SMBUD 2021 - Project work 2

Aman Gabba - 10793117
Andrea Cerasani - 10680486
Giovanni Demasi - 10656704
Pasquale Dazzeo - 10562130
Vlad Marian Cimpeanu - 10606922

POLITECNICO
MILANO 1863

Contents

1 Introduction

1.1 Problem Specification

The aim of this project was to design a 'query document data structure' in MongoDB for supporting a certification application for COVID-19. The database must register all the necessary information about the users including their tests and vaccination status in order to know if a certain person can be considered 'harmless' and so able for instance to visits some public spaces. The application using this database will be able to exploit all the data coming from test results and vaccination status.

1.2 Hypothesis

The assumptions taken into account are the following:

- The recovery document has been introduced in the document DB to distinguish the usual negative test result to the one did after an infection to consider a certain person healed from Covid-19. Using this document can be computationally useful because otherwise, to know if a person has a recovery certificate, it would be necessary to scan all the tests and check the relative results and dates to deduce an infection and a consequent healing. It has been also assumed that the recovery certificate doesn't get automatically released, but it is registered manually by whom of competence after a specific iter. The recovery certificate was not expressly required by the problem specification but it has been introduced because it can be used as support to calculate the expiration date of the certificate.

2 ER diagram

The designed ER diagram contains the following entities: Person, Vaccine facility, Certificate and Medical Stuff. Nurse and Doctor have been designed as sub-entities of Medical Stuff, Vaccine and Swab have been designed as sub-entities of Certificate.

3 Document diagram

The designed Document diagram is shown below. Certificate and Sanitary Operator are represented as collections of documents. Contact is a sub-document of both Certificate and Sanitary operator documents.

The certificate document represents an unique person and it contains him/her Covid information chronology. It can have Recovery, Swab and Vaccination as sub-document and these last two have both an event details sub-document that contains all the information related to the place where the determined event happened.

The Recovery document has been implemented to make the recovery certificate validation easier and faster. Otherwise, doing this validation through the scan of the tests information would have been more complex due to the possibility of a person to do an unlimited amount of tests. With this design choice the time complexity of validate a recovery certificate has been reduced.

4 Dataset description

The Dataset has been built by making a script in Python, using some useful packages like: random-italian-person that automatically generates people with random attributes (this package has been modified due to our necessity); the official PyMongo driver for Python for the communication with the Database since it allows to make queries directly from Python.

We have also used this piece of code to generate hospitals and pharmacies data. This query has been used in overpass turbo (a web-based tool for extracting OpenStreetMap data) selecting the map area of Rome, Naples and Milan.

CODE

5 Queries and Commands

5.1 Queries

5.1.1 Query 1

Query description...

```
CODE
```

5.1.2 Query 2

Query description...

```
CODE
```

5.1.3 Query 3

Query description...

```
CODE
```

5.1.4 Query 4

Query description...

```
CODE
```

5.1.5 Query 5

Query description...

```
CODE
```

5.2 Commands

5.2.1 Command 1

Command description...

```
CODE
```

5.2.2 Command 2

Command description...

```
CODE
```

5.2.3 Command 3

Command description...

```
CODE
```

6 Generator database

The "MongoDB-populator/main.py" file is responsible of generating a random database. In order to correctly run the generator, it is mandatory to...
This script exploits the following packages: `pandas, numpy, PyMongo and python-codicefiscale`.

6.1 Facilities generation

This generator will create facilities data for the following cities: Rome, Milan and Naples. For each city it will create event details representing hospitals and pharmacies available in that specific town.

6.2 Certificate generation

In order to consistently generate people certificates...

6.3 ... generation

...

7 Application description

Goal of the application `"GUI/App.py"` is to give some information about the Covid-19 Certification. The user can use the application in order to:

- Given information about a specific people, print all his valid certifications.
- Given a UCI, check if the related certificate is still valid, who it belongs to and its expiration date according to the actual government regulation.

8 User guide

The application is meant to work on pc or macOS. As the application has been entirely developed in Python, it is not operative system dependent.
In order to correctly run the application it is mandatory to...
The user must check to have installed correctly in the virtual environment the following pacakges: `tkinter`, `pandas`, `PyMongo`, `numpy` and `python-codicefiscale`.

9 Conclusion

Some interesting conclusions can be drawn from the development of this project: Document databases are, if well designed, easy to use and really scalable.

...

10 References and Sources

- Random-italian-person package: https://pypi.org/project/random-italian-person
- PyMongo package: https://docs.mongodb.com/drivers/pymongo/
- Tkinter package: https://docs.python.org/3/library/tkinter.html
- Overpass turbo API: https://wiki.openstreetmap.org/wiki/Overpass_API