlinked data. Merging the three aforementioned categories, we obtain that 14,481 (88.16%) free applications are using data of any kind. A considerably lower frequency of sensitive data usage is observed for paid applications: 93 (10.48%) are using tracking data, 120 (13.53%) are using linked data, and 352 (39.68%) are using unlinked data. Overall, 388 (43.74%) paid applications are using sensitive data of any kind.

Descriptive statistics of sensitive data usages for the apps in S are presented in Table 2. Differences in frequency of usage are noticeable, with a total of 97,496 and 1,353 sensitive data usages observed for free and paid applications, respectively. This difference is reflected in the descriptive statistics: free applications have a median value of 6 usages of sensitive data, with a mean of 5.94 and 4.02 standard deviation; whereas, paid apps exhibit lower values with a value of 0 for the median, 1.53 for the mean, and 2.44 for the standard deviation.

Focusing on the most used collection type, we notice a further difference: linked data is the most commonly used data type for free applications with 44, 625 total and 1.86 mean usages per app, while paid apps rely more often on unlinked data with 798 and 0.9 mean usages per app.

5.2 RQ2: Which is the usage distribution of sensitive data across App Store categories?

The barplot in Figure 3 summarizes data usages by App Store category for apps in *S*. For each category, the percentage of apps employing tracking, linked, and unlinked data is reported. In the case of apps that are listed in multiple categories, apps were assigned to the developer-defined primary category.

The top three categories for usage of tracking data are *Games*, *News*, and *Magazines & newspapers*. Usage of tracking data appears to be especially common in these categories, with a percentage of usage of 88.6%, 72.34%, 68.17%, respectively. Together with categories *Shopping*, *Entertainment*, and *Weather*, these constitute the only categories for which tracking data is employed by more than 50% of apps. On the opposite side, the categories for which usage of tracking data is less common are *Medical* (25.8%), *Business* (20.88%), and *Kids* (2.92%).

Focusing on linked data, high percentages of usage can be noted for categories *Shopping* (80.8%), *Magazines & newspapers* (80.56%), *Travel* (74.25%), *Food & drink* (72.91%), and *Business* (71.58%), all exhibiting with over seventy percent of apps that employ linked data. Categories with an especially low percentage of applications that rely on linked data are *Utilities* (37.14%), *Photo & video* (37.09%), *Kids* (36.04%), *Reference* (33.75%), and *Weather* (33.54%), all with a percentage of less than forty percent.

Regarding unlinked data, categories that exhibit an high percentage of usage are *News* (78.22%), *Photo & video* (75.58%), *Magazines & newspapers* (74.37%), *Shopping* (73.08%), *Weather* (72.47%), and *Utilities* (71.13%), all with over seventy percent of apps employing unlinked data. Categories that exhibit the lower usage rates are *Finance* (59, 96%), *Kids* (54, 87%), and *Business* (53.86%).

5.3 RQ3: Which sensitive data types are more frequently collected by App Store apps?

The heatmaps of Figure 5 summarizes the percentages of usage for each data type by App Store category for applications in *S*. The three heatmaps reports percentages of usage for tracking, linked, and unlinked data, respectively.

Focusing on the data used for tracking (heatmap A in Figure 5), we can immediately notice that the most frequently used data types are *Usage Data* and *Identifiers*, being used in 24.4% and 30.1% of all data usages for tracking purposes, respectively. This is unsurprising as these data types contain the user and device IDs, plus information about the advertisements the user has seen, hence we expect this data to be frequently used for tracking purposes. However four other data types are also frequently used: *Location* (10.3%), *Contact Info* (9.4%), *Diagnostics* (8.3%), and *Purchases* (8.5%). Moreover, even if less common, we can notice that all kinds of data are used for tracking purposes. This includes extremely sensitive data types and, in some cases, likely unrelated to the app purpose (e.g., *Health & Fitness* data used by apps in the *Weather* category, *Browsing History* data used by apps in the *Medical* category).

Examining individual categories in detail, we observe that the *Kids* category is the one with the highest usage of *Identifiers* for tracking. However, this is likely due to the low sample size, with only a total of 9 apps that perform tracking in the category. Other categories have considerably higher than average usage of some data types such as, for instance, *Location* for the *Weather* category (31.7% opposed to 10.3% average) and *Search History* for the *Shopping* category (5.9% opposed to 1.5% average).

Regarding linked data (heatmap B in Figure 5), we observe that among the most used data types the presence of *Usage Data* (14.1%) and *Identifiers* (18.9%), as observed for tracking. However, alongside these data types we can also observe higher usage percentages for

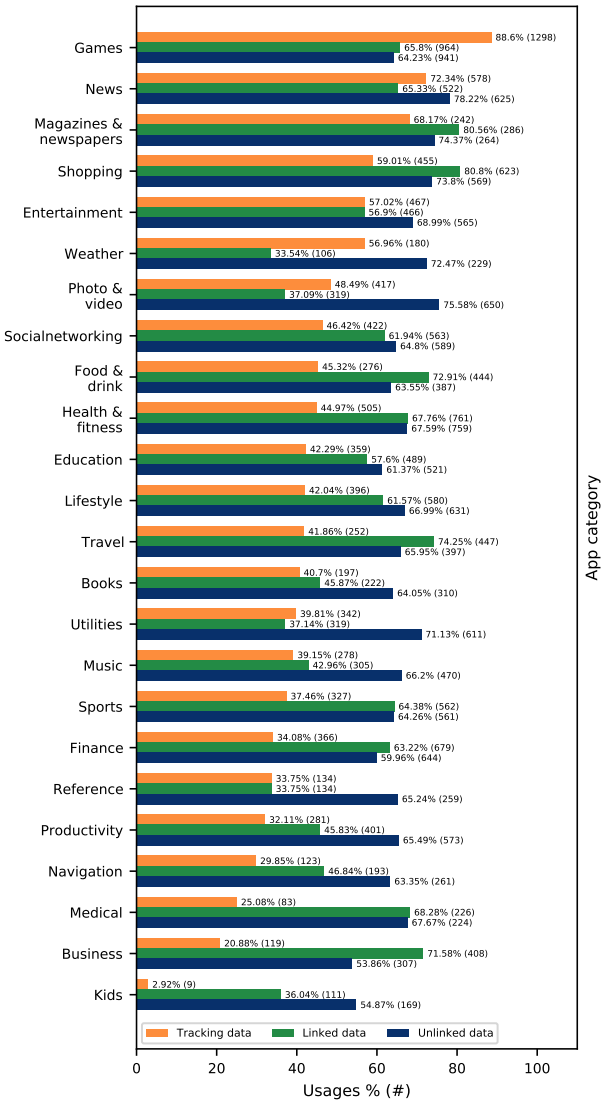

Figure 3: Data usages by App Store category

data types *User Content* (10.7%) and *Contact Info* (16.4%). Also in this case differences among categories can be observed for some data types, e.g., *Sensitive Info* are more used in categories *Medical*, *Socialnetworking*, *Lifestyle* and *Health & Fitness*.

Concerning unlinked data (heatmap C in Figure 5), we observe a low usage frequency for the majority of data types, with the exception of *Diagnostics* (29.9%), *Usage Data* (22.3%), and *Identifiers* (15.3%). Less common but still more frequently used are data types *Location* and *User Content*.

5.4 RQ4: Which intents more frequently characterise collecting sensitive data?

Table 3 summarizes the purposes declared by apps in S for their data collection. For each purpose, we report the number of apps in S that declare collecting linked and/or unlinked data for that purpose, as well as the related percentage. Tracking is not reported in the table as, according to the definition given by Apple and reported in Section 2, data collected that way is used exclusively for third-party advertising purposes.

We can observe that *App Functionality* is the most frequently declared goal for data collection for both free and paid applications, with 12,985 and 310 apps declaring it respectively. Analogously, *Analytics*, *Developer Advertising* and *Product Personalization* follow as popular goals for both types of applications, while *Other Purposes* is the least common one. However, a considerable difference can be observed between the percentages for free and paid applications: the former have considerably higher percentages of data usage, with some purposes being declared by almost 80% of apps, while the latter has a maximum usage percentage of 34.95%. Moreover, a substancial difference in the nature of used data can be observed: free applications rely more on linked data, with categories *App functionality* and *Analytics* having almost an even ratio between linked and unlinked data, while paid applications use more often unlinked data.

Figure 4 shows the percentage of apps declaring each purpose by App Store category. It is noteworthy that the *App Functionality* and *Analytics* purposes are commonly declared by apps in all categories, with at least 50% of apps in each category claiming to collect data for these purposes. On the other hand, we observe a less homogeneous distribution for purposes *Third-Party Advertising* and *Developer Advertising*. Both are in fact particularly common in the categories *Games* (80.4% and 65.5%, respectively), *News* (73% and 59.4%), *Magazines & Newspaper* (58.3% and 70.4%). On the contrary, they are particularly less common in categories *Business* (5.4% and 20.7%), and *Kids* (2.6% and 28.9%). Noteworthy is the the *Weather* category, the only category exhibiting an high rate for *Third-Party Advertising* (49.4%) and a considerably lower rate of *Developer Advertising* (27.2%). We observe considerable variations among categories also for the *Product Personalization* purpose, with maximum usage in categories *Shopping* (70.7%) and *Magazines & Newspaper* (70.1%), and minimum usage in categories *Reference* (18.4%) and *Utilities* (18.9%). Finally, *Other Purposes* is the least used purpose, and has minor variations across categories, with maximum in *Games* (27.6%) and minimum in *Navigation* (5.8%).

Table 3: Apps in S by purpose

	Free apps			Paid apps		
	Linked	Unlinked	Any	Linked	Unlinked	Any
App Functionality	9,421 (57.36%)	8,744 (53.24%)	12,985 (79.06%)	109 (12.29%)	264 (29.76%)	310 (34.95%)
Analytics	7,965 (48.49%)	8,930 (54.37%)	12,963 (78.92%)	63 (7.1%)	264 (29.76%)	265 (33.26%)
Product Personalization	5,598 (34.08%)	3,034 (18.47%)	7,312 (44.52%)	31 (3.49%)	26 (2.93%)	53 (5.98%)
Other Purposes	1,575 (9.59%)	1,077 (6.56%)	2,322 (14.14%)	15 (1.69%)	14 (1.58%)	28 (3.16%)
Developer Advertising	5,177 (31.52%)	2,512 (15.29%)	6,661 (40.55%)	40 (4.51%)	33 (3.72%)	60 (7.44%)
Third-party Advertising	3,392 (20.65%)	3,537 (21.53%)	5,980 (36.41%)	19 (2.14%)	27 (3.04%)	43 (4.85%)

5.5 RQ5: What was the impact of having introduced in-app notifications on collecting tracking data?

In the following, we highlight the principal differences observed in S' with respect to S.

Concerning the way in which sensitive data are collected, we observe a reduction in tracking data for both free and paid applications. For the former, we observe that 7,231 apps (46.3%) are using tracking data, down from the original 8,013 (48,79%). Regarding the latter, 53 apps (6.21%) employ tracking data against the original 93 (10.48%). However, this trend does not hold for the usage of linked and unlinked data. Indeed, among free apps, 9,977 (63.89%) and 11,008 (70.49%) are using linked and unlinked data respectively. This is an increase from the original percentages of 60.94% and 67.97%. Paid applications exhibit a different trend, with 108 (12.66%) using linked data and 350 (41.03%) using unlinked data, from the original 13.53% and 39.68%. Merging the three kinds of data usages, taking into account overlaps, we observe 14,241 (91.19%) free and 383 (44.9%) paid applications that employ sensitive data of any kind in S'.

These trends are confirmed by the statistical tests, that compare pairs of apps that are present both in S and S'. Indeed, we obtain a statistically significant result ($p\text{-}value < 0.01$) only for the test of Equation 1, concerning the usage of tracking data. This signifies that the introduction of the app tracking controls resulted in a statistically relevant reduction of tracking data usages. For it, the effect size calculation resulted in a value of $r = 0.133$, indicating a small effect size. Instead, for what concerns linked and unlinked data, the null hypothesis cannot be rejected, since, as further described in the following, for these kinds of data an increase is observed in place of the hypothesized reduction.

Exploring the distribution of data usages across categories, we observe differences in the magnitude of reduction of tracking data. Indeed, while for 18 out of the 24 categories we record a reduction, we observe that some experienced a significant decrease (e.g., −10.49% for *Health & Fitness*, −10.1% for *Business*, −7.7% for *Medical*) while other experienced a milder reduction (e.g., −2.93% for *Sports*, −2.47% for *Weather*). In contrast with the general trend, categories *Books* (+6.34%), *Games* (+5.10%) and *News* (+3.24%) record a significant increase. Focusing on usages of linked data, we observe a growing usage across most categories, with only categories *Weather* (−0.53%) and *Navigation* (−0.21%) recording a slight decrease. However, the

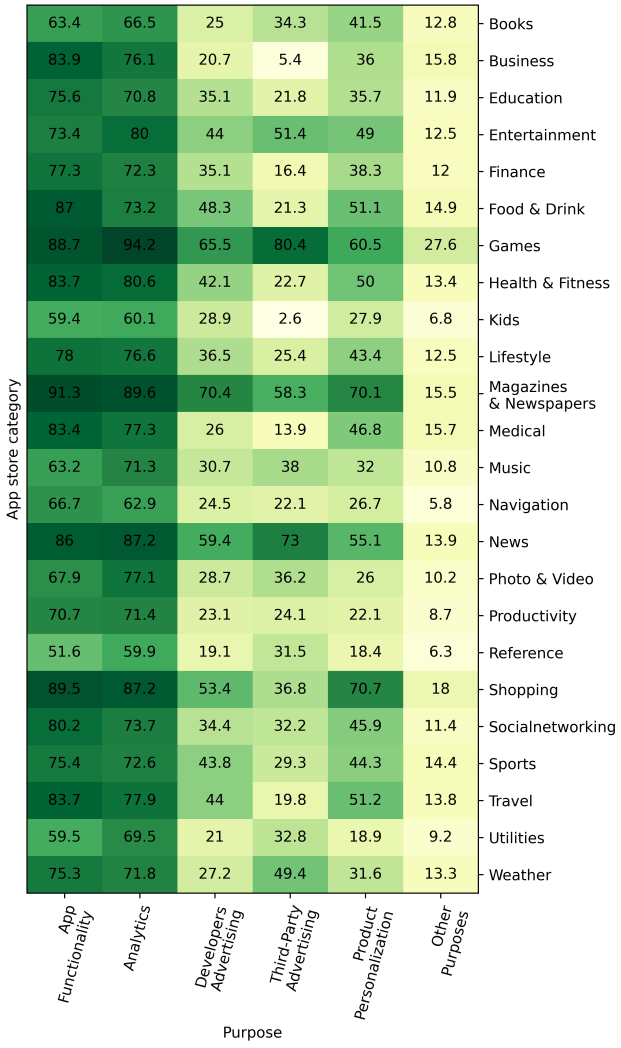

Figure 4: Percentage of apps declaring purpose by App Store category

increase is not enough to compensate for the reduction of tracking data, ranging from a maxium of 5.33% for category *Entertainment*, to a minium of 0.44% for category *Education*. Similarly, we observe an increase in the use of unlinked data, ranging from 5.15% for category *Sports* to 0.53% for category *Health & Fitness*. The only category recording a reduction in usage of unlinked data is the *Games* category (−0.32%).

Examing the differences in collected data types, we notice that all data types experienced a reduction in collection for tracking purposes, with *Contact Info* (−3.32%), *Identifiers* (−2.67%), and Usage Data (−2.19%) being the ones that experienced the greater reduction. However, when also considering linked and unlinked data, we find that for all data types there has been an increase in collection, with *Usage Data* (+2.79%), *Diagnostics* (+2.79%), *Location* (+2.46%), and *Identifiers* (+2.41%) having the highest increase.

Finally, comparing the declared purposes for the data collection between S and S', we observe a slight decrease in usage for almost all purposes, ranging from −2.12% of *App Functionality*, to −0.49% of *Product Personalization*. Noteworthy is the *Third-Party Advertising* purpose that, in contrast to the others, observes a slight increase (+0.43%).

6 DISCUSSION

In the following, we highlight and discuss the main implications of our findings.

App transparency – Apps collect and manipulate an increasing amount of personal data. Frequently, data is collected for reasons not directly tied to the user interest, but rather to carry on the interests of developers, advertisers, or third-party stakeholders. This is evidenced in Table 3, in which we can observe that among free applications 78.92% collect data for *Analytics* purposes, 40.55% for *Developer Advertising*, and 36.41% for *Third-party Advertising*. For instance, referring to Figure 5, we can observe that *Diagnostics* data is sometimes collected for tracking purposes, suggesting that this data is shared with third parties for fingerprinting [16, 51] purposes. Similarly, we can observe that some outliers in the *Weather* category collect *Health & Fitness* data for tracking purposes. With respect to these *"opaque"* behaviors, privacy labels and runtime tracking notifications are a step in the right direction, providing some visibility into practices that previously were completely invisible to the eyes of the users. However, past research has shown that most users do not pay attention to the data used by an app before its installation [20, 22] and that runtime pop-up can be easily overlooked [2, 6, 37]. Therefore, an open research direction is in (i) understanding how the information made available by privacy labels may be used to create privacy indicators that can more easily be interpreted by users, and (ii) designing solutions that can assist the user in choosing applications, e.g., through easy-to-understand, yet jargon-free, privacy level indicators.

Lack of anonymization – Our analysis shows that most of the data collected by applications in the Apple App Store are not anonymized. In particular, the results of RQ1 (Section 5.1) show that free applications on average perform 1.36 usages of tracking data and 2.72 usages of linked data, both of which are unanonimyzed. The same applications on average perform only 1.86 usages of unlinked data. Paid applications have a better ratio, with 0.2 usages of tracking data, 0.43 of linked data, and 0.9 of unlinked data. A more detailed breakdown of differences for the different data types collected is provided in Figure 5. From it, we can notice further differences in how some data types are collected. For instance, within the *Weather* category, location data is collected more frequently in a linked form than in an anonymized fashion. We find this trend to be concerning, as the use of unanonymized data may compromise the user's privacy [5, 24, 54]. Although this is unavoidable for those cases when data cannot be anonymized without compromising the functionality of the app (e.g., *Contact Info* naturally contains the user address, which cannot be stripped out when a billing address is needed), it also highlights a lack of attention from developers towards data anonymization. Future research is needed to understand the reasons behind this behavior and how developers can be nudged towards privacy-preserving data collection practices that

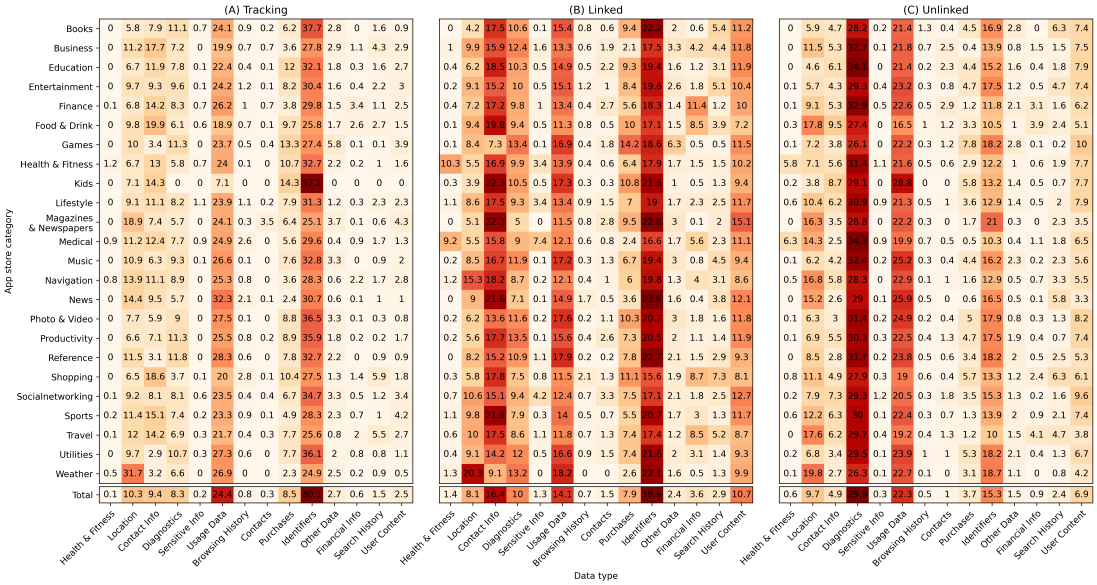

Figure 5: Percentage of data usage by App Store category

account for more effective anonymization as a priority. In this direction, an alternative could be that of adopting privacy-preserving approaches by employing synthetic data generation techniques, which enable the automatic generation of data that mimic the original data (and this is good for developers) while making it impossible to (re-)identify users (and this is good for privacy). The main challenge here resides in the word *mimic* since the synthetic data must have the same statistical properties of the original data without disclosing the original data as was explored in the medical domain [36].

Tracking in the App Store – The first consideration that emerges from our results is that tracking of users for advertising purposes is commonplace for apps in the Apple App store. Indeed, from the results of RQ1, described in Section 5.1, it emerged that almost 50% of the free applications in our dataset collect tracking data. This ratio is much lower in paid applications, stopping at 10.48%. However, analyzing the results of RQ2, provided in Section 5.2, we observe glaring differences among categories, with categories like *Games* and *News* having the majority of apps performing tracking, and others like *Business* and *Kids* in which tracking is performed only by a limited minority. Similarly to differences between free and paid applications [14], these discrepancies are partially attributable to the different business models employed by apps in different categories. Indeed, some types of apps (e.g., *Games*) are likely to rely more heavily on advertising for revenue, as evidenced in Figure 4, thus explaining the high utilization of tracking data in some categories. Another reason behind these differences is that apps in other categories are subject to tighter constraints on data collection or undergo a stricter review process prior to publication. This is the case of the *Kids* category, for which regulations explicitly prohibit the collection of minor's data for marketing and users profiling [19], and the *Medical* category, for which Apple self-imposes tighter review guidelines [28]. Given the limited use of tracking found within these categories, we consider that these measures are indeed effective in protecting the privacy of users and hence we encourage privacy regulators and platform maintainers to gradually expand these regulations to other categories.

In addition, the results discussed above lead us to recommend those Apple iOS users, who pay more attention to privacy and are more worried about sensitive data, to install paid applications in place of free ones, when possible. To some extent, our recommendation is a direct consequence of the monetization model adopted by free versus paid apps. It is in fact well known that paid apps make money (mainly) with customers paying for downloading the app; whereas, free apps, beyond paid subscriptions, make money by means of advertising and in-app purchases that are more incisive and remunerative if (sensitive) personal data are (directly) exploited. The different impact that advertising has on free and paid applications is immediately noticeable in Table 3: among free applications, 40.55% and 36.41% declare collecting data for developer and third-party advertising, respectively; for paid applications, only 7.44% and 4.85% declare data collection for the same purposes. In this sense, the bitter end is that privacy protection, in most cases, lies "between the hammer and the anvil", with end-users either required to pay or forced to accept reduced privacy.

Effects of the App Tracking Transparency – From the results of RQ5, presented in Section 5.5, we observed a reduction in usage of tracking data after the introduction of the App Tracking Transparency framework. The number of free applications employing tracking data experienced a reduction of about two percentage points (dropping to 46.3% from an original 48.79%), while paid applications experienced a reduction of about four percentage points

(from 10.48% to 6.21%). We believe this change is positive, albeit limited, and it highlights that having to explicitly ask for the user's permission discourages developers from performing tracking while providing more control to end-users. However, differences across categories have to be taken into account: the reduction in the use of tracking data is not uniform among categories, with the *Books*, *Games* and *News* categories reporting an increase, in contrast with the general trend. Again these categories are among those that are more likely to rely on advertising to generate revenue, as evidenced in Figure 4. Still, this highlights a contrast of interests between the various stakeholders involved in the mobile domain, as refraining from user tracking would equate to renounce part of the possible revenue for developers of these applications.

7 THREATS TO VALIDITY

In the following we discuss the threats to validity of our study according to the Cook and Campbell categorization [13].

Internal validity refers to the causality relationship between treatment and outcome [53]. In our study, we analyzed the data data-handling practices of iOS applications relying on privacy labels. However, the latter are generated based on what the developer declares about the application and, therefore, may be inaccurate. Developers may also use workarounds and covert techniques to collect data without declaring it [41]. The impact of these threats is limited by the fact that the code of all applications undergoes a review process by Apple prior to publication on the App Store.

Moreover, in RQ5 we attribute the differences between S and S' to the introduction of the App Tracking Transparency framework. However, this difference may also be due to other factors, e.g., developers gaining better knowledge and familiarity of privacy labels over time. However, the introduction of the App Tracking Transparency framework was the most visible privacy-relevant change introduced on the iOS platform in the considered time frame.

External validity deals with the generalizability of obtained results [53]. To ensure that our subjects are representative of the population of iOS apps, we considered a sample of 30,753 apps available in the United States selected from all categories of the App Store. Since the apps are selected from the top free and top-grossing app rankings of all categories, we can expect that they have a high number of users. Of these 17,312 were found to have updated their product page with privacy labels. Our final datasets have an imbalance in the number of free and paid applications, favoring the former. Hence our results might not be representative of the general population of paid applications. However, free apps were found to be the majority of apps in the top-grossing ranking and are generally downloaded more often [23].

Construct validity deals with the relation between theory and observation [53]. We mitigated potential construct validity threats by defining all details related to the design of our study (e.g., the goal, research questions, hypotheses, statistical analysis procedures) before starting the study execution.

Conclusion validity deals with issues that affect the ability to draw the correct conclusions from the outcome of experiments [53]. To mitigate this threat, while answering RQ5, we relied not only on the results of statistical tests but also examined differences between S and S' by means of descriptive statistics. The computed statistics support the experiment results. In addition, we provide a publicly available replication package for independent verification of our findings.

8 CONCLUSIONS AND FUTURE WORK

We conducted a large-scale empirical study to characterize the data collection practice of mobile apps in the Apple App Store.For this purpose, we collected and analyzed the privacy labels of 17,312 apps available on the store. From the results of our analysis, it is highlighted that on average free applications collect more sensitive data than paid ones, the majority of data is collected in an unanonimyzed form, and a wide range of sensitive information are collected for tracking purposes. Furthermore, the introduction of additional run-time tracking controls by Apple resulted in a statistically significant decrease in the number of apps that collect data for tracking purposes.

Future work involves pursuing the future research directions discussed in Section 6, namely designing privacy indicators that can more easily be interpreted by users and investigating how developers can be nudged towards privacy-preserving data collection practices. In addition, an open research direction is employing the privacy labels to characterize the data collection practices of desktop applications and how these differ from the ones of mobile apps ones.

9 REFERENCES

[1] Yuvraj Agarwal and Malcolm Hall. 2013. ProtectMyPrivacy: detecting and mitigating privacy leaks on iOS devices using crowdsourcing. In Proceeding of the 11th annual international conference on Mobile systems, applications, and services. 97–110.

[2] Devdatta Akhawe and Adrienne Porter Felt. 2013. Alice in warningland: A large-scale field study of browser security warning effectiveness. In 22nd {USENIX} Security Symposium ({USENIX} Security 13). 257–272.

[3] Panagiotis Andriotis, Gianluca Stringhini, and Martina Angela Sasse. 2018. Studying users? adaptation to android?s run-time fine-grained access control system. Journal of information security and applications 40 (2018), 31–43.

[4] Manoj Balasubramanian. 2021. App Tracking Transparency Opt-In Rate - Monthly Updates. https://www.flurry.com/blog/att-opt-in-rate-monthly-updates/

[5] Rebecca Balebako, Jaeyeon Jung, Wei Lu, Lorrie Faith Cranor, and Carolyn Nguyen. 2013. " Little brothers watching you" raising awareness of data leaks on smartphones. In Proceedings of the Ninth Symposium on Usable Privacy and Security. 1–11.

[6] Rainer Böhme and Jens Grossklags. 2011. The security cost of cheap user interaction. In Proceedings of the 2011 New Security Paradigms Workshop. 67–82.

[7] Theodore Book, Adam Pridgen, and Dan S. Wallach. 2013. Longitudinal Analysis of Android Ad Library Permissions. CoRR abs/1303.0857 (2013).

[8] Lionel C Briand, Christiane M Differding, and H Dieter Rombach. 1996. Practical guidelines for measurement-based process improvement. Software Process: Improvement and Practice 2, 4 (1996), 253–280.

[9] Paolo Calciati, Konstantin Kuznetsov, Alessandra Gorla, and Andreas Zeller. 2020. Automatically Granted Permissions in Android apps: An Empirical Study on their Prevalence and on the Potential Threats for Privacy. In Proceedings of the 17th International Conference on Mining Software Repositories. 114–124.

[10] Victor R Basili1 Gianluigi Caldiera and H Dieter Rombach. 1994. The goal question metric approach. Encyclopedia of software engineering (1994), 528–532.

[11] The New York Times Company. 2021. What We Learned From Apple's New Privacy Labels. https://www.nytimes.com/2021/01/27/technology/personaltech/apple-privacy-labels.html

[12] William Jay Conover. 1999. Practical nonparametric statistics. Vol. 350. john wiley & sons.

[13] Thomas D Cook, Donald Thomas Campbell, and Arles Day. 1979. Quasi-experimentation: Design & analysis issues for field settings. Vol. 351. Houghton Mifflin Boston.

[14] Soteris Demetriou, Whitney Merrill, Wei Yang, Aston Zhang, and Carl A Gunter. 2016. Free for All! Assessing User Data Exposure to Advertising Libraries on Android.. In NDSS.
[15] Zhui Deng, Brendan Saltaformaggio, Xiangyu Zhang, and Dongyan Xu. 2015. iris: Vetting private api abuse in ios applications. In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. 44–56.
[16] Peter Eckersley. 2010. How unique is your web browser?. In International Symposium on Privacy Enhancing Technologies Symposium. Springer, 1–18.
[17] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the experts: What should be on an IoT privacy and security label?. In 2020 IEEE Symposium on Security and Privacy (SP). IEEE, 447–464.
[18] Zheran Fang, Weili Han, and Yingjiu Li. 2014. Permission based Android security: Issues and countermeasures. computers & security 43 (2014), 205–218.
[19] Álvaro Feal, Paolo Calciati, Narseo Vallina-Rodriguez, Carmela Troncoso, and Alessandra Gorla. 2020. Angel or devil? a privacy study of mobile parental control apps. Proceedings on Privacy Enhancing Technologies 2020, 2 (2020), 314–335.
[20] Adrienne Porter Felt, Erika Chin, Steve Hanna, Dawn Song, and David Wagner. 2011. Android permissions demystified. In Proceedings of the 18th ACM conference on Computer and communications security. 627–638.
[21] Adrienne Porter Felt, Serge Egelman, Matthew Finifter, Devdatta Akhawe, David A Wagner, et al. 2012. How to Ask for Permission. HotSec 12 (2012), 7–7.
[22] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android permissions: User attention, comprehension, and behavior. In Proceedings of the eighth symposium on usable privacy and security. 1–14.
[23] Inc Gartner. 2012. Gartner Says Free Apps Will Account for Nearly 90 Percent of Total Mobile App Store Downloads in 2012. http://www.gartner.com/newsroom/id/2153215.
[24] Clint Gibler, Jonathan Crussell, Jeremy Erickson, and Hao Chen. 2012. Androidleaks: Automatically detecting potential privacy leaks in android applications on a large scale. In International Conference on Trust and Trustworthy Computing. Springer, 291–307.
[25] Michael C Grace, Wu Zhou, Xuxian Jiang, and Ahmad-Reza Sadeghi. 2012. Unsafe exposure analysis of mobile in-app advertisements. In Proceedings of the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks. 101–112.
[26] Jin HAN, Qiang YAN, Debin GAO, Jianying ZHOU, and Robert H DENG. 2013. Comparing Mobile Privacy Protection through Cross-Platform Applications. In Proceedings of NDSS 2013: Network and Distributed System Security Symposium, 24-27 February, San Diego.
[27] Apple Inc. 2020. App privacy details on the App Store. https://developer.apple.com/app-store/app-privacy-details/
[28] Apple Inc. 2020. App Store Review Guidelines. https://developer.apple.com/app-store/review/guidelines/#health-and-health-research
[29] Apple Inc. 2020. User Privacy and Data Use. https://developer.apple.com/app-store/user-privacy-and-data-use/
[30] Ajay Kumar Jha, Sunghee Lee, and Woo Jin Lee. 2017. Developer mistakes in writing android manifests: An empirical study of configuration errors. In 2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR). IEEE, 25–36.
[31] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. 2009. A "nutrition label" for privacy. In Proceedings of the 5th Symposium on Usable Privacy and Security. 1–12.
[32] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. 2013. Privacy as part of the app decision-making process. In Proceedings of the SIGCHI conference on human factors in computing systems. 3393–3402.
[33] 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, 1 (2022), 6–27.
[34] Erich L Lehmann and Joseph P Romano. 2006. Testing statistical hypotheses. Springer Science & Business Media.
[35] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I Hong. 2014. Modeling users? mobile app privacy preferences: Restoring usability in a sea of permission settings. In 10th Symposium On Usable Privacy and Security ({SOUPS} 2014). 199–212.
[36] Helena Montenegro, Wilson Silva, and Jaime S Cardoso. 2021. Privacy-Preserving Generative Adversarial Network for Case-Based Explainability in Medical Image Analysis. IEEE Access 9 (2021), 148037–148047.
[37] Sara Motiee, Kirstie Hawkey, and Konstantin Beznosov. 2010. Do Windows users follow the principle of least privilege? Investigating user account control practices. In Proceedings of the Sixth Symposium on Usable Privacy and Security. 1–13.
[38] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In Proceedings of the 2020 CHI conference on human factors in computing systems. 1–13.
[39] Anthony Peruma, Jeffrey Palmerino, and Daniel E Krutz. 2018. Investigating user perception and comprehension of android permission models. In Proceedings of the 5th International Conference on Mobile Software Engineering and Systems. 56–66.
[40] Abbas Razaghpanah, Rishab Nithyanand, Narseo Vallina-Rodriguez, Srikanth Sundaresan, Mark Allman, Christian Kreibich, Phillipa Gill, et al. 2018. Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem. In The 25th Annual Network and Distributed System Security Symposium (NDSS 2018).
[41] Joel Reardon, Álvaro Feal, Primal Wijesekera, Amit Elazari Bar On, Narseo Vallina-Rodriguez, and Serge Egelman. 2019. 50 ways to leak your data: An exploration of apps' circumvention of the android permissions system. In 28th {USENIX} Security Symposium ({USENIX} Security 19). 603–620.
[42] Lena Reinfelder, Andrea Schankin, Sophie Russ, and Zinaida Benenson. 2018. An inquiry into perception and usage of smartphone permission models. In International Conference on Trust and Privacy in Digital Business. Springer, 9–22.
[43] Jingjing Ren, Martina Lindorfer, Daniel J Dubois, Ashwin Rao, David Choffnes, Narseo Vallina-Rodriguez, et al. 2018. Bug fixes, improvements,... and privacy leaks. In The 25th Annual Network and Distributed System Security Symposium (NDSS 2018).
[44] Gian Luca Scoccia, Ivano Malavolta, Marco Autili, Amleto Di Salle, and Paola Inverardi. 2019. Enhancing trustability of android applications via user-centric flexible permissions. IEEE Transactions on Software Engineering 01 (2019), 1–1.
[45] Gian Luca Scoccia, Anthony Peruma, Virginia Pujols, Ivano Malavolta, and Daniel E Krutz. 2019. Permission issues in open-source Android apps: An exploratory study. In 2019 19th International Working Conference on Source Code Analysis and Manipulation (SCAM). IEEE, 238–249.
[46] Gian Luca Scoccia, Stefano Ruberto, Ivano Malavolta, Marco Autili, and Paola Inverardi. 2018. An investigation into Android run-time permissions from the end users' perspective. In Proceedings of the 5th international conference on mobile software engineering and systems. 45–55.
[47] Yun Shen and Pierre-Antoine Vervier. 2019. Iot security and privacy labels. In Annual Privacy Forum. Springer, 136–147.
[48] Forrest Shull, Janice Singer, and Dag IK Sjøberg. 2007. Guide to advanced empirical software engineering. Springer.
[49] Junwei Tang, Ruixuan Li, Hongmu Han, Heng Zhang, and Xiwu Gu. 2017. Detecting permission over-claim of android applications with static and semantic analysis approach. In 2017 IEEE Trustcom/BigDataSE/ICESS. IEEE, 706–713.
[50] Maciej Tomczak and Ewa Tomczak. 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends in sport sciences 1, 21 (2014), 19–25.
[51] Zhen Tu, Runtong Li, Yong Li, Gang Wang, Di Wu, Pan Hui, Li Su, and Depeng Jin. 2018. Your apps give you away: distinguishing mobile users by their app usage fingerprints. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 3 (2018), 1–23.
[52] Inc. Wirecutter. 2021. We Checked 250 iPhone Apps-This Is How They're Tracking You. https://www.nytimes.com/wirecutter/blog/how-iphone-apps-track-you/
[53] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. 2012. Experimentation in software engineering. Springer Science & Business Media.
[54] Zhemin Yang, Min Yang, Yuan Zhang, Guofei Gu, Peng Ning, and X Sean Wang. 2013. Appintent: Analyzing sensitive data transmission in android for privacy leakage detection. In Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. 1043–1054.