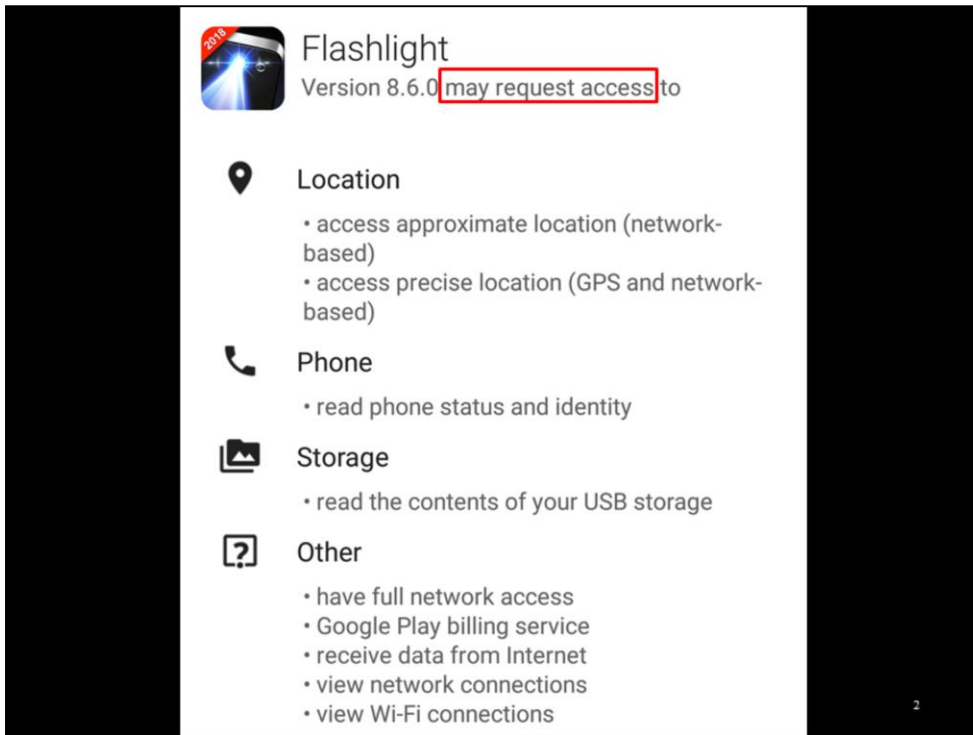


“Won’t Somebody Think of the Children?” Examining COPPA Compliance at Scale

Irwin Reyes, Primal Wijesekera, Joel Reardon, Amit Elazari Bar On,
Abbas Razaghpanah, Narseo Vallina-Rodriguez, Serge Egelman



Privacy analysis at scale



Lots of sensitive info and device resources available to apps.

Disclosures: Not clear if and when permissions are used, and if so, who gets that info.


dynamic analysis platform to observe
how apps **actually access and share** data

3


Solution: dynamic analysis.

Apps run as-is. No need to examine them like static analysis. All actual empirical observations. No false positives.

custom android for logging api calls



lumen app for network flow analysis



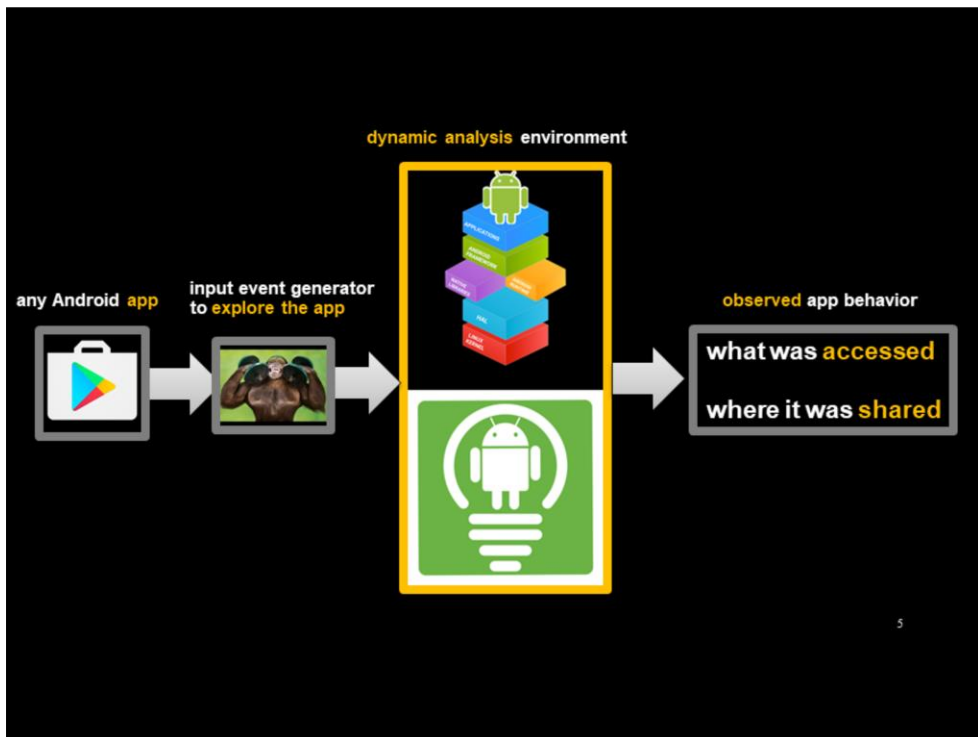
P. Wijesekera, A. Baokar, L. Tsai, J. Reardon, S. Egelman, D. Wagner, K. Beznosov, *The Feasibility of Dynamically Granted Permissions: Aligning Mobile Privacy with User Preferences*, IEEE Security and Privacy (Oakland) 2017

A. Razaghpanah, R. Nithyanand, N. Vallina Rodriguez, Srikanth Sundaresan, M. Allman, C. Kreibich, P. Gill, *Apps, Trackers, Privacy, and Regulators: A Global Study of the Mobile Tracking Ecosystem*, Network and Distributed System Security (NDSS) 2018

4

Custom Android 6 ROM for observing access to sensitive resources.

Lumen Privacy Monitor to see who gets that info.



We run any Android app in this environment and observe its behavior.

Not enough to just launch the app. Solution: explore with monkey. It's dumb!

Monkey did as well as undergrads 60% of the time in children's games. Results are a lower bound.

current deployment runs **1,000 apps/day**

6

We deployed this environment onto a cluster of physical smartphones, running 24/7.

PERSONAL INFORMATION	PERSISTENT IDENTIFIERS
Owner Email Address	Hardware Serial Number
Phone Number	IMEI
GPS Latitude/Longitude	Wi-Fi MAC
Wi-Fi Router BSSID (MAC)	Android ID
Wi-Fi Router SSID (Name)	SIM Card ID
	Google Services Framework (GSF) ID
	Android Advertising ID (AAID)

7

The platform can detect different kinds of personal information and persistent identifiers.



As a case study COPPA: one of the few comprehensive US privacy laws.
Applies to online services (e.g., apps) used by children under 13.

Prohibits collecting contact info and location.

No building profiles of children over time across different services---enabled by persistent identifiers.

Need parental consent for data collection. Consent like credit card verification or phone calls.

Protect the security and privacy of end-users.

Violations are costly.

The image is a screenshot of the Federal Trade Commission (FTC) website. At the top left is the FTC logo, which features a shield with scales of justice and a sword, surrounded by the text "FEDERAL TRADE COMMISSION" and "UNITED STATES OF AMERICA". To the right of the logo, the text "FEDERAL TRADE COMMISSION" and "PROTECTING AMERICA'S CONSUMERS" is displayed in white and blue. In the top right corner, there are links for "Contact" and "Stay Cor". Below the header is a navigation menu with five items: "ABOUT THE FTC", "NEWS & EVENTS", "ENFORCEMENT", "POLICY", and "TIPS & ADVICE". The main content area shows a breadcrumb trail: "Home » News & Events » Press Releases » Two App Developers Settle FTC Charges They Violated Children's Online Privacy". Below this is the title of the press release: "Two App Developers Settle FTC Charges They Violated Children's Online Privacy Protection Act". The sub-headline reads: "Companies' Apps **Shared Kids' Information with Ad Networks** Will Pay \$360K In Civil Penalties". The text "Shared Kids' Information with Ad Networks" is highlighted with a red rectangular box. At the bottom right of the page, there is a small number "9".

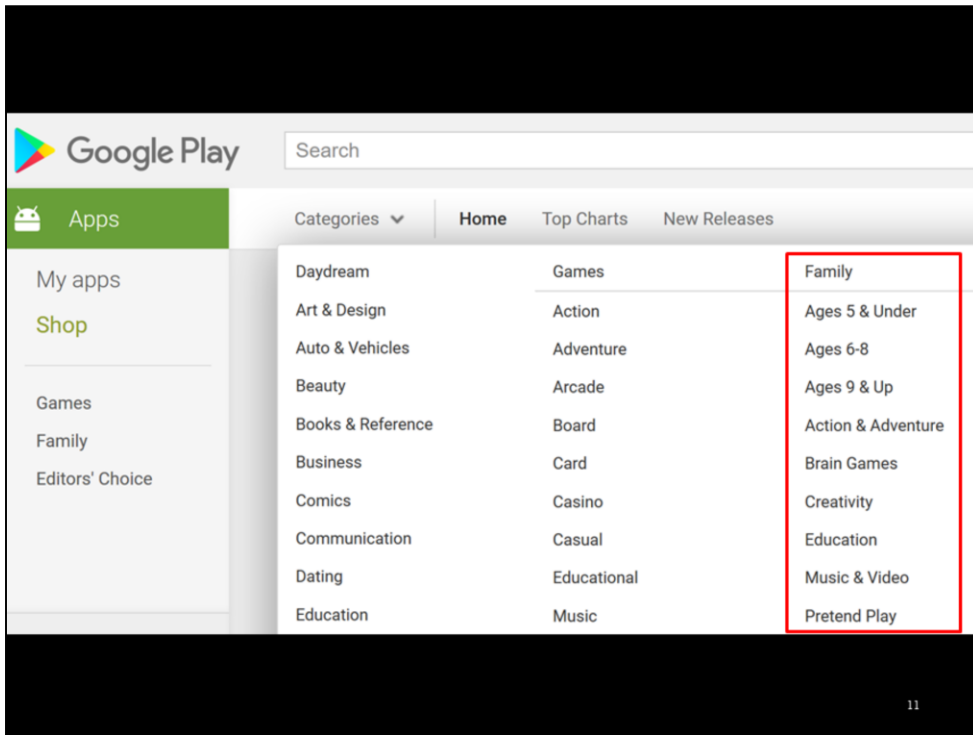
US Federal Trade Commission enforces.

In 2015, \$360K fine for app devs LAI Systems and Retro Dreamer. Persistent identifiers to advertisers.

The image is a screenshot of the Federal Trade Commission (FTC) website. At the top left is the FTC logo, which features a shield with scales of justice and a sword, surrounded by the text "FEDERAL TRADE COMMISSION" and "UNITED STATES OF AMERICA". To the right of the logo, the text "FEDERAL TRADE COMMISSION" is displayed in large, bold, white letters, with "PROTECTING AMERICA'S CONSUMERS" in smaller, blue letters below it. In the top right corner, there are links for "Contact" and "Stay Con". Below the header is a navigation menu with five items: "ABOUT THE FTC", "NEWS & EVENTS", "ENFORCEMENT", "POLICY", and "TIPS & ADVICE". The main content area shows a breadcrumb trail: "Home » News & Events » Press Releases » Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of M Permission". The headline of the press release reads: "Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of Millions of Consumers' Locations Without Permission". Below the headline, a sub-headline states: "Company Will Pay \$950,000 For Tracking Children Without Parental Consent", where the phrase "Tracking Children Without Parental Consent" is enclosed in a red rectangular box. In the bottom right corner of the page, the number "10" is visible.

Third-party services can be liable too.

inMobi handed \$1M fine for collecting location data from children.



So we have this system that allows us to identify potential violations of this law. How do we find COPPA apps?

Starting in late 2016, scraped the Play Store's Top Charts in the Family Friendly categories; like "Ages 6-8" and "Pretend Play"



Those are apps that have opted into the Designed for Families Program, or DFF for short.

DFF is opt-in. Participation is the dev saying kids are in the target audience. Google can reject or remove DFF apps not relevant to children.

DFF's requires devs to represent their apps ****and bundled SDKs**** are COPPA compliant. For example, SDKs for graphics, communications, analytics, and ads.

5,855 free “**Designed for Families**” apps





13

From November 2016 to March 18, crawled the Play Store. Found:

- Over 5,800 free DFF apps
- 750K installs each on average
- Represents nearly 1900 devs

We tested them...

57% of “Designed for Families” apps are in potential violation

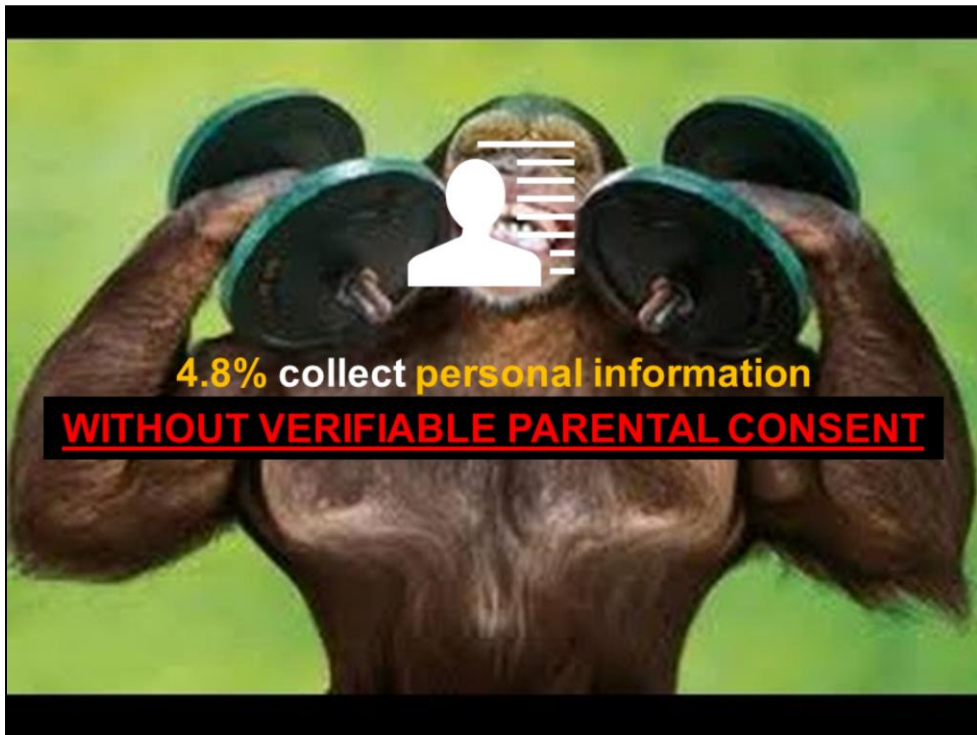
POTENTIAL VIOLATION	RATE (n=5,855)
→  Personal information	4.8%
→  Non-resettable identifiers	39%
→  Potentially non-compliant SDKs	19%
→  Failure to take security measures	40%

14

The majority of our corpus was seen to be in potential violation of COPPA, in that they:

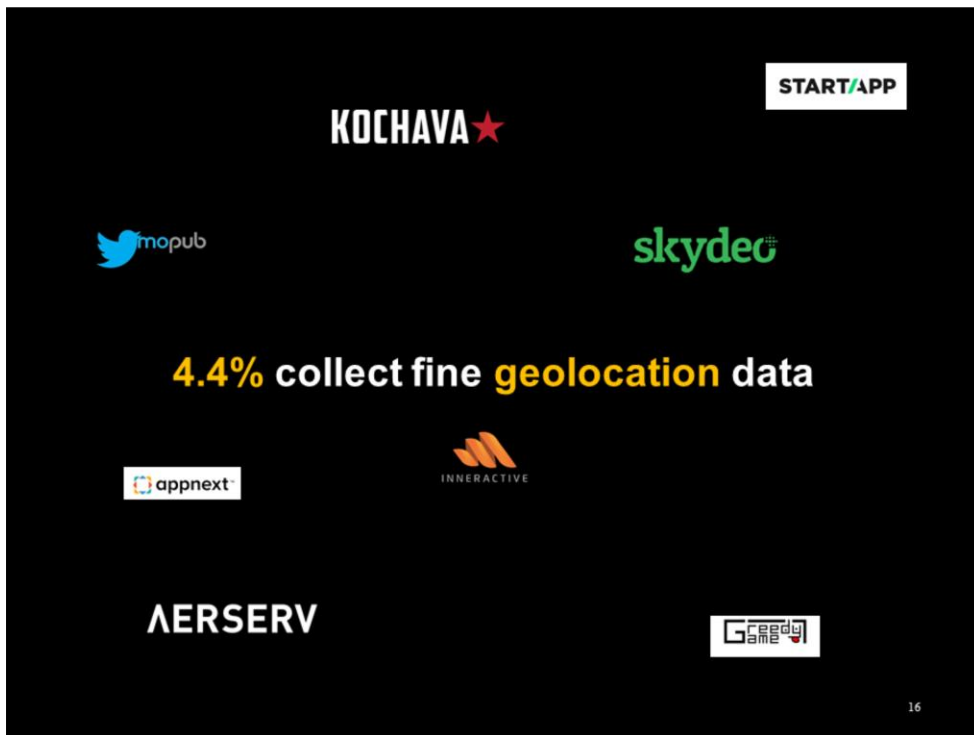
- Accessing and collecting email addresses, phone numbers, and fine geolocation
- Potentially enabling behavioral advertising through persistent identifiers
- Sharing user data and identifiers with SDKs that are themselves potentially non-compliant
- Not using standard security technologies

Note that some apps were observed engaging in more than one of these behaviors, so the percentages will add up to more than 57%.



We observed 282 DFF apps collecting and sharing personal data. Our system can identify when fine geolocation data and contact information are accessed and shared.

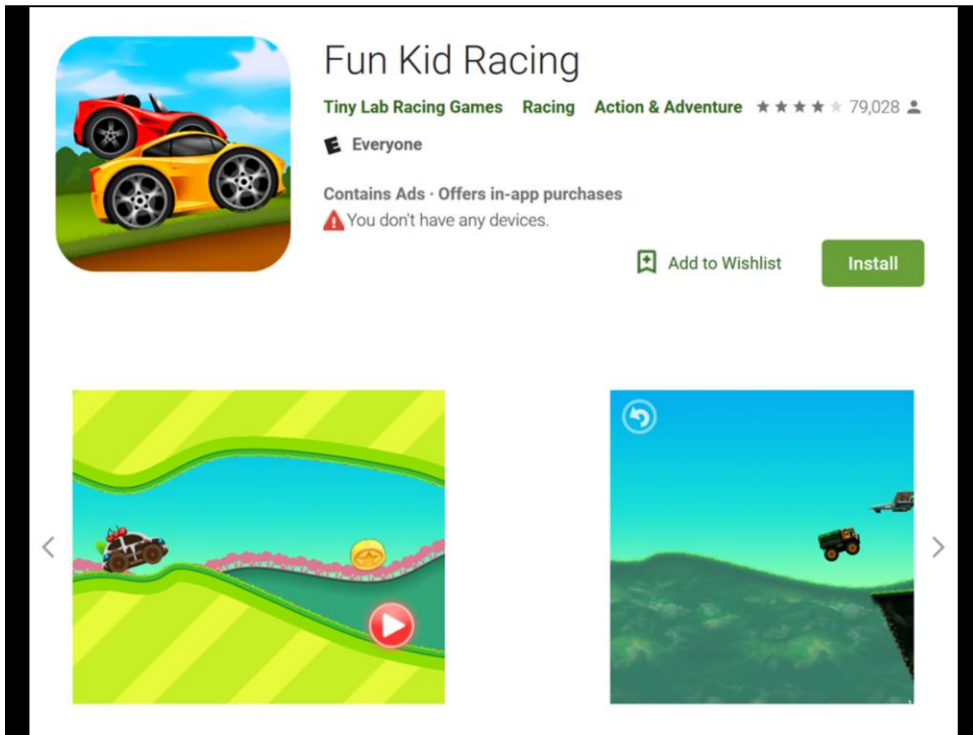
Recall that we're using a dumb exerciser monkey to drive these apps; what it does cannot constitute verifiable parental consent. Also, if the monkey can cause the results blindly, then so can a child.



We observed DFF apps collecting location data accurate enough to identify the device's city and street..

We looked at the collection and sharing of fine GPS coordinates, and found that the top domains receiving this data from DFF apps belonged to ad networks.

We also looked at the collection and sharing of wi-fi router identifiers and names, which can be used to infer location with high accuracy. The top domains receiving wi-fi router data also belonged to ad networks.



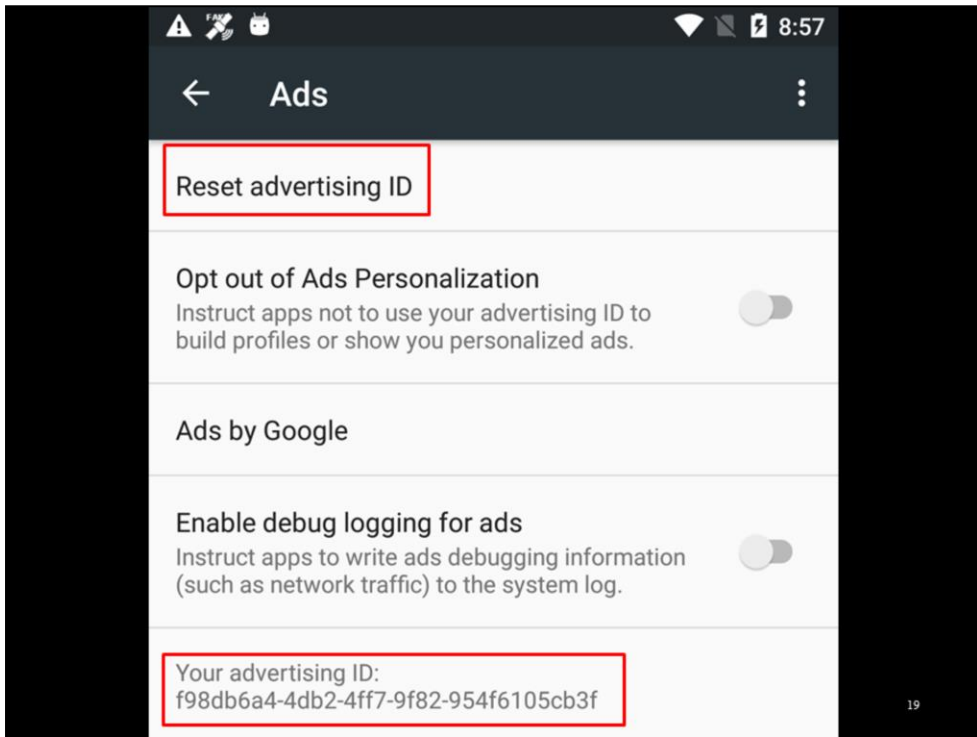
Popular apps were among those observed collecting and sharing geolocation data with advertisers. The game Fun Kid Racing has over 10M installs, and was seen accessing and sharing fine GPS coordinates. This behavior was seen in 81 of 82 of this developer's DFF apps.

In response to our results, the developer stated to CNET that their games aren't specifically for kids.



COPPA prohibits the collection of contact information as well. We were able to identify over 100 apps that accessed the device-registered email address or the device's phone number, or both.

This data most often went to various developer services, as well as ad networks and app recommendation services.



Beyond personal information, COPPA prohibits behavioral advertising for children. Behavioral advertising relies on persistent identifiers to build profiles of users by tracking individuals across different services over time.






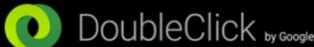
Google recognizes the privacy implications of persistent identifiers, and in 2014 introduced the resettable Android Advertising ID (AAID) to give users control over how advertisers track them. Google requires developers and advertisers to use this in lieu of non-resettable device identifiers like the IMEI and Wi-Fi MAC address.



**39% share the AAID along another identifier,
negating its privacy preserving benefits**

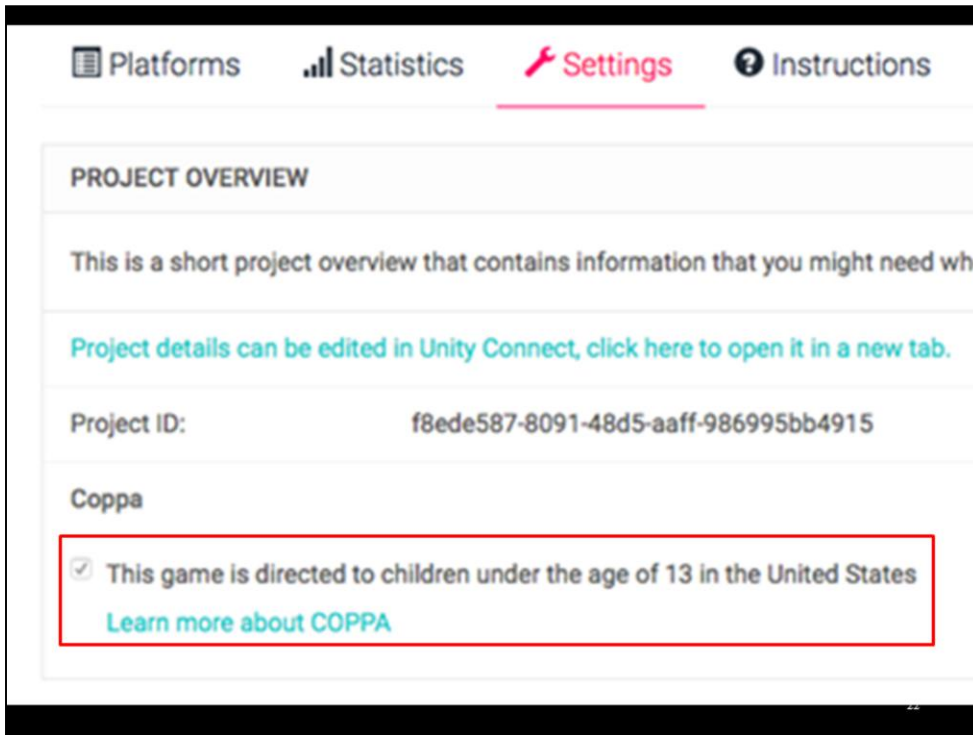
20

However, a large chunk of DFF apps were seen sharing the AAID with another non-resettable identifier to the same destination, which defeats the purpose of the AAID.

AD PLATFORM	VIOLATION OF IDENTIFIER POLICY
	> 99%
	> 99%
	98%
...	...
	3%
	2%
	1%

We found adherence to this AAID-only policy to vary among ad networks themselves. From nearly constant violation with Chartboost to nearly full compliance with Doubleclick (which is a Google company).

Full table in paper.



As noted before, it's not just app developers that are subject to COPPA. The FTC has pursued enforcement actions against third-party SDKs. Some third-party SDKs attempt to comply with COPPA by allowing app developers to specify that the end product is directed at children, and so the SDK will adjust their data access and collection behaviors accordingly.

In some cases, we're able to observe these options be passed between the app and the SDK's servers.

50% **used Unity** (from DFF corpus of 5,855)

84% of Unity apps did **NOT** get `coppaCompliant=true`

23

For example, nearly half of our corpus used Unity, which offers a COPPA option.

However, this option was not set consistently among DFF apps. 84% of Unity apps did not receive an explicit "`coppaCompliant=true`," suggesting that they're potentially operating in a non-compliant mode.

not for children's apps



24

There are third-party SDKs that don't even offer COPPA options at all.



Developer further agrees it will not integrate the Software into any Application or Beta Application (i) with end users who Developer has actual knowledge are under the age of 13, or (ii) that may be deemed to be a “Web site or online service directed to children” as defined under the Children’s Online Privacy Protection Act of 1998 (“COPPA”) and the regulations promulgated thereunder.

25








Those SDKs instead have terms of service with explicit language prohibiting their use in children's apps. Presumably, this is because these services collect and process user data in ways prohibited by COPPA, so the services prefer if developers of children's apps didn't use them.



19% share identifiers or personal information
with **SDKs not allowed in children's apps**

26

However, we found nearly 1 in 5 DFF apps sharing personal information or identifiers with a number of these "verboden" SDKs. Recall that DFF is an opt-in program; developers go out of their way to join this program and signal that their app is meant for users under 13, among others. Developers intend for children under 13 to be in their audience.

SDK	TOTAL DFF INSTALLS
 crashlytics	556M
 <i>supersonic</i> +  ironSource	481M
	386M
 mopub	296M
 INNERACTIVE	239M
 heyZap	150M

Still, "verboden" SDKs can be found in many self-declared DFF apps, accounting for hundreds of millions of installations in aggregate.



**40% share identifiers and personal info
without using encrypted HTTP**

28

We've quantified how apps collect and share sensitive data---often through third-party SDKs. When sharing data, COPPA also requires apps to take reasonable security measures to protect end-users. For our study, we interpret that as something as basic as using encrypted HTTP.

We found 40% DFF apps transmitting potentially sensitive information to remote services without using encrypted HTTP as a basic security measure.







Overall, **57%** of “Designed for Families” apps
are in potential violation

29

Again, between the collection of personal information without verifiable parental consent, the use of persistent identifiers even when resettable ones are available, integration with potentially non-COPPA-compliant third-party SDKs, and failure to implement basic security measures, we find a majority of free apps in the Designed for Families program is in potential violation of COPPA.



Potential COPPA violations are widespread, but unfortunately regulatory agencies like the FTC have finite enforcement capability. COPPA, however, allows for industry self-regulation in the form of review and certification from designated safe harbor industry groups.

	DFF (n=5,855)	SAFE HARBOR (n=237)
 SHARE PERSONAL INFO	4.8%	10%
 SHARE AAID + ANOTHER ID	39%	39%
 USE VERBOTEN SDK	19%	33%
 UNENCRYPTED HTTP	40%	49%

We scoured those safe harbors' websites to identify which apps and developers they've certified. In aggregate across the 7 safe harbors, we found that safe harbor apps were not appreciably any better than DFF apps as a whole.

```
SELECT '(AAID) Transmit AAID + another ID: ', COUNT(DISTINCT qry.pkg)
FROM
```

```
(SELECT apps.packageName AS pkg,appReleases.versionCode AS
vers,testTransmissions.ipAddress AS ip,COUNT(testTransmissions.dataType)
AS identifiers,(testTransmissions.dataType='aaid') AS hasAaid FROM
appReleases
```

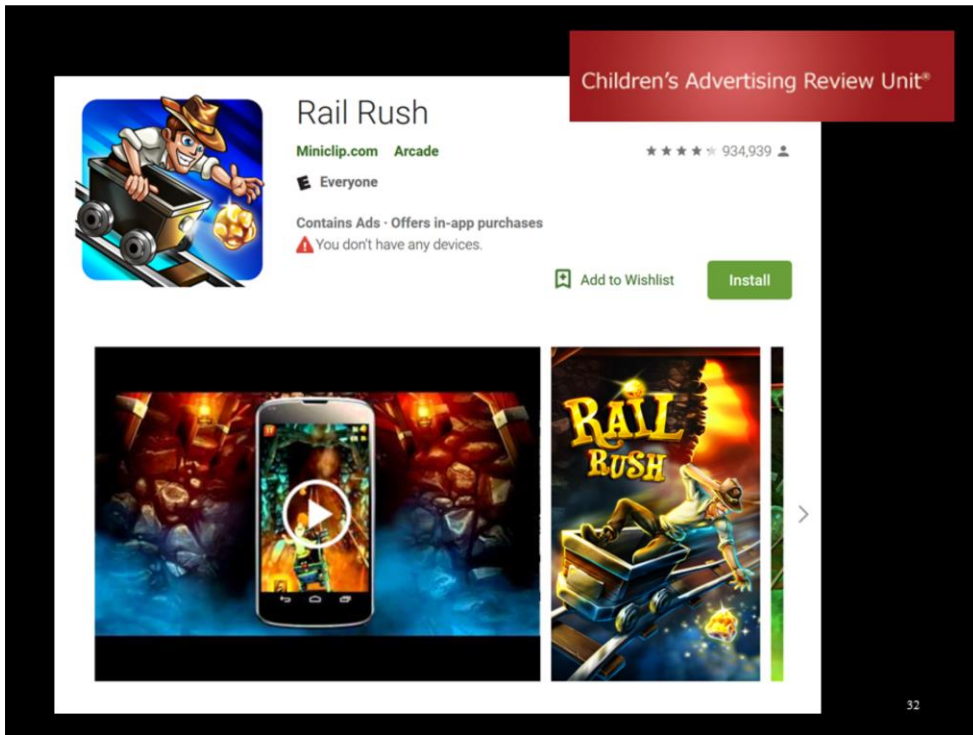
```
INNER JOIN apps ON apps.id=appReleases.appId AND apps.packageName
IN (SELECT safeHarbor.packageName FROM safeHarbor)
```

```
INNER JOIN testTransmissions ON
testTransmissions.releaseId=appReleases.id AND appReleases.id
```

```
AND testTransmissions.dataType IN
('aaid','androidid','hwid','wifimac','imei','simid','imsi','gsfid')
```

```
GROUP BY testTransmissions.releaseId,testTransmissions.ipAddress
```

```
HAVING identifiers >= 2 AND hasAaid=1  
ORDER BY identifiers DESC)  
AS qry;
```

For example, CARU reviewed Rail Rush, which has over 50M installs. We observed Rail Rush not only collecting location data without verifiable parental consent, but also sharing that data with Amplitude, whose terms prohibit its use in children's apps.

closing recommendations

developers: use compliant SDKs and options

SDK providers: enforce terms of use

platform providers: stricter security and analysis

33

Given all these results, what can be done? We offer recommendations to the stakeholders in the mobile app ecosystem.

First, developers need to take care when integrating third-party SDKs into their products. This means selecting COPPA compliance options where available, and avoiding SDKs whose terms prohibit their use in children's apps.

The other side of that equation is that SDK providers need to identify when their partner developers are violating terms of use, specifically terms that prohibit integration with children's apps. Behavioral advertising networks, for example, can pressure developers by freezing payments to partner developers who make children's apps.

As gatekeepers to the mobile app ecosystem, companies such as Google can do more to improve compliance. For example, stricter restrictions for apps to access personal information, empowering users with upgraded permissions systems, and integrating our methods into existing pre-release security and malware scans in app stores.

Crashlytics

From Wikipedia, the free encyclopedia

Crashlytics is a Google-owned [Boston, Massachusetts](#)-based software company founded in May 2011 by entrepreneurs [Wayne Chang](#) and [Jeff Seibert](#).

34

As a side note, we identified Crashlytics as an SDK whose terms prohibit its use in children's apps. Google owns Crashlytics, Android, and the Play Store. Google should be able to detect when its own service is integrated with children's apps, then take necessary steps to address that.

closing recommendations

developers: use compliant SDKs and options

SDK providers: enforce terms of use

platform providers: stricter security and analysis

<https://appcensus.mobi>

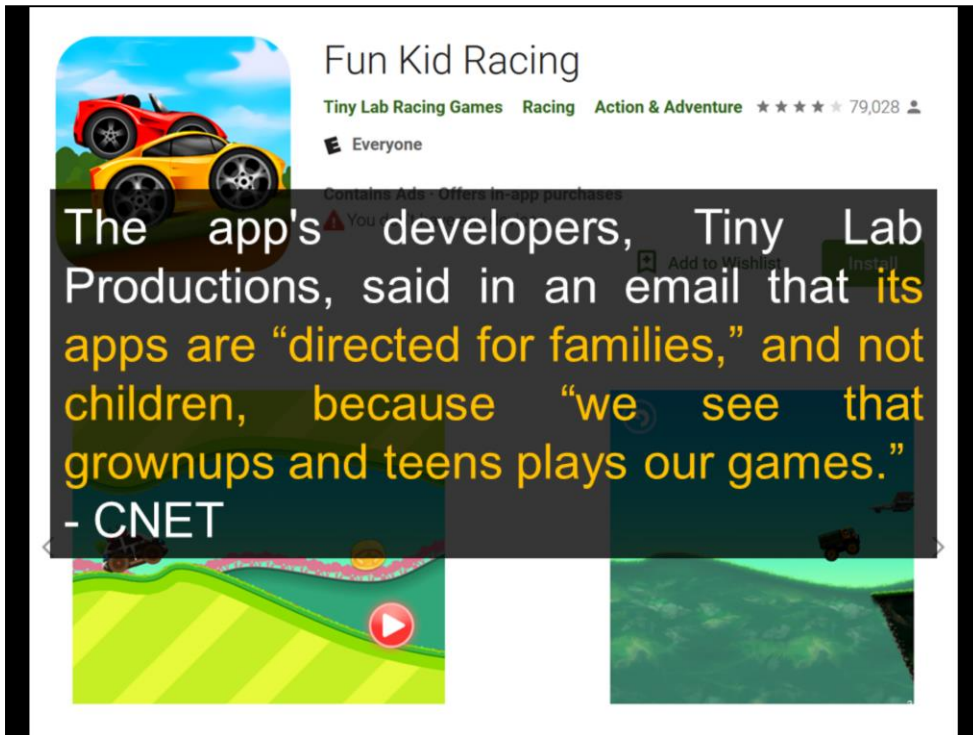
<https://blog.appcensus.mobi>

35

Finally, for regulators, researchers, and the public at-large, we make all our results and newest findings available online. Our results offer a continuously-updated birds-eye view of data collection in the mobile app marketplace.

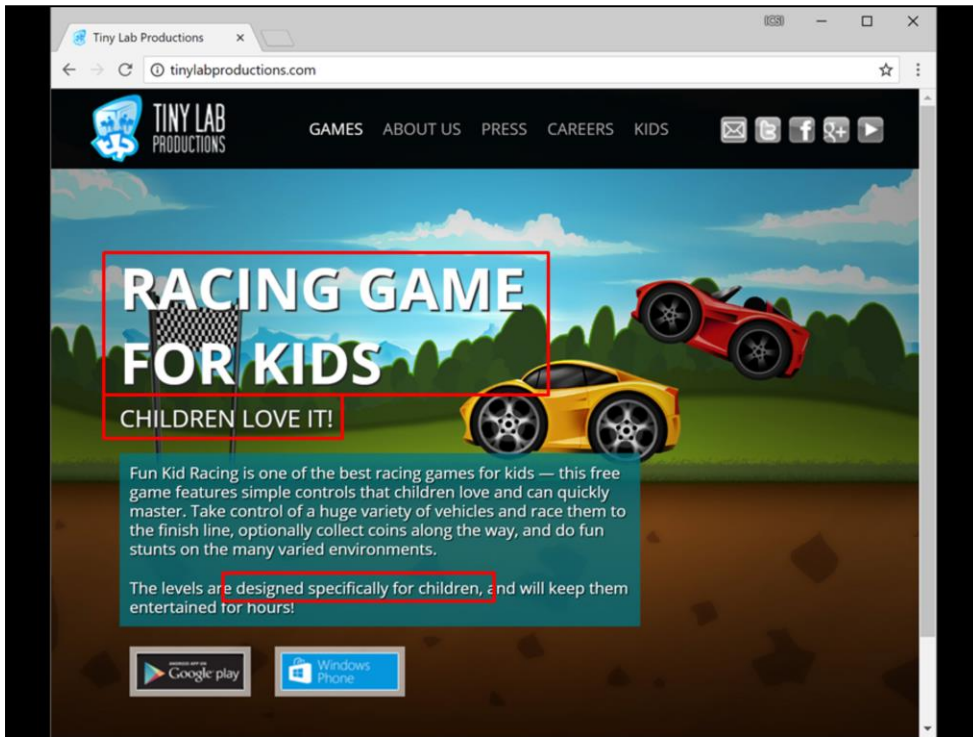
Ultimately though, we believe that our results reveal an opportunity for the other stakeholders---developers, third-party SDK providers, and platform providers---to step up and address what's truly a systemic privacy issue in children's apps.

BACKUP

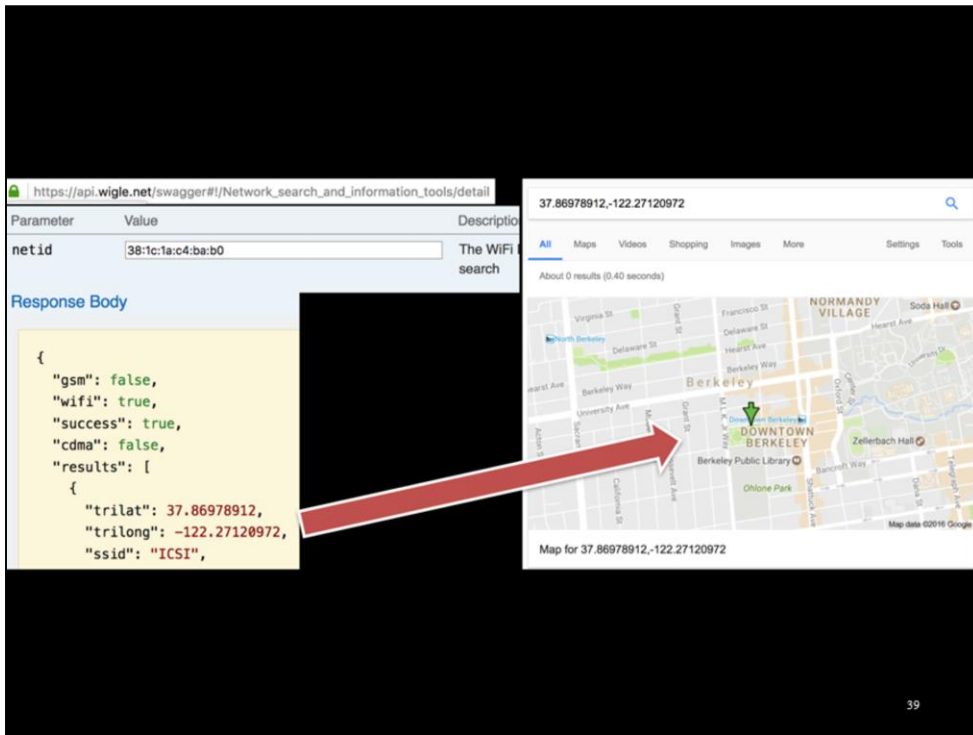


Popular apps were among those observed collecting and sharing geolocation data with advertisers. The game Fun Kid Racing has over 10M installs, and was seen accessing and sharing fine GPS coordinates. This behavior was seen in 81 of 82 of this developer's DFF apps.

In response to our results, the developer stated to CNET that their games aren't specifically for kids.



Their website might cast some doubt on that one...



Geolocation data isn't just limited to GPS coordinates. Information about the currently-connected Wi-Fi router can also be used to deduce location with high accuracy. Wi-fi routers tend to stay in place, and there are geocoding services like wigle.net and the Google Maps Geolocation API that allow lookups of known wi-fi routers' locations. Even wi-fi network names can leak location information.

DOMAIN	APPS SENDING IDs	APPS SENDING NON-AAID IDs	COMPLIANCE WITH GOOGLE POLICY
doubleclick.net	168	1	99%
lkqd.net	65	1	98%
mopub.com	148	3	97%
...
adcolony.com	557	108	80%
supersonicads.com	465	144	69%
tapjoy.com	98	96	2%
tapjoyads.com	95	94	1%
chartboost.com	859	858	< 1%
greedygame.com	59	59	< 1%

40

(full table in paper)

Compliance with Google’s AAID policy varied heavily based on the ad network involved.

Among ad networks observed in at least 50 apps, most complied with the AAID policy reliably: ranging from 69% with Supersonic (now ironSource) to 99% with Doubleclick (a Google company).

However, some such as the widely-installed Chartboost almost always failed to comply.

ROUTER MAC SENT TO DOMAIN	APP COUNT
greedygame.com	61
startappservice.com	60
startappexchange.com	57
kochava.com	30
app-nxt.net	13

41

The most commonly observed destinations for Wi-Fi MAC were all advertising services.

2,909 apps **used Unity** (from DFF corpus of 5,855)

1,068 received **"coppaCompliant" flag** from Unity server

479 have coppaCompliant=**true**

589 have coppaCompliant=**false**

42

Nearly half of our corpus used Unity.

Among Unity apps, only a third received a "coppaCompliant" flag from the Unity config server.

This flag was not set consistently among DFF apps. 83% of Unity apps did not have an explicit "coppaCompliant=true," potentially operating in a non-compliant mode.

1,280 apps **integrated with Facebook** (from DFF corpus of 5,855)

444 sent **"coppa" flag** to Facebook server

75 have coppa=true

342 have coppa=false

27 have coppa=true and false (both!)

43

Other SDKs have COPPA compliance options client-side, where the developer sets the value in the implementation of the app.

These options are sometimes included in the outbound network traffic when the SDK communicates with its home server. This is the case with the Facebook social and ads SDK.

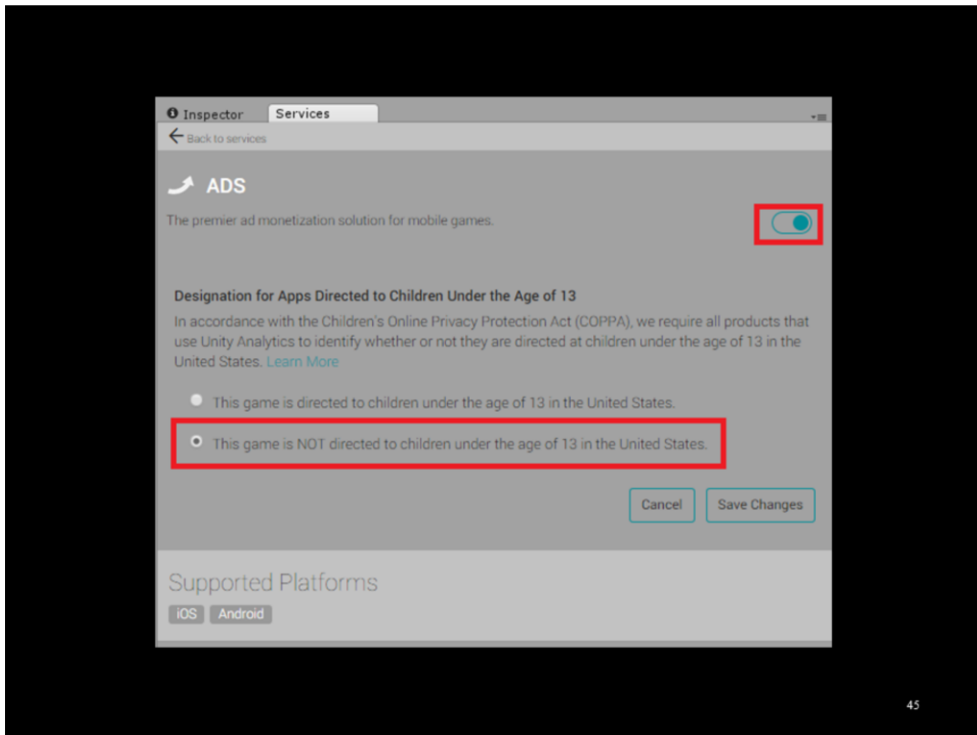
Like Unity, not all apps that integrate with Facebook have the coppa options set. And only a small number of apps consistently set this value to true.

*Supply Partners who sign up using this website may not provide MoPub with data from end users under age 13. Supply Partners **must not register for MoPub's services** using this website **if any of their apps are either: (1) directed to children under age 13** (even if children are not the app's primary audience), or (2) collect information from children that Supply Partners know are under age 13.*

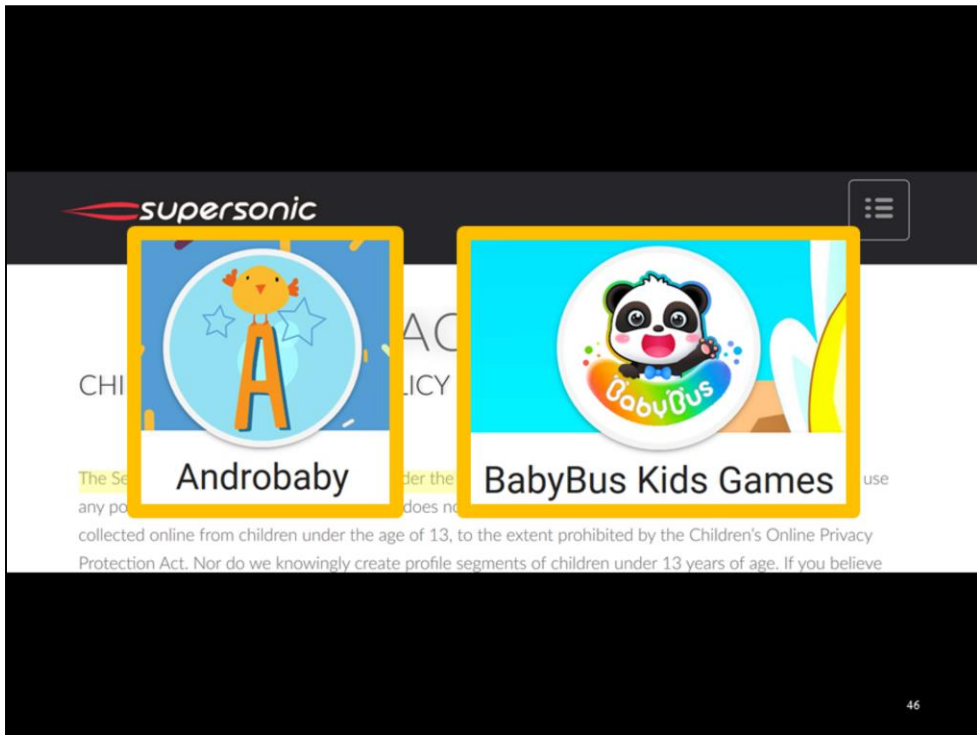
- MoPub Terms of Service

44

Those SDKs instead have terms of service with explicit language prohibiting their use in children's apps. Presumably, this is because these services are for behavioral advertising, or otherwise collect and process user data in ways prohibited by COPPA.

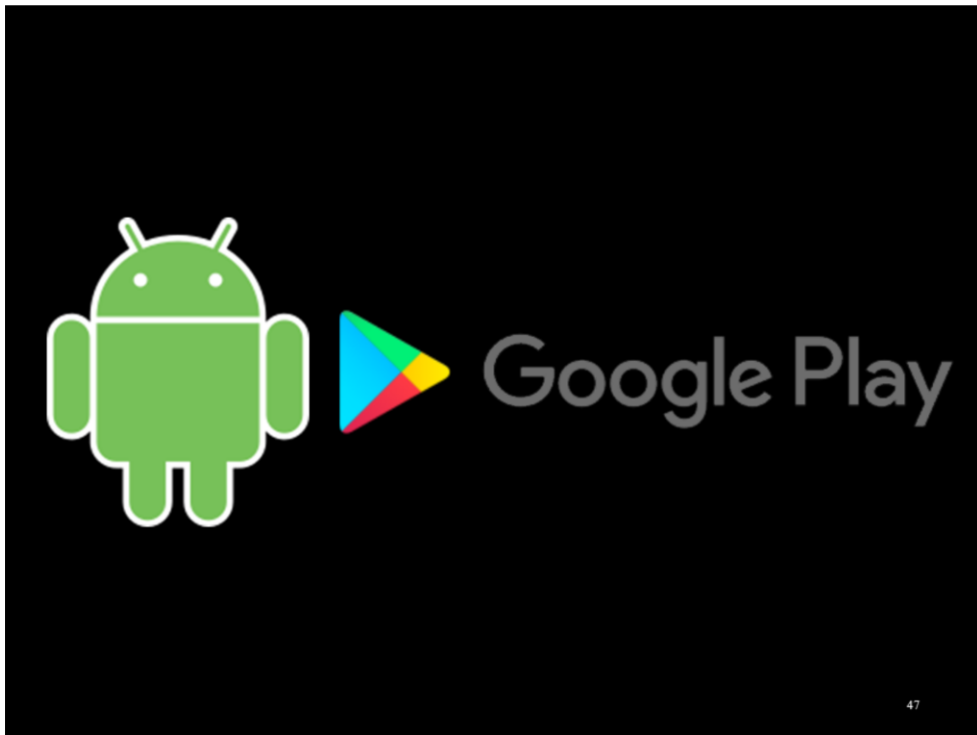


Developers have a responsibility to be aware of the data collection options that their third-party SDKs offer. We identified Unity and Facebook as two popular SDKs that have these options. However, our data suggests that only a small fraction of apps appear to have put these services into COPPA-compliant modes.

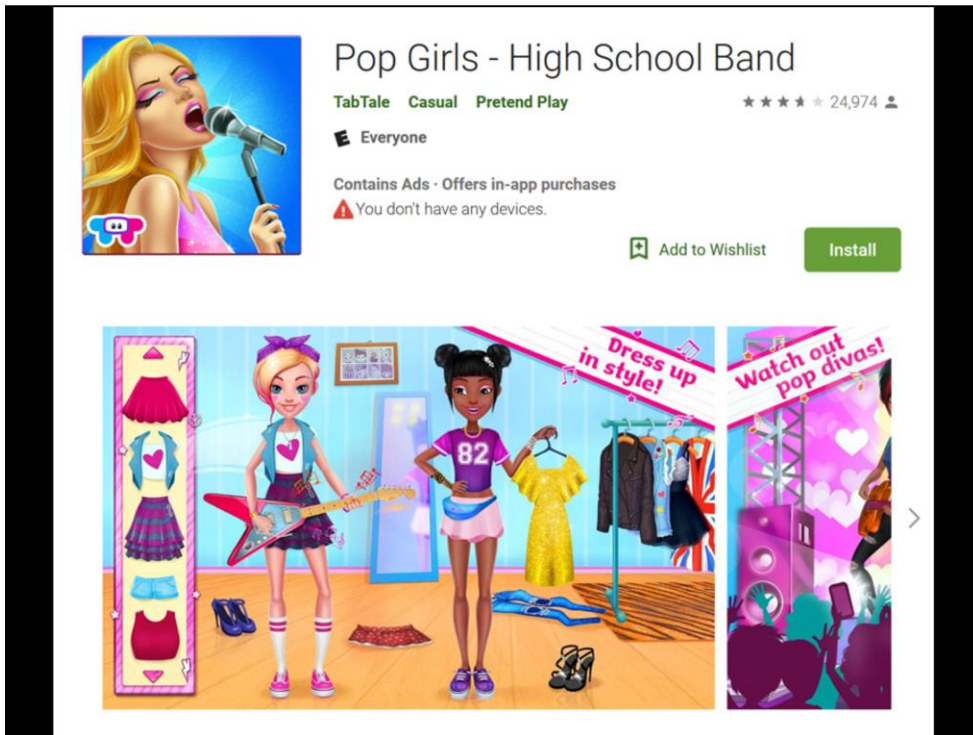


More fundamentally, developers need to know if the SDKs they integrate into children's apps are indeed appropriate for that audience. We observed nearly 1 in 5 DFF apps using SDKs that prohibit their use in apps directed at children.

On the other hand, the providers of those "verboten" SDKs have a responsibility to communicate these restrictions to developers and root out non-compliant ones. Shortly after the publication of this work, we received a legal letter from the ad tech company ironSource, which is SuperSonic's parent company. In our response, we informed them that their partnered developers---all of whom had to register with SuperSonic---included such organizations as "Androbaby" and "BabyBus Kids Games." Ad networks can improve COPPA compliance by terminating payments to developers that don't abide by ad networks' own terms of use.



Platform providers such as Google have a key role too. They develop and maintain the underlying OS, which determines how easily apps and bundled third-party services access private data on the device. Stronger platform-level data protection is needed, such as improved permissions models and tighter restrictions on accessing sensitive data. Google also operates Android's primary app distribution platform, the Play Store. Apps already undergo malware analysis upon submission to the Play Store. Our techniques should be integrated into the app security testing pipeline, which would allow developers to find and address privacy issues before apps are made public.



This included some popular children's apps, such as this TabTale game with over 1M installs.

It collected and shared the router's BSSID (MAC address) with the ad network StartApp.

	DFP APPS (n=5,855)	SAFE HARBOR APPS (n=237)
SEND IDENTIFIERS	73%	66%
SHARE PERSONAL DATA	4.8%	10%
USE VERBOTEN SDK	19%	33%
DON'T ENCRYPT COMMS	40%	49%

49

We scoured those safe harbors' websites to identify which apps and developers they've certified. In aggregate across the 7 safe harbors, we found that safe harbor apps were not appreciably any better than DFP apps as a whole.