CODE V. WADE: A Privacy Evaluation Platform for Reproductive Health Applications

School of Engineering & Applied Science
THE GEORGE WASHINGTON UNIVERSITY

Olivia Legault†, Rubin Roy†, Miles Grant†, Yasemin Acar*
Department of Computer Science
†All authors contributed equally *Principal Investigator

Introduction

- **Reproductive health applications** (RHAs) such as period trackers and fertility apps **collect the personal data of millions of Americans**.
- Data from these platforms have been used to **prosecute people seeking abortions**.

Objectives

- Build a **core understanding** of how RHAs protect personal health information
- Determine if RHA developers **accurately represent** the privacy their product provides
- Provide a **searchable website** for users to make an informed decision on which apps to use
- Provide an **objective ranking** that researchers & advocates can use when analyzing RHAs

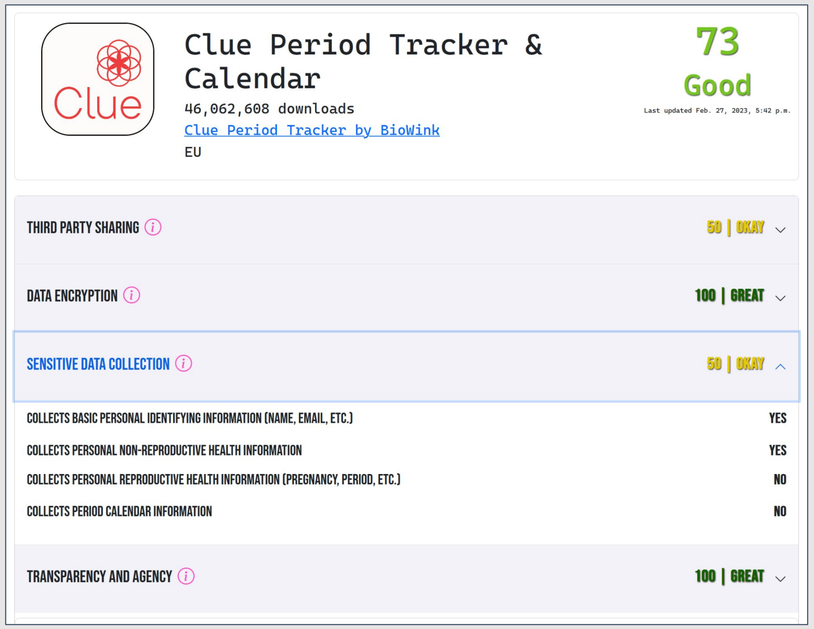
Methodology

Framework: We employ a **mixed-methods** approach to determine the privacy of RHAs.

1. Packet Analysis: We use an Android emulator to run a series of tests on the application that simulate actions of a **real-life customer** (such as creating an account, deleting an account, etc.). We then use **mitmproxy** to intercept packets leaving the emulator. These packets of data can contain **personal identifying information** and sensitive health data such as medical conditions and period calendar information.

2. Privacy Policy Analysis:
We use **natural language processing** (NLP) to **parse privacy policies** and determine if they permit data sharing with different entities. If the privacy policy contradicts the findings of the packet analysis, this can indicate the **truthfulness of the policy**. We currently use **ChatGPT+** (based on **GPT-4**) as our model.

3. App Store Analysis:
We scrape the **Google Play Store** for basic app information (downloads, rating, etc.) and developer info (name, continent, email)



Privacy Scores: Each of the three above methods is used to evaluate a list of privacy metrics (e.g. "does not allow data deletion", "shares data with third party"). If any method indicates that the app fails to meet the privacy standard for that metric, points are deducted from the Code V Wade privacy score, measured out of 100.

Displaying Results: The scores and information for each app are hosted on a **public open-source** website with a **searchable database**. Users can also **submit a request** for an app to be reviewed or updated through the site.

Future Work and Deployment

- **Streamline** manual packet analysis process
- **Weight** different privacy metrics based on user sentiments and concerns
- **Provide** more options for searching and visualizing our database

References and Acknowledgements

- https://mitmproxy.org/
- https://chatonai.org/chatgpt-plus
- https://openai.com/blog/chatgpt-plus
- https://play.google.com/about/howplayworks/
- https://techcrunch.com/2023/01/27/digital-data-roe-wade-reproductive-privacy/
- https://deepai.org/machine-learning-model/summarization
- Special thanks to Masood Ali (UIUC), Adellar Irankunda, Professor Adam Aviv (Department of Computer Science), and the GWU Usable Security and Privacy Lab

olegault@gwmail.gwu.edu rroy1@gwmail.gwu.edu milesgrant@gwmail.gwu.edu acar@gwu.edu codevwade@gmail.com