

Safetypedia: Crowdsourcing Mobile App Privacy and Safety Labels

Lisa LeVasseur
Internet Safety Labs
San Diego, CA

Lisa.LeVasseur@internetsafetylabs.org

Irene Knapp
Internet Safety Labs
San Diego, CA

Ireneista@internetsafetylabs.org

Bryce Simpson
Internet Safety Labs
San Diego, CA

Bryce.Simpson@internetsafetylabs.org

George Vo
Internet Safety Labs
San Diego, CA

George.Vo@internetsafetylabs.org

Saish Mandavkar
Internet Safety Labs
San Diego, CA

Saish.Mandavkar@internetsafetylabs.org

Abstract—Transparency is a foundation for consumer product safety. Users of mobile apps deserve accurate safety labels that identify the unavoidable risks present when using technology as intended. How do we generate accurate and current safety labels for mobile apps at scale? The authors describe a pilot project called Safetypedia, a crowdsourcing software platform and worldwide community of certified “safety inspectors” to collect mobile app behavior data, that semi-automatically generates robust and freely available app safety labels. Several challenges exist around this approach such as: (1) how to ensure high quality data collection from the crowdsourced community, (2) how to automate more parts of the data collection and assessment process, (3) how to “taxonomize everything”, (4) if AI be valuable in the process, and (5) will the crowdsourcing approach work?

Keywords—software product safety, privacy, transparency, accountability.

I. INTRODUCTION

Transparency is a foundation for consumer product safety. People deserve to know the safety and privacy risks that are inherent in the products that they use on a daily basis. Technology products are anything but transparent. Privacy policies and terms of service are both designed to protect the company, not the users of technology. Moreover, the data in those policies is descriptive and doesn’t communicate risk. What is needed are accurate product safety labels, that convey objectively-measured innate risks users face when using the technology.

Since 2019, Internet Safety Labs has been developing standards and infrastructure to measure safety and privacy risks in websites and mobile apps. The authors have supported the development of a published baseline software safety standard [1], manually tested more than 1,700 mobile apps, and developed a web service called App Microscope (<https://appmicroscope.org/>) to display mobile app safety labels. The next objective is to scale the infrastructure to produce accurate and robust safety labels for all mobile apps (and

eventually websites). In 2024, the authors launched a pilot project, “Safetypedia”, to crowdsource privacy risk data collection. The project trains and certifies people to perform privacy inspections on mobile apps, submit findings and collected network traffic files into the Safetypedia database using the custom dashboard. With some additional manual research support, a new safety label for the mobile app is generated and viewable via App Microscope.

II. WHY SAFETYPEDIA

A. User Safety

There are misconceptions and misappropriations around the meaning of “safety” when it comes to technology. We offer the following disambiguation to address the common conflation of safety with security (or more specifically, cybersecurity):

- *Security* is concerned with protecting the integrity of the system; to ensure the system works as designed; to ensure the system isn’t infiltrated, isn’t interfered with, and that digital assets aren’t stolen.
- *Safety*, however, is concerned with ensuring that the system as designed is reasonably safe for the people who use and are impacted by the system; that the system as designed doesn’t itself harm people.

While cybersecurity standards are quite mature, currently, there are no agreed-upon standards for software product safety, though there is increasing awareness and agreement on the fact that software products and platforms are harmful in many ways. Data privacy risks are the most pervasive and widely recognized safety risks in digital products today, though not the only safety related risks. Privacy harms are just one of what we call *programmatic harms*—i.e. invisible and unavoidable harmful behaviors inherent in software. Other such programmatic harms include those from deceptive user interface patterns (aka *dark patterns*) and algorithmic decision-making.

Like all commercial products, consumers of technology deserve to know the objectively measured safety risks in digital products and services.

B. Safety Labels at Scale

This project builds off mobile app safety labels as seen in App Microscope by Internet Safety Labs (ISL). Fig.1 shows the summary section of an ISL mobile app safety label.

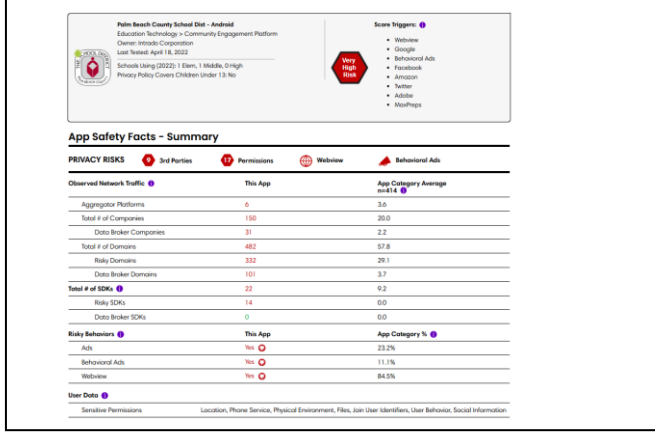


Fig. 1. Example Safety Label (Partial, <https://appmicroscope.org/app/1582/>)

ISL notes in the App Microscope help text that today's safety labels only provide privacy risks but are designed to include *all* safety risks for all apps over time [2]. The authors observe that a key challenge to this goal is having a worldwide, accurate, and current supply of safety labels, and therefore app "testings". Two potential solutions have been considered by the researchers: (1) automated testing, such as those being developed in the ProperData NSF projects [3], and (2) crowdsourced human-performed safety inspections.

This proposal shares the researchers' experience with prototyping a crowdsourced solution called Safetypedia, which fosters a community of certified safety inspectors to audit mobile apps anywhere in the world, resulting in freely available safety labels. The idea is inspired by crowdsourcing data á la Wikipedia but includes more rigor and quality assurance around the contributors and the data. Also, the contributors provide structured data that is programmatically transformed into safety labels.

III. SAFETYPEDIA

The Safetypedia pilot consists of three functions:

1. Safety inspector training and certification.
2. Individual app inspecting & data collection.
3. Data processing and safety label generation.

A. Safety Inspector Training & Certification

Participants in the Safetypedia program are trained (via video call) to perform a privacy inspection on an app and to enter the necessary data into the Safetypedia portal. Participants are also "certified" by being tested on their understanding prior to moving on to the next stage. Anyone can be trained to perform a safety inspection; it does not require special technical knowledge.

B. App Inspecting & Data Collection

Once certified, participants can audit any apps (from anywhere in the world). In our first group of inspectors, one participant was in New Zealand, for example. Testing location is important information to be presented in the safety label since app behavior can vary due to varying privacy laws across geographic jurisdictions.

Along with app behavioral details and a network traffic file from the inspection, inspectors provide inspection metadata such as the location (state and country) of the inspection, and the age of the user conducting the inspection. Note that some inspectors test as a "child" as well as an adult to collect app behavioral differences for different audiences.

C. Data Processing & Safety Label Generation

1) Generating an App Risk Score

The app risk score reflects privacy riskiness of the first party and the first party's parent organization, and the riskiness of all third parties either observed in the network traffic or in included software development kits (SDKs). The score relies on company risk scores (currently shown in <https://internetsafetylabs.org/resources/references/company-privacy-risk-dictionary/>). The company privacy risk score includes assessment of the following risks: (1) monetizing of personal information, (2) occurrence of data breaches, and (3) privacy-related regulatory action or civil legal action, among other observed corporate behaviors.

When a new app inspection is submitted by an inspector, the ISL research team checks for any new [unscored] subdomains or companies associated with the inspection.

Generating privacy risk scores for the related entities requires a substantial amount of additional [manually-performed] research to collect and process privacy risks including:

- Manual determination of corporate ownership of new domains found in the network traffic.
- Manual collection of privacy risk data for new companies found in the network traffic or associated with an app not previously inspected.

Once this is complete, the data is available via a new safety label for the app in App Microscope. Thus, over time, multiple labels (i.e. "testings") will be presented for each app.

D. Discussion

We propose having a collaborative discussion sharing Safetypedia results to date and soliciting feedback on the approach, including the following challenges in particular:

1. Can crowdsourcing work as an effective tool for product safety transparency?
2. How to scale data quality assurance?
3. How to automate data collection?
4. The researchers have been exploring using machine learning tools to help (especially with taxonomizing) but so far, they've been ineffective. How much can these tools help in the future?

ACKNOWLEDGMENT

Safetypedia is supported by funding from Wrethinking, the Foundation and the Internet Society Foundation.

REFERENCES

- [1] ISL Software Safety Standards Panel, “ISL Safe Software Specification for Websites and Mobile Apps v1.1,” July 5, 2022. [Online]. <https://internetsafetylabs.org/resources/specifications/isl-safe-software-specification-for-websites-and-mobile-apps-v1-1/>
- [2] App Microscope About Page, Internet Safety Labs. [Online]. <https://appmicroscope.org/help>
- [3] ProperData. University of California, Irvine. [Online]. <https://properdata.eng.uci.edu/>