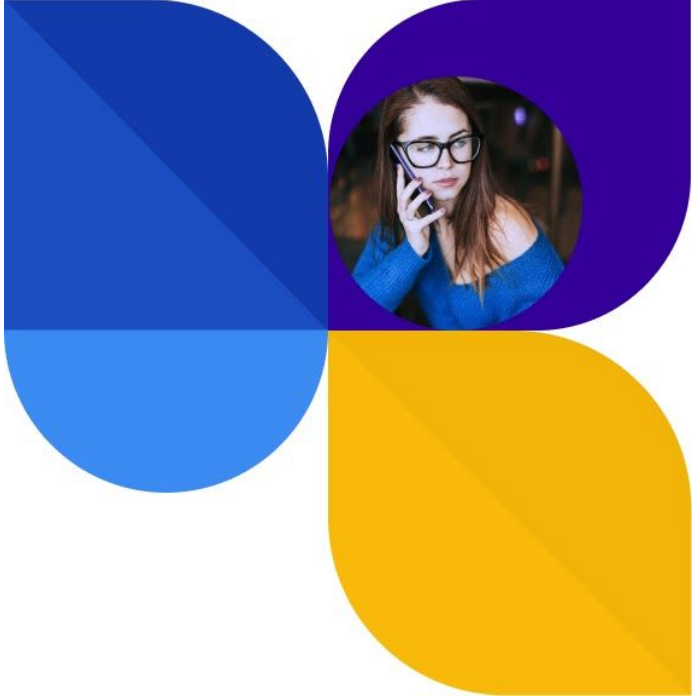

LEMON

The COVID-19 Tracing-App

Survey Overview

10/09/20

Introduction

This document outlines, in detail, the survey functionality of the LEMON application. It deals with the purpose of this feature, how the data is processed and where it is passed to, alongside how the information is stored whilst still protecting the privacy of the user.

Purpose

The purpose of the survey is to provide a wide range of epidemiological data in order to contribute to the medical community in order to better understand the development of symptoms and the spread of COVID-19. This information can be shared anonymously by the user with relevant health authorities.

Due to the nature in which this information is gathered, not every data point can be confirmed as 100% reflective of a user's real-life situation. Although steps have been taken in order to minimise the possibility of false data being recorded, there is still potential for user error when submitting their information. With this in mind, LEMON still believes that this data has value and will provide a unique insight into near real-time data on how the virus is affecting people.

User Flow

This section describes the user experience around the survey submission. For more information, A detailed and complete user flow for the entire app can be found in the user journey document.

- A user chooses to update their COVID-19 status. They are presented with the following options:
    - I don't know
    - Negative
    - Positive
    - Recovered
- A user selects their status and are taken to the next screen
- Should the user have selected positive, they are taken through a separate flow first
    - A verification code is required in order to confirm that they have indeed tested positive. This code is linked back to a specific test result and verified on central health authority servers.
    - Should the positive test be verified, they are asked whether they want to share their contact tracing information and notify other users that have been in contact with them.
    - They are then taken back into the normal flow.

- The user is taken to a survey screen
- The user can choose what information they want to enter or completely skip this step
- The user is asked whether they want to share the survey information that they have completed.
    - Should they choose to do so, the information is sent to the server and stored anonymously.
    - Should they refuse the data is stored securely on the device only.
- The user is taken back to the dashboard where they will be able to see their updated status.

Data Flow

In order to align with the core values of LEMON, strict protocols are put in place over what data flows where. Following the decentralised design, data only leaves the device should the user give permission to share this information. Once this data is donated it is stored completely anonymously and therefore cannot be recovered.

Should the user choose to submit their data, the app sends the survey information to the server behind end-to-end session-based encryption. It must also be noted that in order to submit data to the server, a valid session between an application and server must be established with a JSON Web Token (JWT) submitted alongside the survey data.

This JWT contains the hashed instance app-id that is used to anonymously identify a specific application to the server. This is required in order to validate the number of submissions that a specific device has made over a defined period of time. Click here for more detailed information on the submit survey endpoint.

Data Storage

Due to the nature of the project, the data can be stored in up to 2 places, the server and the device. All data is stored to the device and only shared should the user give permission.

Device Storage

Upon entering the information and confirming on the survey page, all data entered is stored onto the device. This data is encrypted within storage with the application key pair to ensure its privacy. Only when a user logs into the application is this data unencrypted and becomes accessible to the application only.

The submitted information is also stored alongside the date of submission. This is required for validation purposes as explained in the section below.

Server Storage

Should the survey data be stored it is sent to the server with the hashed app instance Id. It must be noted though that the survey data itself is not stored against this id in any way. This is to endure that the submitted data remains completely anonymous and can in no way be traced back to the user.

However, in order to maintain some sort of validation on information submitted by a device, a submission log stores each submission by hashed app instance id against a generalised time (chopped to display only the day). This log is in no way linked to the submitted survey data meaning there is no way to match a submission log to a set of survey information.

Data validation

When submitting statistical data like this it is essential to put in place some validation on both the server and application in order to reduce and prevent any malicious or accidental attacks occurring in order to skew or ruin the data set. In light of this, the following validation is put in place on the survey flow.

App Validation

As the app has full access to the data that is submitted a full validation process can take place. This validation includes the following:

- A specific status cannot be submitted more than twice in one month
- A device can only submit up to and including 3 surveys in 1 week
- A device can only submit up to and including 4 surveys in a month

Server Validation

Due to the nature in which the data is anonymised on the server, this makes it hard to implement validation as there is no way to trace submitted data to a specific device. Therefore some of the above validations that take place on the application itself cannot be implemented on the server.

However, by storing a submission log it is possible to put some validation on the number of times a device can submit survey data over a specified time period. Therefore the following validation are put in place:

- A device can only submit up to and including 3 surveys in 1 week
- A device can only submit up to and including 4 surveys in a month

This validation is put in place to limit the effect that a malicious user could have by attempting to spam the LEMON survey data with multiple surveys.

Based upon the fact that the infectious period for COVID-19 lasts at least 2 weeks, it was deemed that the maximum number of statuses in a month should not exceed 4 as shown in a few scenarios below:

1. I Don't Know
2. Tested Positive
3. Recovered

1. Tested Positive
2. Recovered
3. Tested Positive

1. I don't Know
2. Tested Negative
3. Tested Positive
4. Recovered

On top of this, it was deemed unlikely that a specific user would take more than 4 tests in a month should they want to update their status to reflect their latest test information.