

Examining Privacy Violations in Children's Apps

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



INTERNATIONAL
COMPUTER SCIENCE
INSTITUTE



Berkeley
UNIVERSITY OF CALIFORNIA



THE UNIVERSITY OF
BRITISH
COLUMBIA



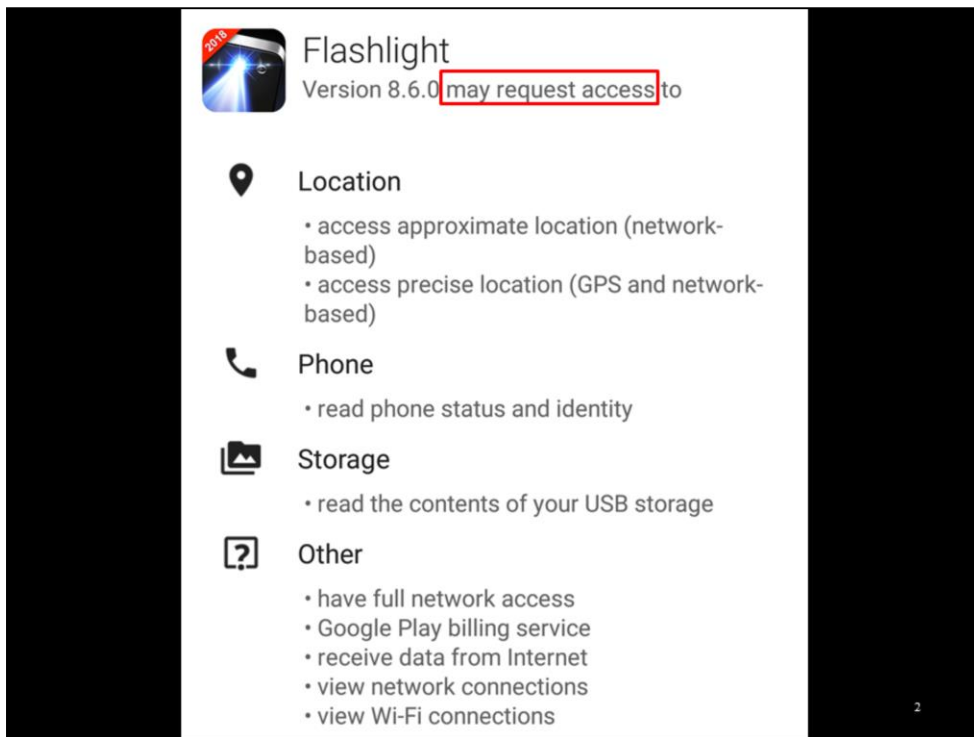
UNIVERSITY OF
CALGARY



Stony Brook
University



idea
networks



Apps are able to request access to private user data and sensitive device resources.

In their app store listings (such as this one from the Google Play Store), apps disclose their capabilities. However, these disclosures don't tell the full story. Do apps actually use these privileges? With whom do they share sensitive data?

automated run-time analysis to observe how apps **actually access and share** data

I. Reyes, P. Wijesekera, J. Reardon, A. Elazari Bar On, A. Razaghpanah, N. Vallina-Rodriguez, S. Egelman. *"Won't Somebody Think of the Children?" Examining COPPA Compliance at Scale*, Privacy Enhancing Technologies Symposium (PETS) 2018

Available at <https://appcensus.mobi/about>

3

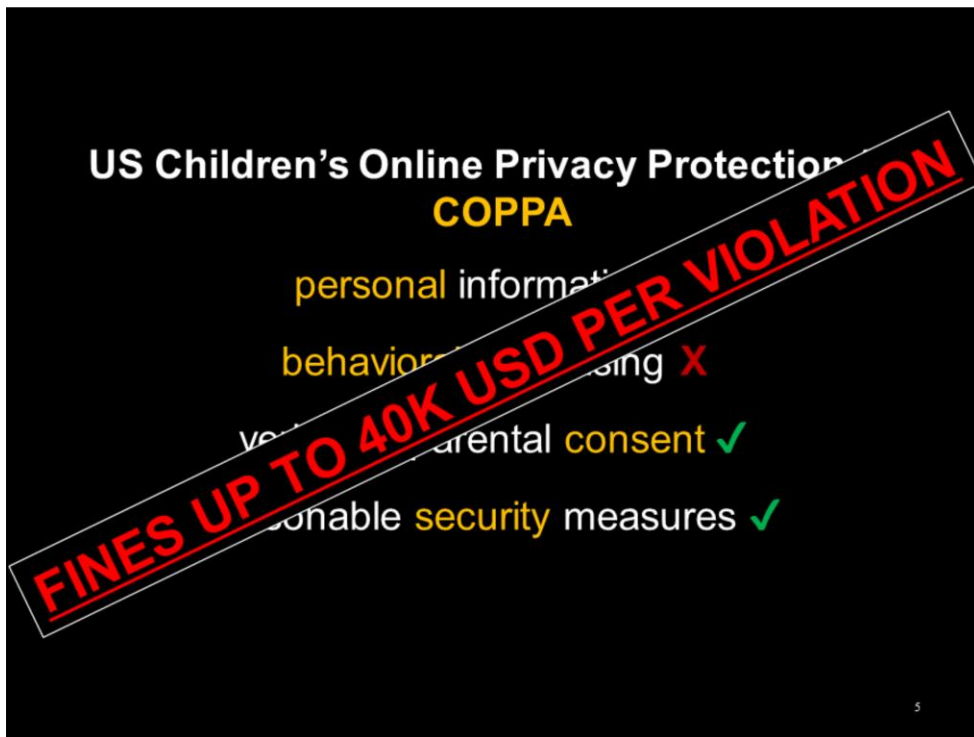
We developed a fully automated platform to analyze how apps actually collect and share sensitive data.

We instrumented the Android operating system and used advanced network traffic monitoring tools. Apps are run and evaluated without any human interaction. Technical details in the paper.

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)

4

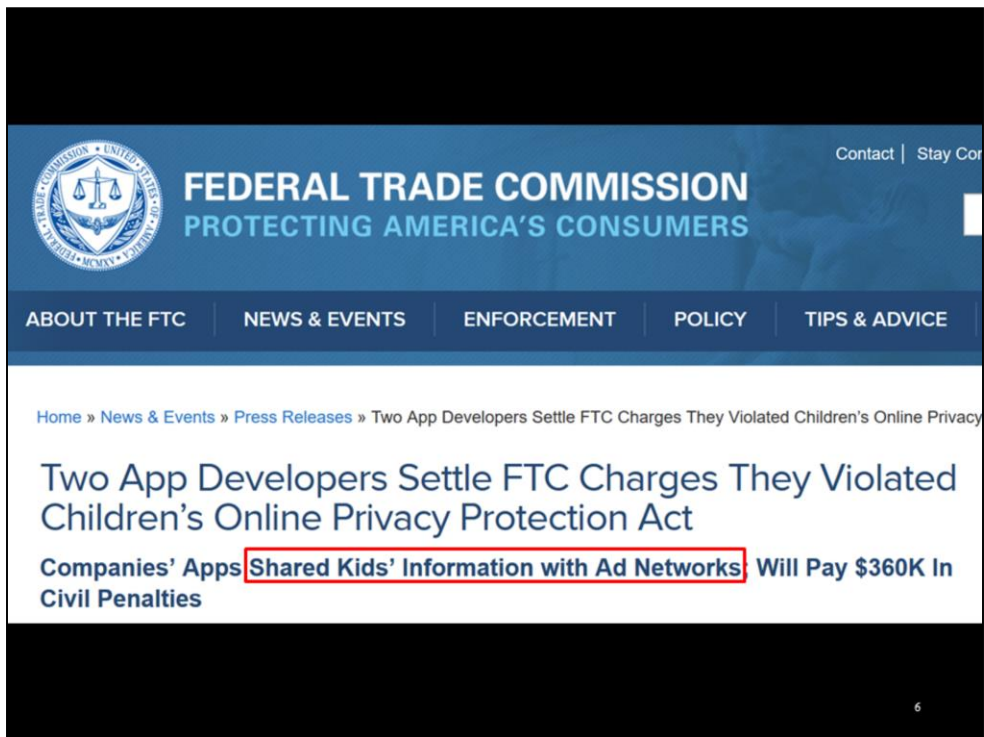
Our system observes when apps access and share personal information, as well as unique persistent identifiers that can be used to track users over time and across services.



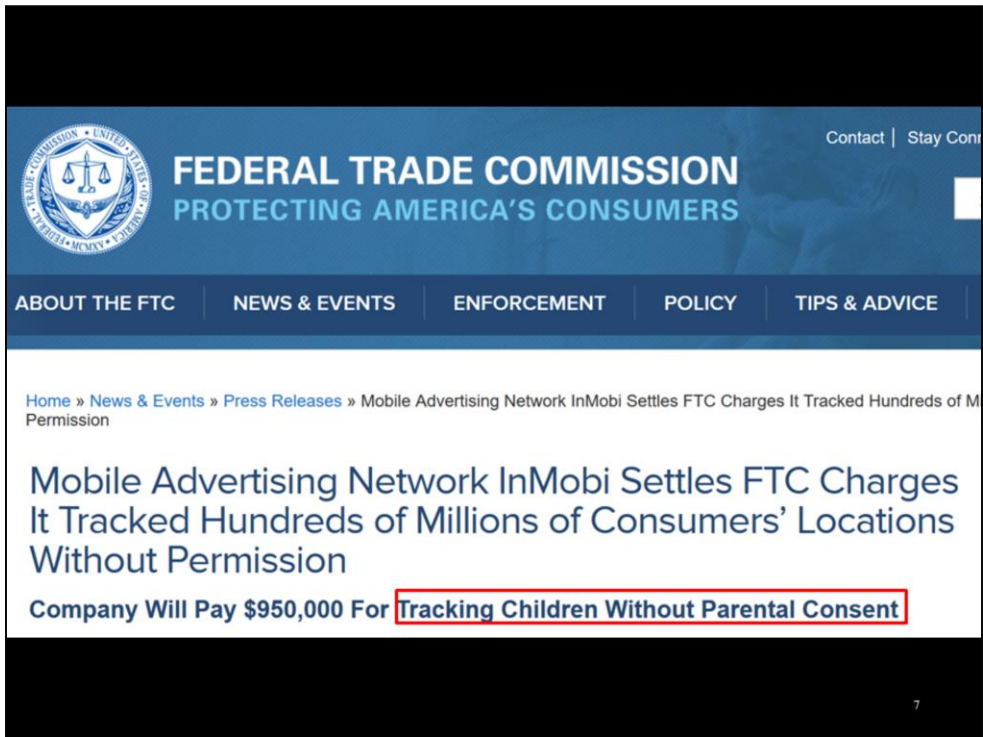
COPPA is one of the few comprehensive privacy laws in the US. It covers online services (like apps) that have users under 13 years of age.

Verifiable parental consent: Can take on the form of out-of-band methods like credit card verification or a phone call. Our system is fully automated with no direct human input, so observed data collection did not have consent.

Note that our analysis system is not specific to COPPA. It can be adapted to other regulatory measures such as GDPR and California's new online privacy law.



LAI Systems and Retro Dreamer in 2015. Sharing persistent identifiers from child users of apps to ad networks.



The image is a screenshot of the Federal Trade Commission (FTC) website. At the top, there is a blue header with the FTC logo on the left, which features a shield with scales of justice and the words "FEDERAL TRADE COMMISSION" and "UNITED STATES OF AMERICA". To the right of the logo, the text "FEDERAL TRADE COMMISSION" is written in large white letters, and "PROTECTING AMERICA'S CONSUMERS" is written in smaller white letters below it. In the top right corner of the header, there are links for "Contact" and "Stay Connected". Below the header is a navigation bar with links for "ABOUT THE FTC", "NEWS & EVENTS", "ENFORCEMENT", "POLICY", and "TIPS & ADVICE". The main content area has a breadcrumb trail: "Home » News & Events » Press Releases » Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of Millions of Consumers' Locations Without Permission". The headline of the press release is "Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of Millions of Consumers' Locations Without Permission". Below the headline, it says "Company Will Pay \$950,000 For Tracking Children Without Parental Consent", with the phrase "Tracking Children Without Parental Consent" highlighted in a red box. In the bottom right corner of the page, there is a small number "7".

FEDERAL TRADE COMMISSION
PROTECTING AMERICA'S CONSUMERS

[Contact](#) | [Stay Connected](#)

[ABOUT THE FTC](#) | [NEWS & EVENTS](#) | [ENFORCEMENT](#) | [POLICY](#) | [TIPS & ADVICE](#)

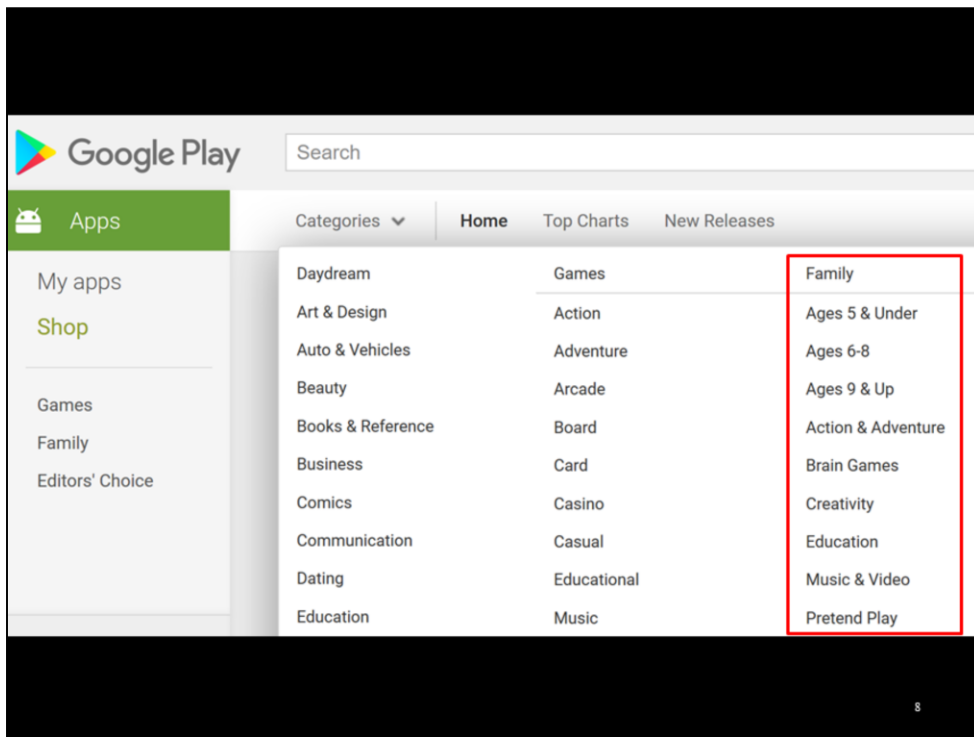
[Home](#) » [News & Events](#) » [Press Releases](#) » Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of Millions of Consumers' Locations Without Permission

Mobile Advertising Network InMobi Settles FTC Charges It Tracked Hundreds of Millions of Consumers' Locations Without Permission

Company Will Pay \$950,000 For **Tracking Children Without Parental Consent**

7

InMobi is an ad network found bundled in children's apps. In 2016 fined for collecting location data from children.



What apps does this law apply to? We looked at the “Family” category in the Google Play Store.

Designed for Families

☐ Opt in to Designed for Families

Designed for Families is a developer program for apps and games designed specifically for children and family audiences. [Learn More](#)

Eligibility

All apps participating in the Designed for Families program must be relevant for children under the age of 13 and comply with the eligibility criteria below. App content must be appropriate for children. Google Play reserves the right to reject or remove any app determined to be inappropriate for the Designed for Families program.

7. You represent that apps submitted to Designed for Families are compliant with COPPA (Children's Online Privacy Protection Rule) and other relevant statutes including any APIs that your app uses to provide the service.

<https://play.google.com/about/families/designed-for-families/program-requirements/>

9

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 services** are COPPA compliant. For example, graphics, communications, analytics, and ads.



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





10

Apps collected between November 2016 and March 2018

Average 750K installs

Representing nearly 1900 developers

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 services	19%
➔	 Failure to take security measures	40%

11

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%.

**potential violations often arise
from **third-party services** included with apps**

12

We attributed most of these violations to various third-party services bundled with apps.

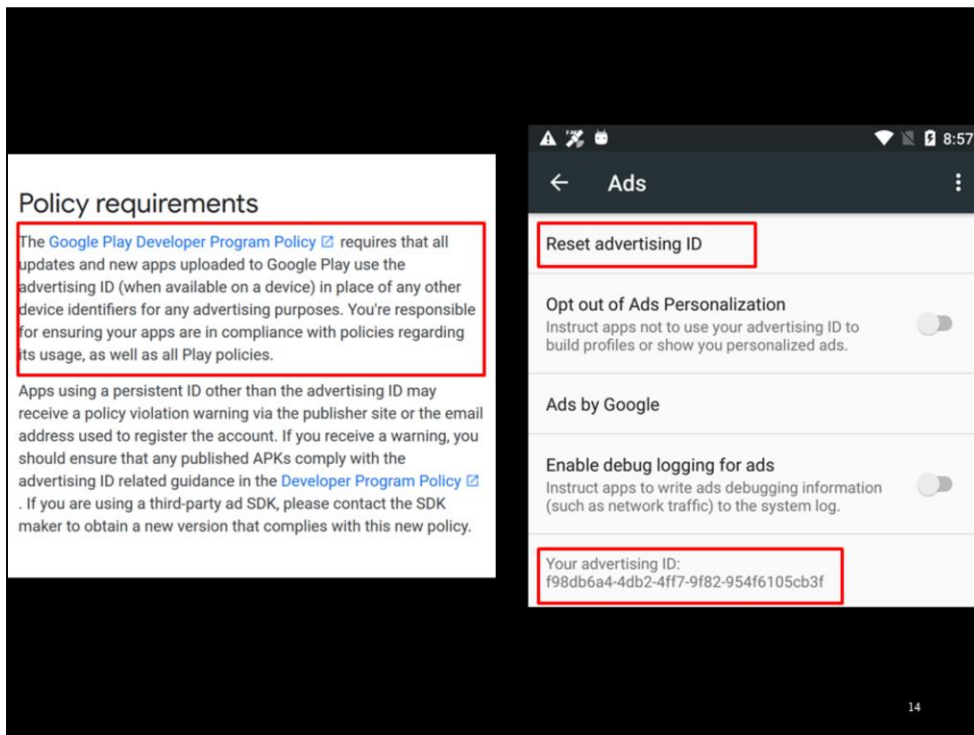
These services allow developers to expedite production by offering drop-in functionality, whether for graphics, communications, advertising, or analytics, among others.

**potential violations persist
due to platform providers not enforcing terms**

13

We believe that these violations are prevalent because the gatekeepers in the mobile app space are not enforcing their own terms meant to protect end-users. (recall DFF requirements)

Google controls the Android operating system and the Play Store, which is the primary app distribution channel for Android. They are in an excellent position to conduct analysis similar to ours on all apps submitted to the Play Store, as well as secure the operating system to prevent potential abuses.



For example, COPPA prohibits behavioral advertising for children. Behavioral advertising uses persistent identifiers to build profiles of users by tracking individuals over time and across services.






Google has recognized the privacy implications of persistent identifiers, and in 2013 introduced the resettable Android Advertising ID (AAID) to give users (or parents) control over how advertisers track them. Since 2014, 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**

15

However, a large chunk of children's apps were seen sharing the AAID with another non-resettable identifier to the same destination, which defeats the purpose of the AAID. Although Google requires the use of the AAID, non-resettable identifiers remain available to apps.

AD PLATFORM	VIOLATION OF IDENTIFIER POLICY
Chartboost 	> 99%
	> 99%
	98%
...	...
mopub	3%
	2%
 DoubleClick <small>by Google</small>	1%

16

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

Full table in paper.



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

17

Not all third party services are appropriate for children, as claimed by those services themselves. We found nearly 1 in 5 DFF apps sharing personal information or identifiers with third-party services whose own terms of use prohibit their deployment in children's apps.

Recall that the apps we studied were opted into the Designed for Families program, indicating that the developers intended to include children in their apps' audience. Still, these same developers were found including these prohibited services.

not for children's apps



18

Presumably, these services prohibit their use in children's apps because these services may engage in non-COPPA-compliant data collection and processing.



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.

19

Crashlytics is a crash reporting service that allows developers to receive usage information about their apps in the wild. Crashlytics terms prohibit its use in children’s apps.

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](#).

20

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.



Potential COPPA violations are widespread, but the reality is 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 certifying bodies.

**industry self-regulation via safe harbors
has had no measurable positive effect**

22

However, we found that apps certified by safe harbors fared no better than DFF apps as a whole

POTENTIAL VIOLATION	DFF (n=5,855)	SAFE HARBOR (n=237)
 PERSONAL INFO	4.8%	10%
 NON-RESETTABLE IDENTIFIERS	39%	39%
 PROHIBITED SERVICES	19%	33%
 NO BASIC SECURITY MEASURES	40%	49%

23

In fact, they were in some cases were worse.

There's a large body of economics research into adverse selection, in which bad actors are the ones most likely to participate in positive signaling activities.

We suspect safe harbors have had the unintended consequence of allowing potentially non-compliant apps to signal that they are indeed COPPA compliant.

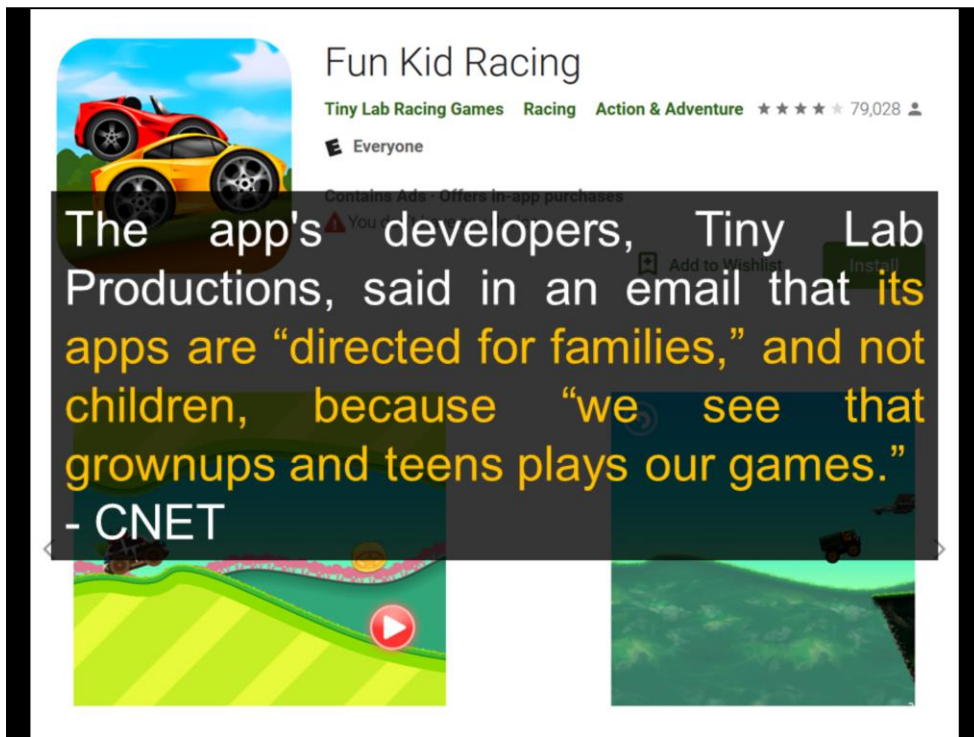


industry and regulators react

24

Our study has had an impact in industry and enforcement since its release last April.

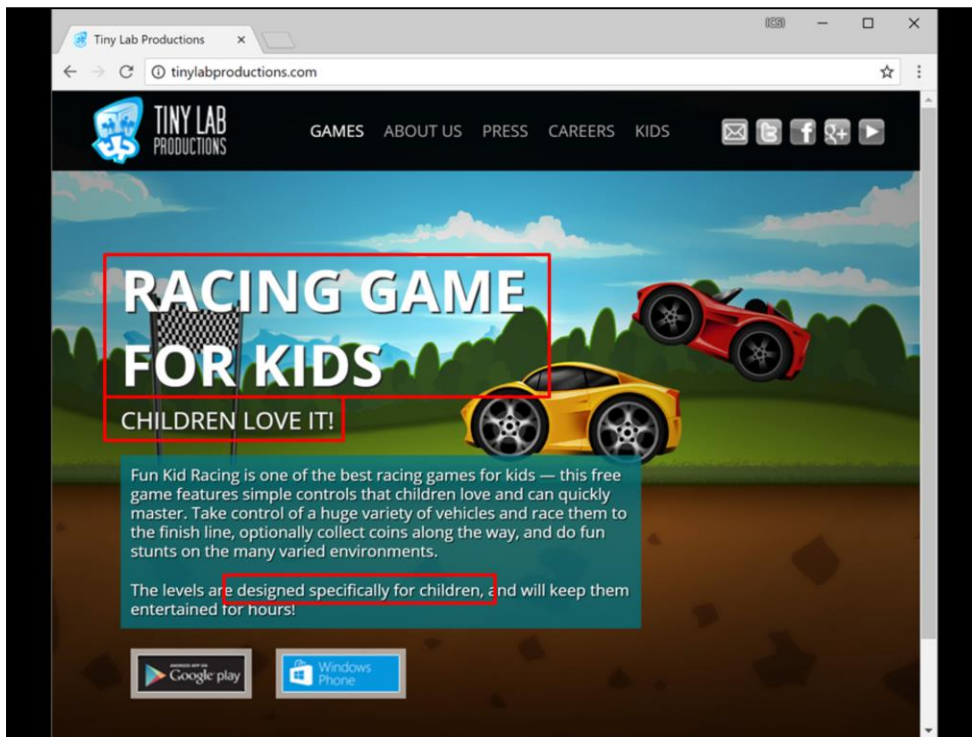
I'll close this presentation with an example of such impact.



In our study, we named Tiny Lab Productions’s games as a popular example of the collection of personal information from children without verifiable consent.

Their game Fun Kid Racing has over 10M installs, and was seen collecting and sharing geolocation data with advertisers. Of Tiny Lab Production’s 82 DFF games, we observed this behavior in 81 of them.

In response to our findings, Tiny Lab Productions stated to CNET that their games are not necessarily for children.



We have identified that 2,667 apps are potentially incorrectly listed as directed to “mixed audiences,” and “not primarily directed to children,” corresponding to ~51% of Designed for Families (DFF) apps from our original sample which are still listed on DFF.

Developers seem to have an incentive to miscategorize their apps as “not primarily directed to children” so they will be able to engage in defective “age gating,” thereby very likely causing children under 13 to enter ages over 13, allowing COPPA-prohibited behavioral advertising.

Email from our team to Google

27

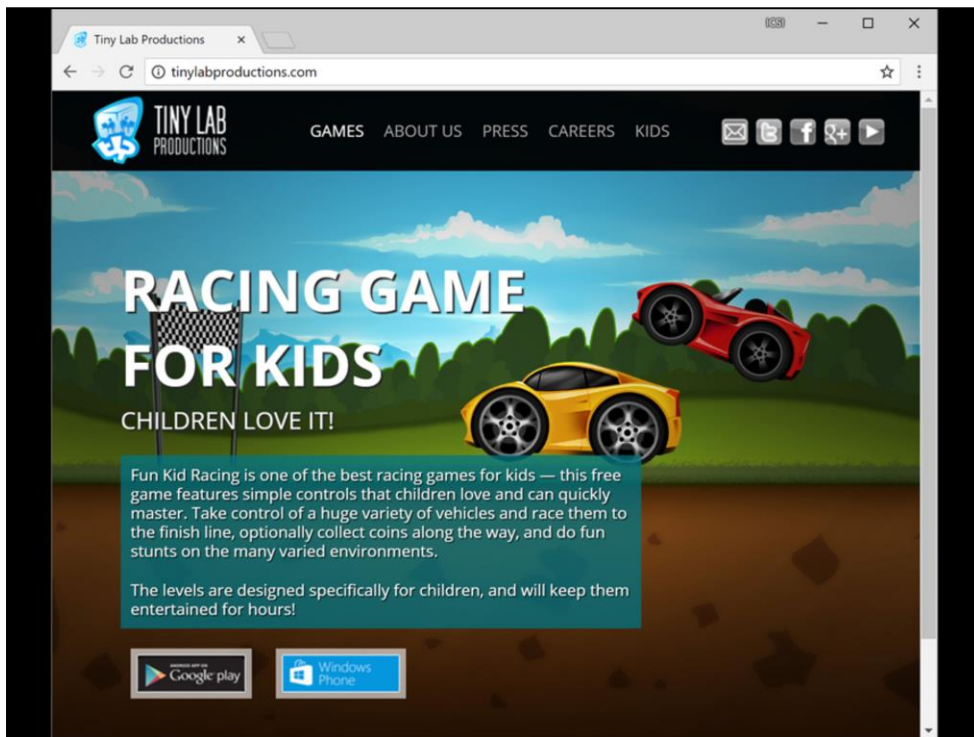
We reported Tiny Labs to Google, along with our results identifying all other DFF apps potentially violating COPPA and failing to meet Google’s own standards for DFF apps

115. In response, Google offered two primary rebuttals: (1) there was no mechanism to detect and prevent the issue “at scale”; and (2) more information was required on appropriate “heuristics” to support the conclusion that the Tiny Lab apps at issue were child directed.

28

Google responded to us saying that there was no way to detect these issues at scale, and that it was unclear that Tiny Labs was offering child-directed apps.

1) This was exactly the technology we developed and deployed in the course of this research



2) Definitely *not* for kids

The New York Times

How Game Apps That Captivate Kids Have Been Collecting Their Data

A lawsuit by New Mexico's attorney general accuses a popular app maker, as well as online ad businesses run by Google and Twitter, of violating children's privacy law.

By JENNIFER VALENTINO-DeVRIES, NATASHA SINGER, AARON KROLIK and MICHAEL H. KELLER SEPT. 12, 2018

<https://www.nytimes.com/interactive/2018/09/12/technology/kids-apps-data-privacy-google-twitter.html>

30

In September, the New Mexico Attorney General filed a suit, with Tiny Lab Productions and Google as co-defendants for violating children's privacy law.

A month later, Google appeared to reverse course: The company told Mr. Abromaitis it had identified a Tiny Lab app that should be designated for children. Google gave Tiny Lab a week to change that app and any others like it. Tiny Lab labeled 10 of its apps for children and used ad networks in them designed for children's apps. Google approved the updates but flagged more apps at the end of August, Mr. Abromaitis said, so he made another round of changes.

Then, this week, after inquiries from The Times, Google terminated Tiny Lab's account and removed all of its apps from the Play store, citing multiple policy violations.

<https://www.nytimes.com/interactive/2018/09/12/technology/kids-apps-data-privacy-google-twitter.html>

31

After facing scrutiny from the New York Times and the New Mexico AG's office, Google recently took a more aggressive stance towards Tiny Labs, taking down their apps after Tiny Labs failed to address the various privacy issues we identified in those products.

closing recommendations

regulators: examine the gatekeepers

platform providers: stricter security and analysis

app developers: use compliant services

parents: ㄟ(ˉ)ㄟ

<https://appcensus.mobi>

32

What did we learn from all this?

The mobile app industry has a small number of gatekeepers who exert much control over the development and distribution of apps. Scrutinizing their practices can be an effective way to achieve compliance at scale.

Google, for example, maintains the Android operating system and its primary app distribution channel, the Play Store. They're well positioned to enforce their own terms of service and deploy security measures meant to protect young users (e.g., DFF requirements).

App developers have a responsibility to be cognizant of the third-party services they use in their products. Nearly all the potential violations we detected arose from these. App developers should verify that the services they integrate into their products are indeed appropriate for child audiences and, if available, are configured with privacy options meant to safeguard children.

Finally, parents are quite limited in what they can do to avoid these potential violations. Our results show that this is a systemic issue. I suppose especially

tech-savvy parents can build their own custom mobile OS, instrument the kernel, and compile data from network monitoring tools, as we did, but that may be out of reach to those who aren't full-time security researchers. Regulation and enforcement are key to protecting consumers in this area.

All our test results and additional findings since the paper are posted at our project website.

BACKUP

33

The image shows a web browser window with two main panels. The left panel displays the Swagger API interface for https://api.wgle.net/swagger#/Network_search_and_information_tools/detail. It features a 'Parameter' table with 'netid' set to '38:1c:1a:c4:ba:b0' and a 'Description' of 'The WiFi search'. Below this, the 'Response Body' is shown as a JSON object:

```
{
  "gsm": false,
  "wifi": true,
  "success": true,
  "cdma": false,
  "results": [
    {
      "trilat": 37.86978912,
      "trilong": -122.27120972,
      "ssid": "ICSI",
    }
  ]
}
```

The right panel shows a Google Maps interface with the coordinates '37.86978912,-122.27120972' entered in the search bar. The map displays a street view of Berkeley, California, with a red arrow pointing from the 'trilat' value in the JSON response to the map's location. The map includes labels for 'Berkeley', 'DOWNTOWN BERKELEY', 'Berkeley Public Library', and 'Ohlone Park'. The bottom of the map shows the text 'Map for 37.86978912,-122.27120972'.

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 wgle.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%

(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

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**

37

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!)

38

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.

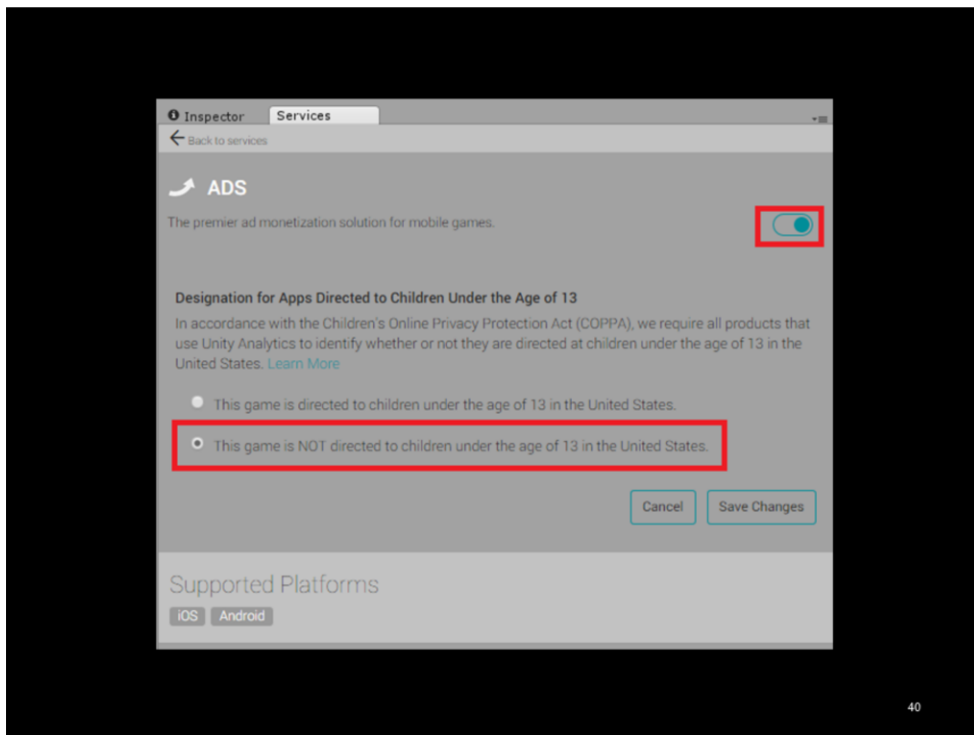
Like Unity, not all apps that integrate with Facebook have the coppa options set. And only a small number of apps consistently set this value to true.

*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.***

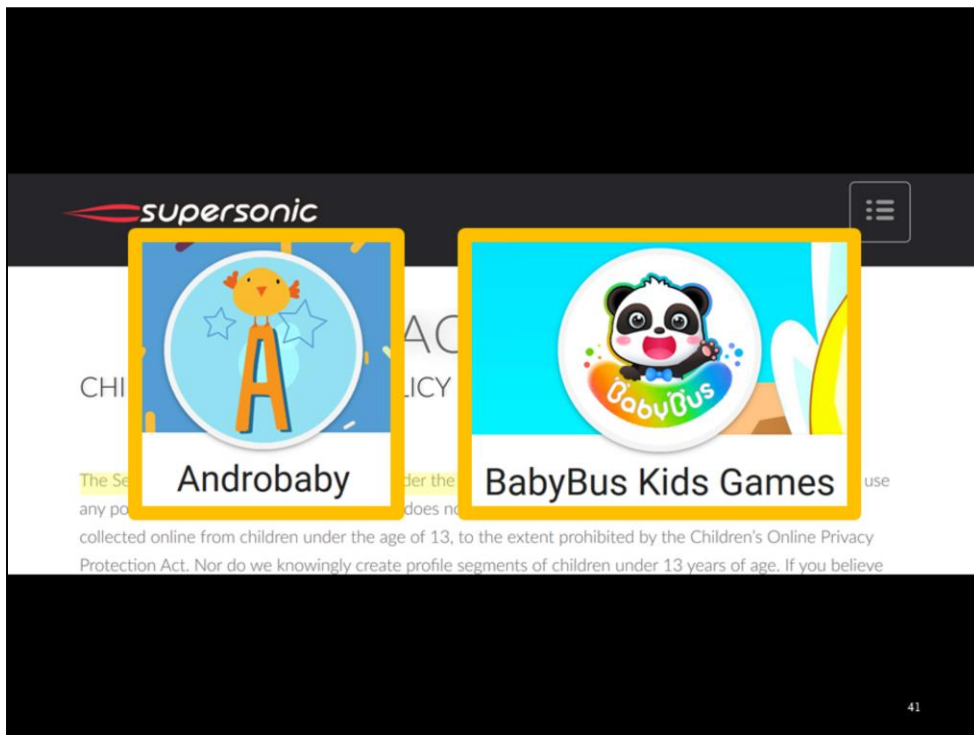
- MoPub Terms of Service

39

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.

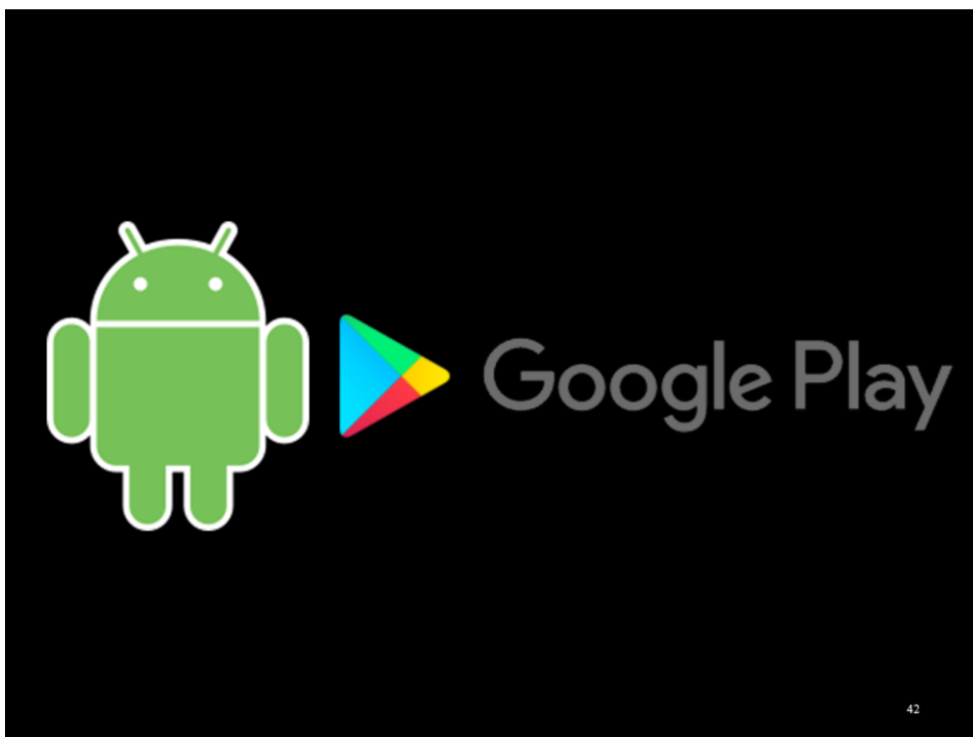


Developers have a responsibility to be aware of the data collection options that their third-party SDKs offer. We identified Unity and Facebook as two popular SDKs that have these options. However, our data suggests that only a small fraction of apps appear to have put these services into COPPA-compliant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

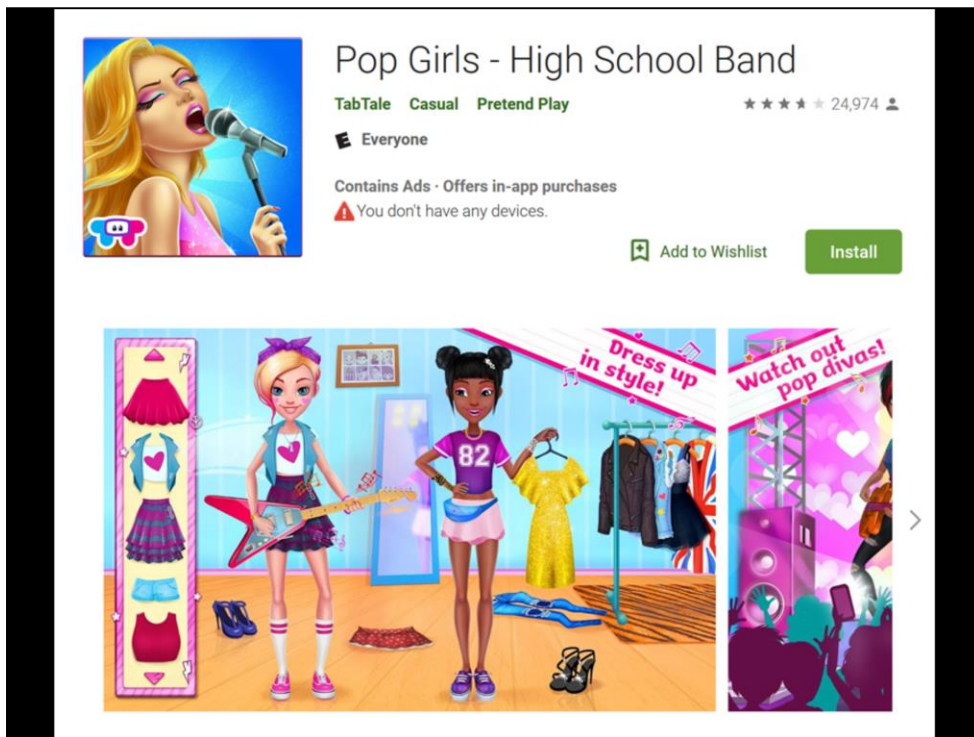


More fundamentally, developers need to know if the SDKs they integrate into children's apps are indeed appropriate for that audience. We observed nearly 1 in 5 DFF apps using SDKs that prohibit their use in apps directed at children.

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.

	DFF 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%

44


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.

**custom android for
logging api calls**



+

**lumen app for
network flow analysis**



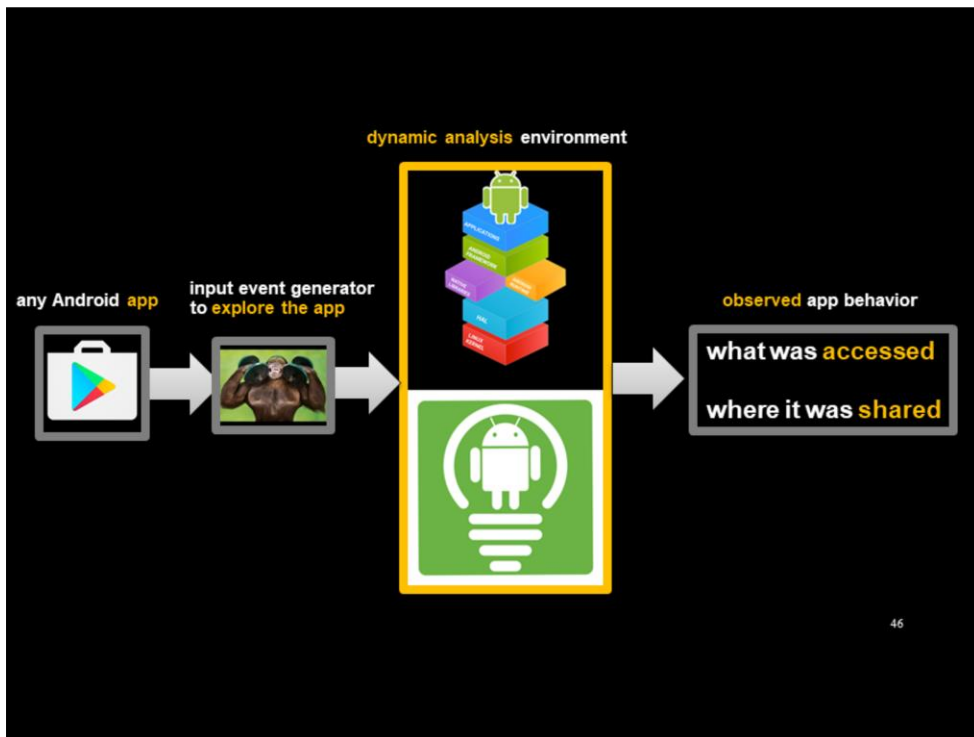
P. Wijesekera, A. Baokar, L. Tsai, J. Reardon, S. Egelman, D. Wagner, K. Beznosov, *The Feasibility of Dynamically Granted Permissions: Aligning Mobile Privacy with User Preferences*, IEEE Security and Privacy (Oakland) 2017

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

45

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.

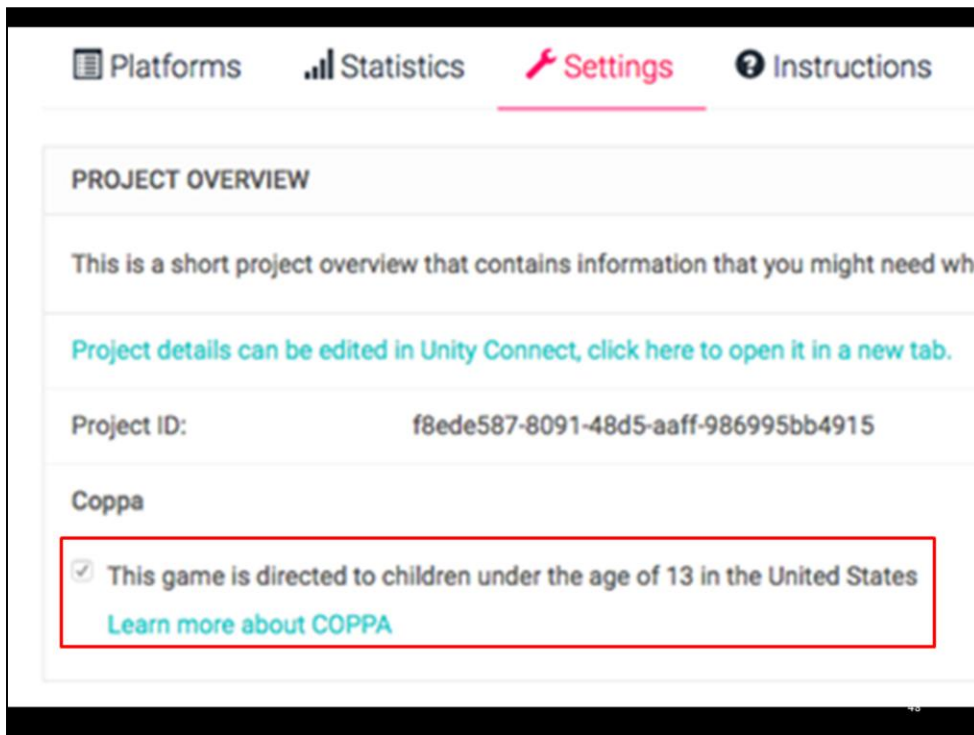
Not enough to just launch the app. Solution: explore with monkey. It's dumb!

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**

47

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



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

49

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.

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

50

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**

51

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.

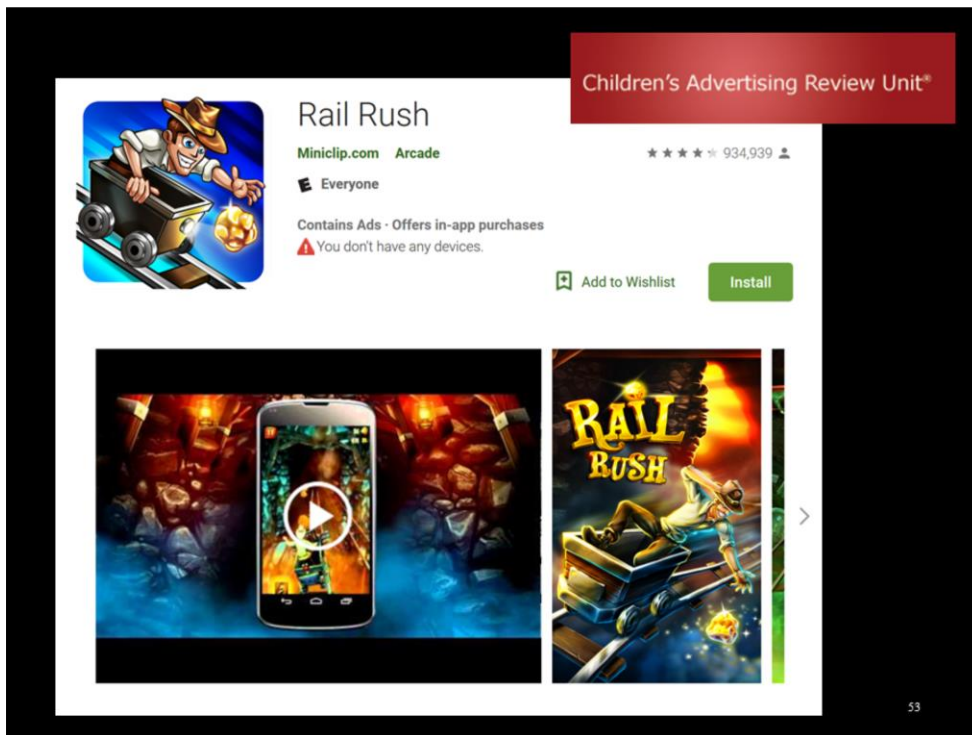
We found 40% DFF apps transmitting potentially sensitive information to remote services without using encrypted HTTP as a basic security measure.



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

52

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.



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.