A Multi-Faceted, At-Scale Analysis of Apps Privacy Disclosures in the Apple and Google App stores

David Rodriguez*, Akshath Jain, Jose M. Del Alamo and Norman Sadeh
* Presenter

1. Abstract

Recently introduced privacy labels have raised **concerns** about their reliability and completeness compared to privacy policies.

This work aims to **unveil potential discrepancies** between privacy labels and policies with ATLAS framework (Automated Privacy Label Analysis System).

We found **misalignments** Label-to-label, label-to-policy and label-to-behavior between matching apps of Apple and Google Play stores.

2. Data collection

➤ App information collected from **918,293** iOS applications.
➤ **183,135** matching Apple & Google Play apps found.
  • **32,127** of them had both privacy labels, policies and were updated within less than two weeks in both stores.
  • **5,697** had identical privacy policies.
  • **3,622** could be downloaded.
➤ **105,131** iOS apps' privacy policies and labels discrepancies analyzed using ATLAS framework.

3. Training classifiers to generate privacy labels from privacy policies

We reduced the noise in the dataset through importance sampling. The following figure shows the sampling procedure for privacy policies regarding the collection of user's name.

(a) Random Sample   (b) Clusters   (c) Importance Sample
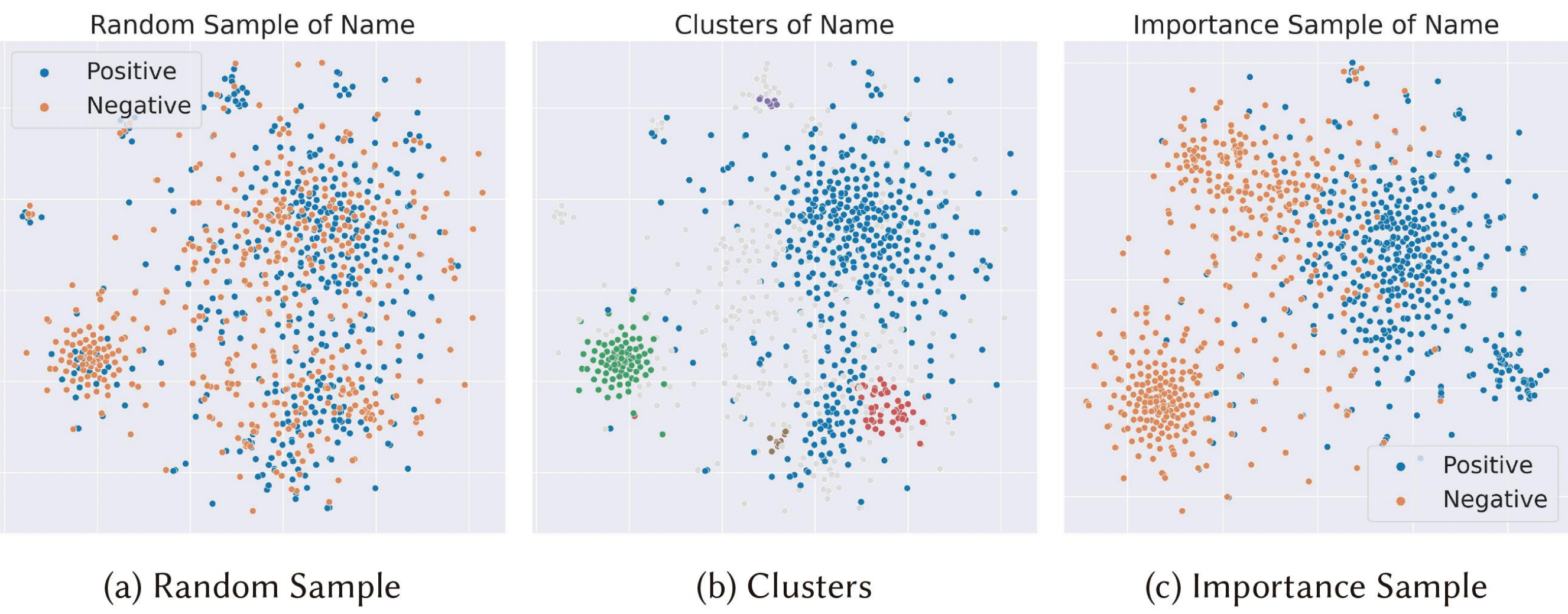
## 4. Performance of ATLAS' ensemble-based classifier
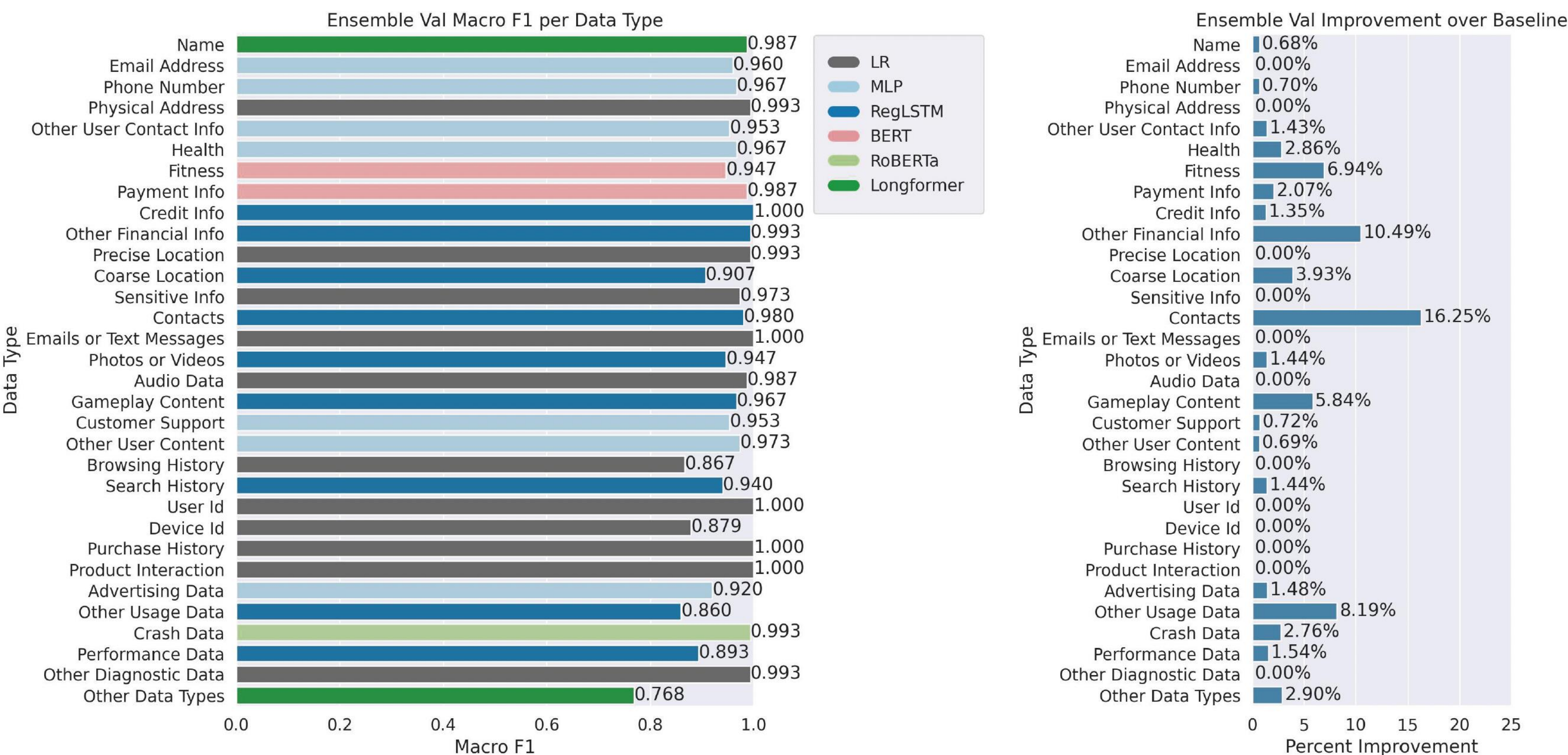
The classifier is able to predict privacy labels from policies for each specific type of data. It was created using several language models as BERT or RegLSTM, achieving an average F1-score of 91.3%.

5. Potential compliance issues comparing labels and policies

We used the trained classifier to perform a discrepancy analysis on 61,596 iOS apps. Discrepancies per each type of data are classified in highly (dark blue) and lowly probable (light blue). Results show that **88% of apps had at least one discrepancy** between policy and labels.

(a) Incomplete Policies   (b) Incomplete Labels
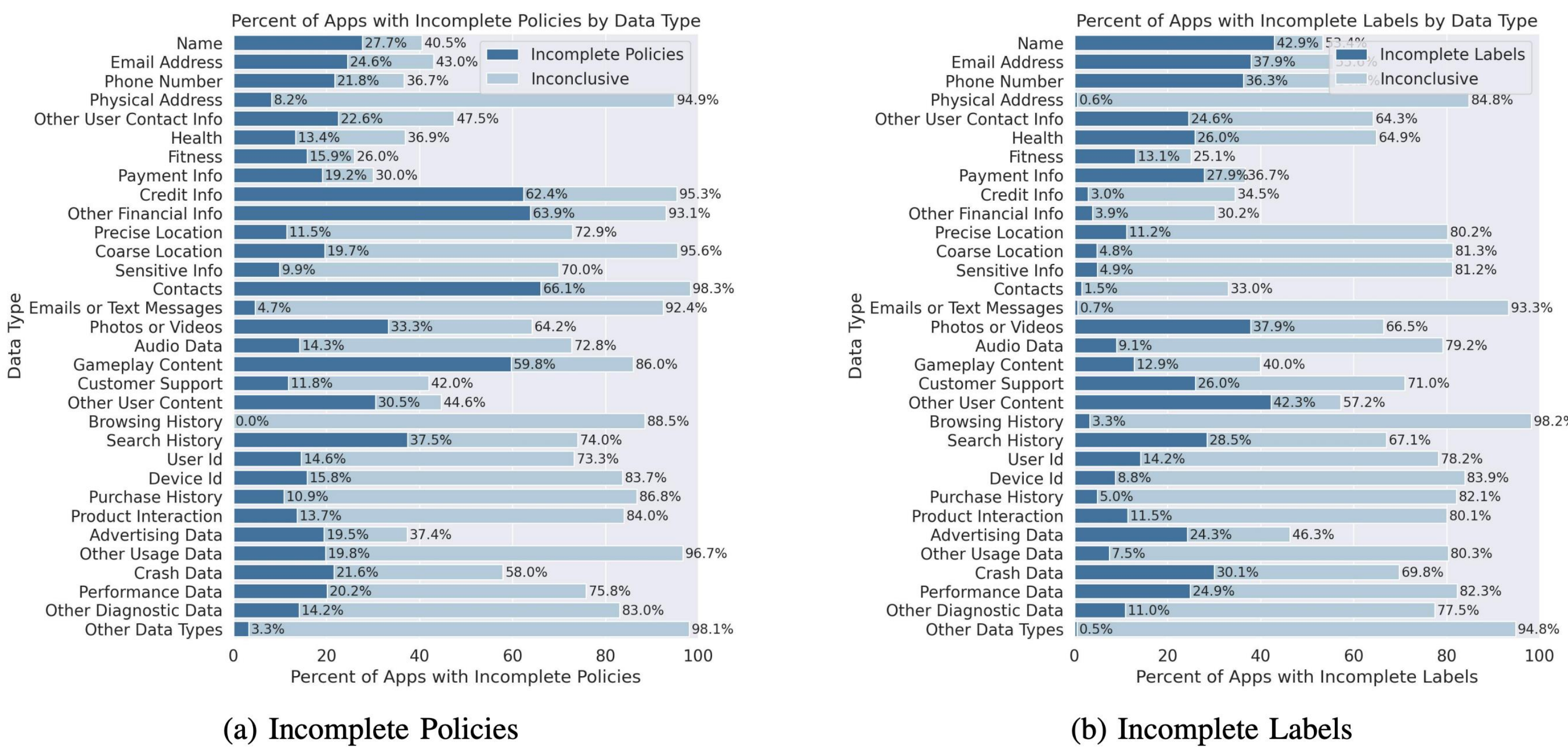
6. Comparing privacy labels for identical Android and iOS apps
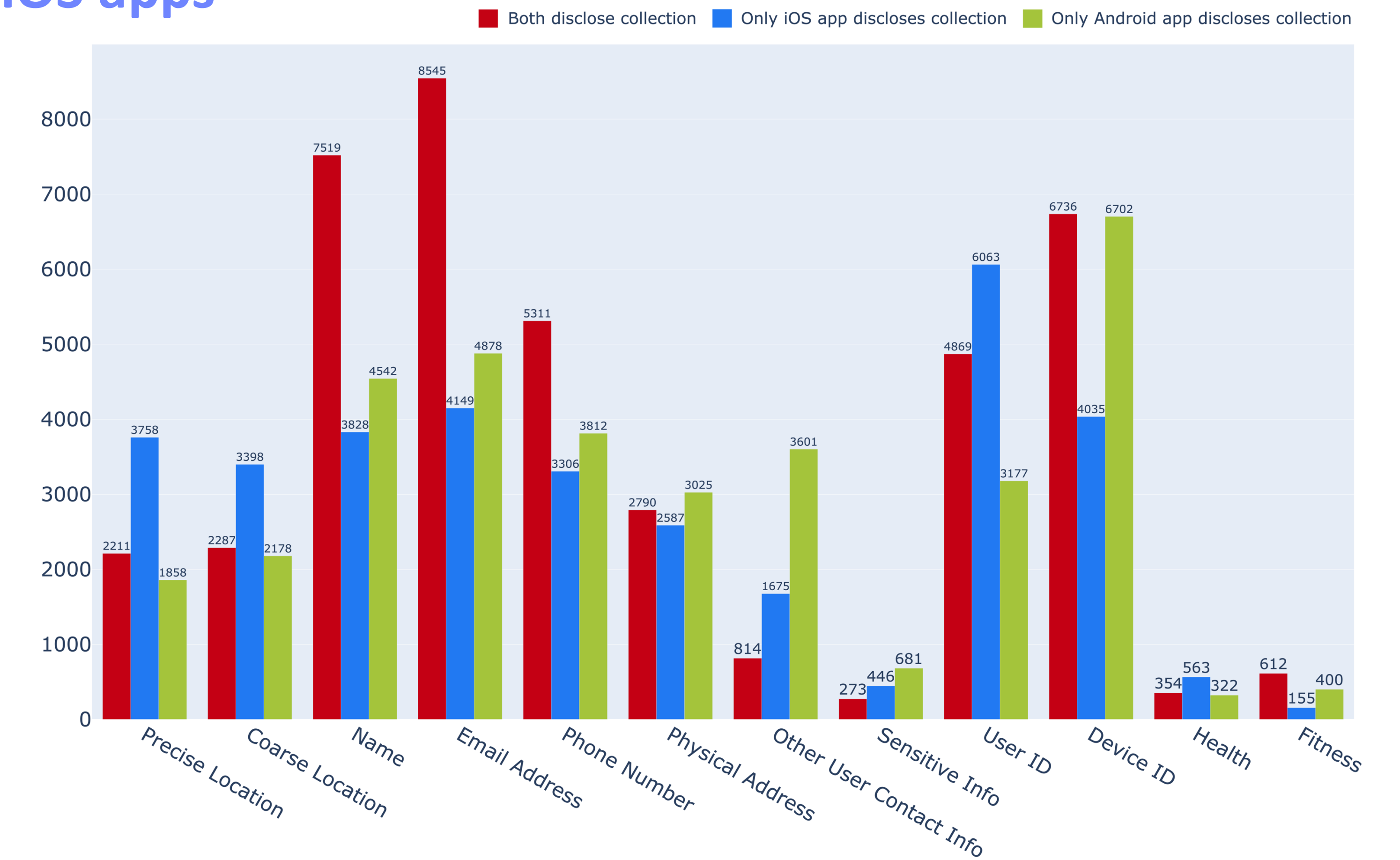
Comparing data claimed to be collected in the privacy labels by the same apps: The blue and green bars show the discrepancies between the Apple and Google Play markets.

7. Static Analysis

Number and percentage of apps accessing each data type but not disclosing collection.

Data type	No. of apps accessing data	No. (%) of apps not disclosing
Name	56	14 (25.00%)
Email address	44	16 (36.36%)
Health or Fitness info	2743	2743 (100.00%)
Contacts	2480	2389 (96.33%)
Other audio files	1759	935 (53.15%)
Photos	3324	1827 (54.96%)
Videos	2651	1666 (62.84%)
SMS or MMS	248	235 (94.75%)
Voice or sound recordings	0	0 (0.00%)
Calendar events	1969	1936 (98.32%)
Files and docs	3480	2465 (70.83%)
Approximate location	987	742 (75.17%)
Precise location	7	7 (100.00%)
User payment info	950	902 (94.94%)
Device or other IDs	3241	1104 (34.06%)
User IDs	1557	955 (61.33%)
Web browsing history	0	0 (0.00%)

8. Dynamic Analysis

Number and percentage of apps **sending** each data **off-device** type but not disclosing collection.

Data type	No. of apps accessing data	No. (%) of apps not disclosing
Approximate location	0	0 (0.00%)
Precise location	9	8 (88.89%)
Device or other IDs	531	247 (46.52%)
User IDs	269	179 (66.54%)

CyLab
Carnegie Mellon University
Security and Privacy Institute