Analyzing Privacy Implications of Mobile Apps Data Collection across Age Groups

Adamo Mariani
University of Twente
a.mariani@student.utwente.nl

Matteo Liberato
University of Twente
m.liberato@utwente.nl

Anna Sperotto
University of Twente
a.sperotto@utwente.nl

Antonia Affinito
University of Twente
a.affinito@utwente.nl

***Abstract*—Mobile applications increasingly access a wide range of personal information, raising significant privacy concerns, particularly for children and teenagers. Previous studies have shown low compliance between privacy policies and permissions in mobile apps. However, current research has not yet explored how an app's target age group influences the permissions it requests, especially among minors. While recent regulatory frameworks like COPPA, CCPA, and GDPR establish clear rules for data acquisition and privacy for specific age groups, their practical application remains uncertain. This research investigates how data collection practices align with privacy policies among mobile applications targeting different age groups. We show that, on average, the same application collects more user data when downloaded from Google Play than the Apple App Store. Furthermore, applications targeting teenagers collect data more frequently than those targeting other age groups, indicating the necessity for strict regulations for this age group.**

***Index Terms*—Mobile Applications, Privacy Policies, COPPA, GDPR, Google Play, Apple App Store, Age Groups**

I. Introduction

In the last few years, mobile application stores have experienced significant growth, which is only expected to continue. Mobile applications (apps) have become an integral part of our daily lives, with as many as 257 billion downloaded apps worldwide in 2023 [1]. However, this proliferation has raised significant privacy concerns, as many apps collect a variety of data often unrelated to their primary functions. To protect against possible misuse, mobile operating systems require apps to request the user's permission before accessing the user's data [2]. Furthermore, it is increasingly common for children to use apps under the supervision of parents or relatives. Regulations like the Children's Online Privacy Protection Act (COPPA)[1], the California Consumer Privacy Act (CCPA) [2] and the EU General Data Protection Regulation (GDPR) [3] specifically aim to protect children's data handling and ensure their privacy is safeguarded. However, many apps still fail to fully comply with these requirements, potentially exposing vulnerabilities that exploit young users [3]. Also, it has been shown that data collected by the apps targeting children below 5 years is even more valuable to advertisers [4]. Therefore, it is key to determine if there is a correlation between an app's target age group and the types and the amount of data it collects. Previous research works investigated the behavior and privacy policies of apps targeting children below the age of 12 [3], [5]–[7]. The ratings included in the app store listings, such as those from the official Google Play Store and Apple App Store, categorize apps based on age suitability. The official Google Play Store guidelines differentiate between the following targeted age groups: ≤ 5, 6-8, 9-12, 13-15, 16-17, $18+$, with additional policies applied to apps targeting children under the age of 12 [8]. The Apple App Store considers similar age groups of $4+$, $9+$, $12+$, and $17+$ [9]. Except for the "Children" category, these ratings (i.e., PEGI) indicate the minimum age for which an app is suitable. For instance, "Booking.com: Hotels and more" has a PEGI rating of 3 despite it targets adults. However, we believe that a clearer division of age groups is needed. In particular, in this work, we categorize the age groups as Children (≤ 12), Teens (13-17), and Adults ($18+$).

This paper aims to analyze the influence of an app's target age group on its data collection practices, uncovering patterns that could inform future regulatory frameworks and raise awareness among end users. Specifically, we explore the privacy policies and the data collection practices of the applications listed in the two main app stores: Google Play and Apple App. The study addresses the following research question: How effectively do data collection practices comply with privacy policies among apps targeting different age groups? To address this question, we investigate the following sub-questions:

1) To what extent can we predict the target age group of mobile applications using data extracted from mobile store listings?
2) What are the differences in data collection practices among mobile applications targeting different age groups on the Google Play and Apple App stores?
3) To what extent do data safety practices disclosed in privacy policies align with mobile store listings?

This paper is structured as follows: we discuss related work in Section II, and detail our methodology in Section III. In Section IV we present our results. Finally, we provide conclusions in Section VI and discuss implications in Section V.

[1]https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa

[2]https://oag.ca.gov/privacy/ccpa

[3]https://gdpr-info.eu/

978-3-903176-66-9 ©2024 IFIP

II. Related Work

Several studies have analyzed the privacy issues of mobile applications [10], [11]. Luo et al. [12] found that games and educational apps, which often target children, have a high rate of privacy violations. Other studies examined apps targeting children but focused only on those available in the "Family" and "Kids" categories of the Google Play and Apple stores [3], [13]. These studies highlight privacy policy violations in such apps, including sharing data with third parties, compliance issues, and missing privacy policies. But these apps represent only a fraction of apps targeting children, which are reportedly more than 90% of all apps [14]. Researchers have attempted to address the presence of apps targeting children that are not included in children's categories of app stores. Sun et al. [7] analyzed 20K apps targeting children, 15K of which were not in the children-specific category, and found that over 36% requested access to location information. This finding highlights the issue of over-claimed permissions, where apps request more permissions than needed, creating significant privacy risks [15]. Multiple studies have assessed permission systems, revealing their vulnerabilities [16], [17], and developed tools to analyze these issues in both generic [18] and specific app categories [19]. Other researchers analyzed the differences between permissions and privacy policies, often using machine learning techniques [20]. Majethiya et al. [21] conducted a review on over-claimed permissions, concluding that semantic analysis is the most efficient method for this type of analysis. Results show a discrepancy between an app's declared requirements and the data it actually collects on both Android and IOS [22], with permissions being manipulated to gather users' data. For instance, Verderame et al. [23] show that more than 95% of Android apps access sensitive information, yet only 1% comply with the Google Play privacy guidelines. To our knowledge, no classifier uses semantic analysis of iOS and Android app store listings to distinguish between target age groups, and most related research does not fully consider age demographics.

III. Methodology

In this section, we describe the methodology and the data sources used in our study. As mentioned in Section I, our goal is to analyze whether data collection practices comply with privacy policies among applications targeting different age groups. To achieve this, we structured our methodology into several steps, each corresponding to our sub-research questions. Specifically, our system involves several components: Google Play and Apple App Store scrapers; a classifier able to determine the age group to which an application is targeted; a cross-platform comparison of data collection practices between Google and Apple; and a privacy policy parser to verify compliance with the app store listing information. Figure 1 illustrates the steps involved in our methodology.

App Store Scrapers. We created a dataset comprising metadata from 1622 applications, 811 from the Google Play Store and 811 from the Apple App Store, as these are the largest app stores globally [24]. To ensure diversity and variability in our dataset, we randomly selected applications across all categories. We retrieved Google Play application metadata using the google-play-scraper[4]. For the Apple App Store, we developed a specific scraper using the selenium web-driver in combination with the app-store-scraper[5]. For each application, we retrieved the following information: app listing details including title, description, age rating, and category; privacy policy; and data safety measures.

Target Age Group of an App Module As mentioned in Section I, both Google Play and Apple App stores specify the intended audience for apps targeting children, categorizing them into age bands. These classifications help users understand the suitability of apps for different age groups and ensure compliance with regulations such as COPPA, CCPA and GDPR, which explicitly target protecting children's data. Differentiating between Teen and Adult categories is more challenging due to the content ratings provided for apps, such as PEGI [6] or ESRB [7] scores. These ratings indicate the minimum age suitability rather than the true target audience. For example, apps rated *PEGI 18* or *Mature* target adults. However, apps rated *PEGI 12* or *Teen*, which are outside the "Children" category, require further analysis to determine whether they are intended for teens or adults.

Therefore, to answer the first sub-question - How accurately can the targeted age of mobile applications be predicted using mobile store listing information? - we developed a classifier that accurately distinguishes between teen and adult age groups. To this end, we selected 500 mobile applications from our dataset. We implemented an algorithm to establish the ground truth for training the classifier. The first step involved checking if an app targets children using two approaches. The initial approach required manually examining the app category and whether the store categorized it under specific children's age bands. If confirmed, the app was labeled as targeting children, and further steps were bypassed. Alternatively, we compiled synonyms like *child*, *kid*, and *toddler* and counted their occurrences in app titles, descriptions, and the first 50 reviews of apps in the "Family" category. Based on a manual investigation, we established that if these words appeared three times within the app's title and description, or six times within the app's first review page, the app would be classified as targeting children. The second step involved labeling applications targeting teenagers and adults. For this purpose, we retrieved the age groups of user accounts using each application. This information was provided by *Data.ai* [25], a platform specialized in data aggregation and analysis of mobile apps and digital markets, including user demographics.

After completing the data labeling process, we proceeded to select features, choosing those corresponding to the listing information - *title*, *description*, *categories*, and *content rating* - of an application. To align the ratings across the two stores, we mapped the ESRB and PEGI ratings of Android applications to

[4]GitHub Repository: https://github.com/facundoolano/google-play-scraper
[5]GitHub Repository: https://github.com/facundoolano/app-store-scraper
[6]Pan-European Game Information score part of IARC
[7]Entertainment Software Rating Board score part of IARC

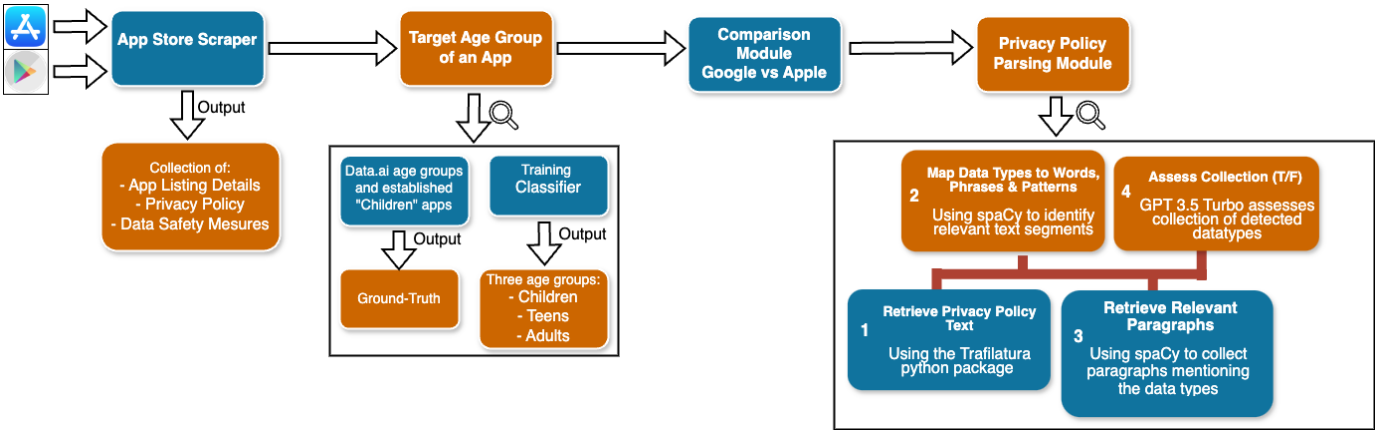

Fig. 1. Methodology Overview

Apple store content ratings using the conversion table provided on the Apple Developer website[8].

For the classification we have relied on the pre-trained GPT-3.5[9] by OpenAI, leveraging its flexible interface for adjusting training parameters. We selected the base model (GTP3.5), provided it with pre-split training and validation datasets, and configured the training parameters accordingly. After fine-tuning our models, we tested them by comparing their output with the previously established ground truth.

Comparison Module - Google Play vs App Store After determining the age group each app targets, we compared the data type collection practices and privacy policies of Apple and Google apps, categorized per age group and platform. Data types refer to specific categories of data that applications collect, store, and use for various purposes. Apple[10] and Google Play[11] classify several data types such as location or contacts, with some methodological alignment. We created a conversion table to map the differences between the two stores' data types, which we omit for brevity. Unmatched types, including "Files and docs" and "Calendar events" are excluded from the analysis.

To compare the data types collected by Apple and Google apps, we matched apps in the Google Play store with the respective app in the Apple Store. We selected 1100 applications from the Google Play store and retrieved the corresponding applications from the App Store. We manually analyzed the titles of a subset of 100 Google Play Store apps on the App Store. While most titles were identical, some showed slight variations. Using the Levenshtein distance formula, we calculated the similarity between these app titles and determined the optimal distance threshold for matching application names. Specifically, we categorized apps that did not find an exact match on the Apple App Store as *False Matches*. The closest match was identified as the first search result when entering the unmatched app's name. Conversely, we labeled *True Matches*

[8]https://developer.apple.com/help/app-store-connect/reference/age-ratings/
[9]https://platform.openai.com/docs/models/gpt-3-5-turbo
[10]developer.apple.com/app-store/app-privacy-details
[11]support.google.com/googleplay/android-dev

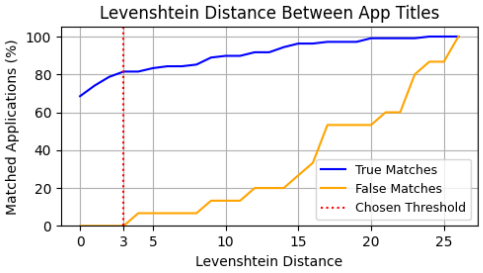

Fig. 2. Levenshtein Distance of App Names: Google Play vs. Apple Store Apps

as applications found in both stores with slight differences in their names. Figure 2 indicates that around 65% of applications have a Levenshtein distance of 0, which means their names were identical. Additionally, by using a distance threshold of 3 (red dotted line in the plot), we successfully identified over 80% of true matches while effectively excluding false matches.

Privacy Policy Parsing Module The final component of our methodology involves parsing the privacy policies of applications from both stores to verify compliance with the collected data types. Figure 1 includes a visualization of the privacy policy parsing methodology. Once the privacy policy URL is retrieved, the first step involves developing a script for extracting the privacy policy. For this purpose, we used Trafilatura, noted for its speed and effectiveness [26]. In the second step, we used the spaCy library for NLP to identify and map text segments in policies to specific data collection types. We created a mapping for various words, phrases, and patterns. For each detection, spaCy retrieved the data type and its surrounding paragraph. Then, we used GPT-3.5 to analyze the paragraphs and confirm if the detected data types were collected. This extra step aims to improve accuracy, with the idea that a contextual analysis of the paragraph may exclude cases where privacy policies mention data types that are not actually collected. Finally, we compared the data types disclosed in the privacy policies to those disclosed on the collected application listings. The matched and unmatched

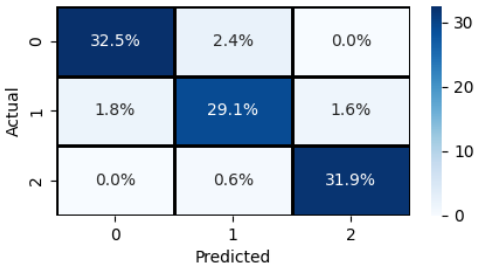

Fig. 3. Confusion Matrix of the Best Performing Model

data types were stored for each app to analyze patterns of undisclosed permissions across age groups and categories.

IV. Results

A. The Classifier

To predict a targeted age group, we trained several classifiers using different combinations of features, train-test splits, and hyperparameters. The performance metrics are recorded in Table I.

Initial results showed that an 80-20 train-test split performed worse than a 70-30 split, possibly due to the small dataset size. Therefore, we excluded those models from the results table. We also varied hyperparameters such as batch size, learning rate (LR) multiplier, and number of epochs. The batch size refers to the number of samples processed before updating the model parameters. The LR multiplier scales the model's weights, influencing the update speed with each batch. Epochs represent the number of full cycles through the training dataset. To assess the significance of false negatives and false positives, we mapped categories as follows: *Adults* $\rightarrow$ 0, *Teens* $\rightarrow$ 1, and *Children* $\rightarrow$ 2. If the classifier predicted a label of 2 on a true label of 1, a distance of $|2-1| = 1$ is recorded. Figure 3 shows the confusion matrix for the best-performing configuration, which achieved a maximum F1 score of 91.9% and an accuracy of 93.6%. This represents a notable improvement over the base model, labeled as model #0 in the table, which achieved an accuracy of 66.2%. Our results indicate that models trained with the feature set comprising app titles, descriptions, and categories generally outperform those trained with other features. The addition of app content rating as a feature negatively affected the classifier's performance, suggesting that target age and content rating may not be closely correlated. Surprisingly, we did not find any instances where the predicted and actual categories differed by two distances (i.e., classifying a Children's app as an Adult one, or vice versa).

B. A cross-platform comparison

As explained in Section III, we used the best-performing classifier to label the 1.622 applications retrieved from the Apple App and Google Play stores, categorizing them into three age groups: Children, Teens and Adults. We then extracted the data safety measures, including the types of data collected and their collection purposes.

Table II displays the percentage of applications that collect specific types of data, categorized by platform and target age group. To examine the dependency between data types and platforms (listed under "By Platform") and between data types and age groups (listed under "By Age Group"), we conducted a chi-squared test of independence. We identified significant associations with p-values below 0.05. A significance level below 0.05 means there is less than a 5% chance of incorrectly concluding there is a difference between categories when, in reality, there is no actual difference. Except for a few data types, age groups have an average significance value of 0.0003, indicating the statistical significance age groups have on data collection practices. Additionally, the analysis also revealed a significant dependency between data types and platform, with p-values lower than 0.05 except for a few cases.

Interesting cases include device ID, product interaction, and advertising data that were collected more frequently. The device ID uniquely identifies devices and tracks user habits across platforms. Purchase history, linked to the device ID, reveals buying patterns and preferences, enabling tailored advertisements. Advertising data evaluates user interactions and engagement with ads. These data types have the highest overall collection rates on both platforms. Teenagers, in particular, show significantly higher rates of data collection in these categories. Furthermore, chi-square test values for these types of data confirm significant differences in how applications treat adults and teenagers as separate age groups. To confirm that our results are not biased by rates for Children, in Table III, we specifically evaluated the chi-square test only for Teens and Adults. We confirmed that the p-value remains below 0.05, indicating that certain data types are more frequently collected in applications targeting teenagers compared to adults.

1) Data Collection Practices by Age Group: As discussed in Section II, existing literature indicates that many apps tend to over-claim permissions to gather extensive user data. Given the high value of children's data and minimal regulatory restrictions for teenagers (often treated similarly to adults), we hypothesized that developers would target teenagers the most for data collection. To explore this hypothesis, we analyzed the data collection practices of the applications included in our dataset. Figures 4 (a) and (b) show the distribution of data collection purposes of Google Play and App Store applications, respectively, revealing teenagers as the primary target across all categories, which supports our hypothesis. Notably, data collected explicitly for third-party advertising stands out the most, with approximately 53% of apps targeting teens compared to 29% targeting adults. The distribution of data collection purposes remains consistent across platforms, indicating that the same application typically collects data for similar purposes on both Google Play and the App Store.

2) Data Collection Practices by Platform: Except for undisclosed data types or collection purposes not recorded on Google Play apps, the mobile platform does not affect data collection purposes. However, it significantly influences the

TABLE I
CLASSIFIER FEATURE COMBINATIONS AND PERFORMANCE METRICS

#	Feature Set	Batch Size	LR Multiplier	Epochs	Precision	Recall	F1 Score	Accuracy
0	[title, description, rating, categories]	-	-	-	0.435	0.998	0.606	0.480
	[title, description, categories]				0.537	0.983	0.695	**0.662**
	[title, description]				0.529	0.975	0.686	0.643
1	[title, description, rating, categories]	1	2	3	0.888	0.883	0.886	**0.909**
2	[title, description, contentRating]	1	2	3	0.890	0.915	0.902	**0.922**
3	[title, description]	2	1.5	5	0.858	0.926	0.891	0.911
		4	1	10	0.866	0.915	0.890	0.911
		2	2	5	0.876	0.920	0.898	0.918
		4	1.5	3	0.870	0.949	0.908	0.924
		1	2	3	0.882	0.932	0.906	0.924
		4	1	3	0.879	0.949	0.913	**0.929**
		4	0.5	3	0.883	0.943	0.912	**0.929**
4	[title, description, categories]	4	0.5	3	0.911	0.850	0.879	0.907
		1	2	4	0.921	0.911	0.916	0.933
		1	2	3	0.927	0.911	**0.919**	**0.936**

TABLE II
DATA TYPE COLLECTION BY TARGET AGE AND PLATFORM

	Apple App Store			Google Play Store			By Age Group		By Platform	
Data Type	**Adults**	**Teens**	**Children**	**Adults**	**Teens**	**Children**	χ^2	**p-value**	χ^2	**p-value**
Browsing History	1.83%	2.12%	1.69%	0.61%	1.33%	0.00%	1.2354	0.5392	3.5217	0.0606
Email Address	39.02%	28.38%	24.58%	54.57%	43.77%	25.42%	23.8681	**<0.0001**	18.9655	**<0.0001**
Name	28.66%	15.12%	5.08%	42.99%	32.10%	14.41%	48.9176	**<0.0001**	34.1376	**<0.0001**
Other User Contact Info	3.05%	2.92%	0.00%	23.48%	16.45%	7.63%	15.5562	**0.0004**	95.4379	**<0.0001**
Phone Number	12.50%	3.18%	1.69%	22.87%	7.69%	4.24%	66.3323	**<0.0001**	17.7805	**<0.0001**
Physical Address	4.57%	1.33%	0.00%	10.37%	1.59%	0.00%	44.8101	**<0.0001**	6.6667	**0.0098**
Contacts	3.05%	7.16%	0.85%	4.88%	7.43%	0.85%	17.3341	**0.0002**	0.5904	0.4423
Credit Info	0.30%	0.27%	0.00%	0.61%	0.00%	0.00%	2.1913	0.3343	0.0000	1.0000
Other Financial Info	3.66%	0.53%	0.00%	10.37%	0.80%	0.00%	54.1751	**<0.0001**	10.3725	**0.0013**
Payment Info	3.05%	1.06%	0.00%	7.32%	3.45%	0.85%	16.0470	**0.0003**	11.0769	**0.0009**
Fitness	0.30%	0.80%	0.00%	1.83%	1.06%	0.00%	2.4226	0.2978	2.5714	0.1088
Health	0.00%	0.80%	0.00%	1.52%	0.27%	0.00%	1.8508	0.3964	1.0000	0.3173
Device ID	49.09%	67.64%	25.42%	64.02%	74.80%	51.69%	34.6680	**<0.0001**	11.4605	**0.0007**
User ID	43.60%	56.50%	29.66%	46.65%	48.81%	26.27%	23.9866	**<0.0001**	0.6970	0.4038
Coarse Location	27.13%	35.54%	20.34%	32.93%	42.71%	34.75%	11.8020	**0.0027**	7.1257	**0.0076**
Precise Location	15.55%	5.04%	1.69%	19.51%	5.31%	0.85%	75.9083	**<0.0001**	1.0764	0.2995
Purchase History	23.48%	40.85%	22.03%	28.96%	45.62%	17.80%	45.1660	**<0.0001**	1.7633	0.1842
Search History	6.71%	7.69%	2.54%	12.80%	14.06%	3.39%	12.3456	**0.0021**	13.2353	**0.0003**
Sensitive Info	12.50%	1.59%	0.00%	29.88%	16.98%	7.63%	56.0412	**<0.0001**	70.5321	**<0.0001**
Undisclosed	0.61%	1.06%	1.69%	0.00%	0.00%	0.00%	1.1079	0.5747	8.0000	**0.0047**
Advertising Data	28.66%	44.30%	15.25%	54.27%	56.50%	45.76%	17.7424	**0.0001**	38.0608	**<0.0001**
Other Usage Data	17.38%	26.53%	5.93%	20.43%	37.67%	24.58%	34.4833	**<0.0001**	13.6219	**0.0002**
Product Interaction	61.28%	70.03%	56.78%	54.27%	56.50%	45.76%	4.8140	0.0901	7.7472	**0.0054**
Audio Data	3.05%	4.24%	2.54%	8.54%	5.84%	3.39%	2.7500	0.2528	7.5301	**0.0061**
Emails or Text Messages	5.79%	4.77%	0.00%	25.30%	21.22%	4.24%	25.4648	**<0.0001**	83.7122	**<0.0001**
Gameplay Content	3.35%	18.04%	8.47%	20.43%	37.67%	24.58%	46.5331	**<0.0001**	67.8930	**<0.0001**
Other User Content	14.33%	11.67%	0.85%	22.87%	22.81%	7.63%	24.0179	**<0.0001**	23.2214	**<0.0001**
Photos or Videos	25.00%	15.65%	3.39%	34.45%	21.22%	4.24%	59.7797	**<0.0001**	8.1895	**0.0042**

type and quantity of data collected.

Figure 5 (a) shows the collection of Advertising Data, which is highly valuable to third parties and exhibits significant disparities across platforms and age groups. The figure highlights that apps on the App Store face challenges in obtaining and selling children's information to third parties. On the other hand, apps on the Google Play store aim to maximize advertising data collection across all age groups. Interestingly, apps downloaded on the Play Store collect coarse location data more frequently in children's applications. Figure 5 (b) shows the frequency of location data collection, with teenage-targeted apps having the highest count. These findings suggest that teenagers are a primary target for location data collection, highlighting the need for stricter data privacy measures and regulations for younger users. Another interesting example is "Gameplay Content", which is the only data type collected more frequently within children's apps than those targeting adults on both stores. On average, apps on the Google Play store collect 2.1 times more data than the same apps listed on the App Store. Given the higher amounts of data collected, it seems that developers have more freedom when publishing their applications on Google Play. Overall, it is evident from Table II that apps targeting children collect less information than the rest. This is likely due to of current regulations

TABLE III
ADULTS VS TEENS DATA TYPE COLLECTION

	Apple App Store		Google Play Store		By Age Group		By Platform	
Data Type	**Adults**	**Teens**	**Adults**	**Teens**	χ^2	**p-value**	χ^2	**p-value**
Device ID	49.09%	67.64%	64.02%	74.80%	11.7158	**0.0006**	6.3612	**0.0117**
Purchase History	23.48%	40.85%	28.96%	45.62%	28.7600	**<0.0001**	2.6024	0.1067
Advertising Data	28.66%	44.30%	54.27%	56.50%	6.0557	**0.0139**	25.9202	**<0.0001**

Fig. 4. Data Collection Purposes by Target Age Group: Apple App (a) vs. Google Play (b)

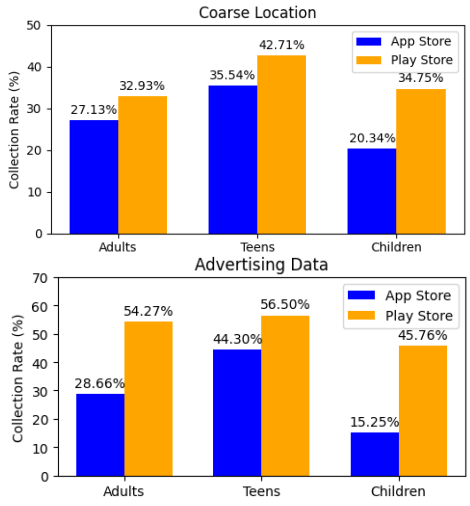

Fig. 5. Advertising Data and Coarse Location Collection Rates

providing a small layer of protection for children, which this research aims to extend to protect teenagers as well.

C. Privacy Policies vs. App Listings

Following the methodology outlined in Section III, we parsed and analyze the privacy policies of applications of our dataset. Due to complexities in mapping words and phrases to specific data types, 'Advertising Data' slightly changes its meaning to indicate applications that collect any data for advertisement purposes. For brevity, we do not report all the results of our analysis. However, the overall average data collection rate increased by about 10.1% for apps aimed at teens and children, and by 5% for apps aimed at adults, compared to what is declared in their app listings. This implies more consistency between the privacy policy and app listing information for apps targeting adults. Processing contact information, such as physical address, phone number, and name shows the highest overall difference, reaching 48% (particularly for apps targeting children). Interestingly, data collection for advertising purposes is present in 100% of the analyzed privacy policies for applications targeting teenagers. Except for Product Interaction and Coarse Location, significant differences (greater than 20%) show higher data collection rates in privacy policies. Analysis with the LLM GPT-3.5 revealed challenges in distinguishing between the declaration of Precise and Coarse Location data collection. This inaccuracy may explain the observed difference, which disappears when considering the average count of both data types.

V. DISCUSSION

In our study, the classifier developed to determine an app's targeted age group achieved a maximum accuracy of 93.6%, significantly higher than the base model's 66.2%. This result demonstrates the effectiveness of the semantic analysis approach used and indicates that analyzing app listings is a viable method for determining targeted age groups. In addition, the comparison between Google Play and Apple App revealed notable differences in data collection practices for the same apps, with apps on Google Play collecting more than twice the user data on average. This indicates a disparity in compliance with data safety regulations, possibly due to the App Store having stricter data safety guidelines, highlighting the need for uniform guidelines across platforms. Significant differences were found in the collection of data types across age groups, with higher rates in apps targeting teenagers. These data types had the highest collection rates among teenagers, suggesting that developers exploit the lack of restrictions on teenagers' data, targeting them more than adults despite similar privacy controls. We found discrepancies between data collection practices disclosed in privacy policies and those listed in app stores, with the most significant impact on apps targeting teenagers and children. While children need parental consent for data sharing, teenagers are expected to manage it themselves. Developers disclose more information in privacy policies than on store listings, indicating a lack of transparency, especially concerning younger users who are

less aware of the implications. Previous studies [19] found no significant differences in permissions requests between Google Play and Apple App stores. However, their comparison of privacy policies to app listings revealed that privacy policies disclose roughly 30% more data types collected. The discrepancy in the cross-platform analysis may be due to differences in methodologies. Also, when analyzing data collection rates by store category, we found no significant differences, likely due to the large number of categories in the dataset. However, *Games*, *Social Networking*, and *Lifestyle* had notably higher data collection rates. Further research is needed to support claims of systematic data collection in these categories.

VI. Conclusion

This study investigated the impact of a user's age group on data collection practices in mobile applications and their alignment with privacy policies. To this end, we tuned a highly accurate classifier to determine the targeted age group of mobile apps. We then analyzed the most frequently collected data types across age groups and platforms and calculated their statistical significance through multiple chi-squared tests of independence. Lastly, we proposed a methodology to parse privacy policies and identified notable inconsistencies between the disclosed data collection practices and those observed in app listings, especially for apps targeting teenagers and children.

Our findings highlight the need for stricter enforcement of data privacy regulations and emphasize the importance of making clearer distinctions between age groups. The significant increase in data collection by apps targeting teenagers should serve to inform future revisions of child data protection regulations, potentially leading to changes in age and permission requirements. Future work will involve expanding the training dataset to include edge cases and formally assessing GPT's accuracy in recognizing collected data types using an annotated set of policies.

References

[1] L. Ceci, "Annual number of mobile app downloads worldwide 2023," May 2024. [Online]. Available: https://www.statista.com/statistics/271644/worldwide-free-and-paid-mobile-app-store-downloads/

[2] M. Lutaaya, "Rethinking app permissions on ios," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–6.

[3] I. Reyes, P. Wijesekera, J. Reardon, A. Elazari Bar On, A. Razaghpanah, N. Vallina-Rodriguez, S. Egelman *et al.*, ""won't somebody think of the children?" examining coppa compliance at scale," in *The 18th Privacy Enhancing Technologies Symposium (PETS 2018)*, 2018.

[4] M. Meyer, V. Adkins, N. Yuan, H. M. Weeks, Y.-J. Chang, and J. Radesky, "Advertising in young children's apps: A content analysis," *Journal of developmental & behavioral pediatrics*, vol. 40, no. 1, pp. 32–39, 2019.

[5] Y. Zhao, T. Liu, H. Wang, Y. Liu, J. Grundy, and L. Li, "Are Mobile Advertisements in Compliance with App's Age Group?" in *Proceedings of the ACM Web Conference 2023*. Austin TX USA: ACM, Apr. 2023, pp. 3132–3141.

[6] Q. Luo, J. Liu, J. Wang, Y. Tan, Y. Cao, and N. Kato, "Automatic Content Inspection and Forensics for Children Android Apps," vol. 7, no. 8, Aug. 2020, pp. 7123–7134, conference Name: IEEE Internet of Things Journal.

[7] R. Sun, M. Xue, G. Tyson, S. Wang, S. Camtepe, and S. Nepal, "Not Seen, Not Heard in the Digital World! Measuring Privacy Practices in Children's Apps," in *Proceedings of the ACM Web Conference 2023*, ser. WWW '23. New York, NY, USA: Association for Computing Machinery, Apr. 2023, pp. 2166–2177.

[8] Google, "Manage target audience and app content settings," 2024. [Online]. Available: https://support.google.com/googleplay/android-developer/answer/9867159

[9] Apple, "Age ratings - Reference - App Store Connect," 2024. [Online]. Available: https://developer.apple.com/help/app-store-connect/reference/age-ratings/

[10] H. Liu, P. Patras, and D. J. Leith, "On the data privacy practices of Android OEMs," *PLOS ONE*, vol. 18, no. 1, p. e0279942, Jan. 2023, publisher: Public Library of Science.

[11] Y. Lin, J. Juneja, E. Birrell, and L. F. Cranor, "Data Safety vs. App Privacy: Comparing the Usability of Android and iOS Privacy Labels," Jan. 2024, arXiv:2312.03918 [cs].

[12] Q. Luo, Y. Yu, J. Liu, and A. Benslimane, "Automatic Detection for Privacy Violations in Android Applications," *IEEE Internet of Things Journal*, vol. 9, no. 8, pp. 6159–6172, Apr. 2022, conference Name: IEEE Internet of Things Journal.

[13] L. Jibb, E. Amoako, M. Heisey, L. Ren, and Q. Grundy, "Data handling practices and commercial features of apps related to children: A scoping review of content analyses," *Archives of disease in childhood*, vol. 107, no. 7, pp. 665–673, 2022.

[14] Pixalate, "90% of mobile apps are for kids aged 12 and under," Dec. 2020. [Online]. Available: https://www.pixalate.com/blog/google-apple-mobile-apps-for-kids

[15] G. L. Scoccia, A. Peruma, V. Pujols, I. Malavolta, and D. E. Krutz, "Permission Issues in Open-Source Android Apps: An Exploratory Study," in *2019 19th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. Cleveland, OH, USA: IEEE, Sep. 2019, pp. 238–249.

[16] I. M. Almomani and A. A. Khayer, "A Comprehensive Analysis of the Android Permissions System," *IEEE Access*, vol. 8, pp. 216 671–216 688, 2020, conference Name: IEEE Access.

[17] H. Bagheri, E. Kang, S. Malek, and D. Jackson, "A formal approach for detection of security flaws in the android permission system," *Formal Aspects of Computing*, vol. 30, no. 5, pp. 525–544, Sep. 2018.

[18] F.-H. Hsu, N.-C. Liu, Y.-L. Hwang, C.-H. Liu, C.-S. Wang, and C.-Y. Chen, "Dpc: A dynamic permission control mechanism for android third-party libraries," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 4, pp. 1751–1761, 2019.

[19] B. Brumen, A. Zajc, and L. Bošnjak, "Permissions vs. privacy policies of apps in google play store and apple app store," in *Information Modelling and Knowledge Bases XXXIV*. IOS Press, 2023, pp. 258–275.

[20] M. S. Rahman, P. Naghavi, B. Kojusner, S. Afroz, B. Williams, S. Rampazzi, and V. Bindschaedler, "Permpress: Machine learning-based pipeline to evaluate permissions in app privacy policies," *IEEE Access*, vol. 10, pp. 89 248–89 269, 2022.

[21] R. J. Majethiya and M. Shah, "Comparative analysis of detecting over-claim permissions from android apps," in *2023 International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*, Feb. 2023, pp. 1–8.

[22] S. Kununka, N. Mehandjiev, and P. Sampaio, "A comparative study of android and ios mobile applications' data handling practices versus compliance to privacy policy," *Privacy and Identity Management. The Smart Revolution: 12th IFIP WG 9.2, 9.5, 9.6/11.7, 11.6/SIG 9.2. 2 International Summer School, Ispra, Italy, September 4-8, 2017, Revised Selected Papers 12*, pp. 301–313, 2018.

[23] L. Verderame, D. Caputo, A. Romdhana, and A. Merlo, "On the (Un)Reliability of Privacy Policies in Android Apps," in *2020 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2020, pp. 1–9, iSSN: 2161-4407.

[24] L. Ceci, "Annual mobile app downloads worldwide by store 2026," Jun. 2022. [Online]. Available: https://www.statista.com/statistics/1010716/apple-app-store-google-play-app-downloads-forecast/

[25] Data.ai, "Mobile Market Data and App Insights | data.ai (fka App Annie)," 2024. [Online]. Available: https://www.data.ai/en/

[26] M. R. Bhutto, "Automating privacy policy extraction and summarization," B.S. thesis, University of Twente, 2022.