

# SMBUD 2021 - Project work 1

Aman Gabba  
Andrea Cerasani  
Giovanni Demasi  
Pasquale Dazzeo  
Vlad Marian Cimpeanu



**POLITECNICO**  
MILANO 1863

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Problem Specification . . . . .	3
1.2	Hypothesis . . . . .	3
<b>2</b>	<b>ER diagram</b>	<b>4</b>
<b>3</b>	<b>Graph diagram</b>	<b>6</b>
<b>4</b>	<b>Dataset description</b>	<b>7</b>
<b>5</b>	<b>Queries and Commands</b>	<b>7</b>
<b>6</b>	<b>References and Sources</b>	<b>7</b>
<b>7</b>	<b>Conclusion</b>	<b>7</b>

# 1 Introduction

## 1.1 Problem Specification

The aim of this project was to design a 'query graph data structure' in Neo4j for supporting a contact tracing application for COVID-19. The database must register all the necessary information about the users including vaccines and Covid swabs in order to have a pandemic trend overview for a given country. The application using this database will be able to exploit all the data coming from tracking applications and from all the public facilities.

## 1.2 Hypothesis

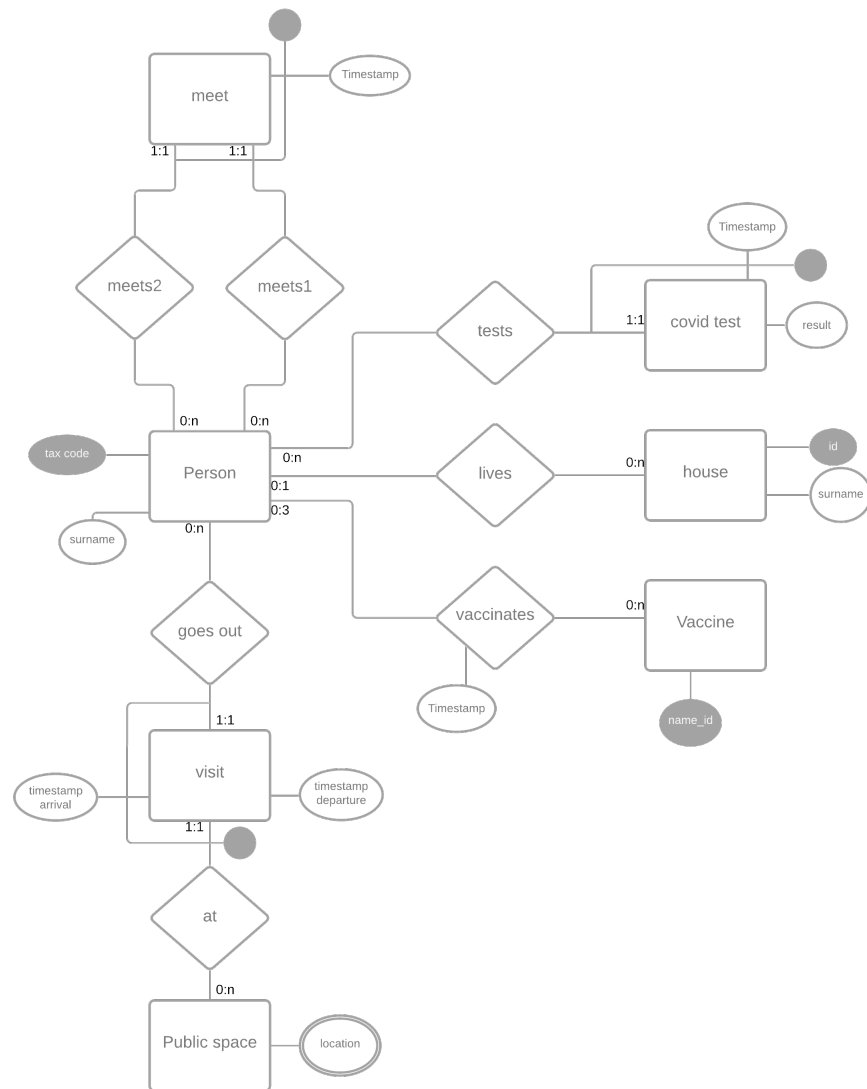
The assumptions taken into account are the following:

- People belonging to the same family live in the same house if not explicitly specified. Using the concept of "house" instead of "family", it is offered the possibility to differentiate domicile from residence.
- All the personal data are verified by an authoritative figure, for instance the government.
- The domicile declaration is assumed to be truthful.
- All the data coming from public spaces are always considered true and complete.
- People always provide all the necessary information to the staff when they visit a certain public space.
- Every MEETS relationship is automatically added to the database by a tracing app when two mobile phones stay in the same range for more than 15 minutes.
- Relationships use the Data type instead of Timestamp to register the relations because of safety and simplicity reasons. The former reason allows people using the Database to trace contacts during all the day and not only during a range of time (for more control). The latter reason is required due to avoid mistakes by the staff of public facilities during the time registration (an error in time registration could lead to a wrong tracing).

## 2 ER diagram

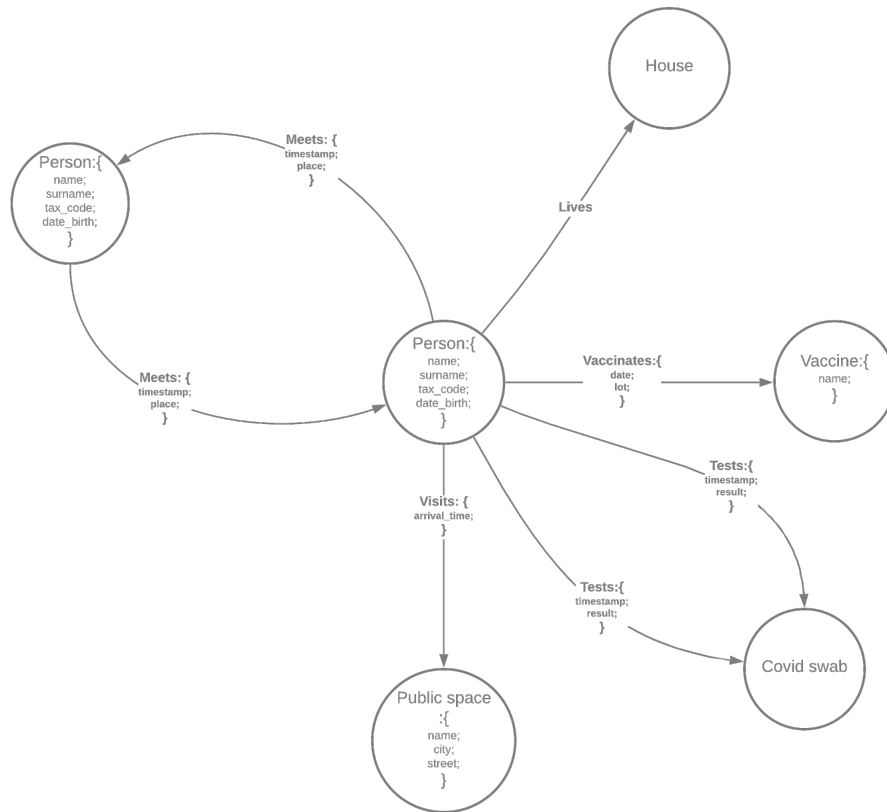
The designed ER diagram contains the following entities: Person, Public Space, Vaccine, House and Covid Test. The other classes have been introduced due to ER diagram correctness reasons, but they haven't been taken into account during the design of the Graph Database because the last one allows a more flexible design.

As said before the concept of 'Family' has been replaced with the concept of 'House' since it is more realistic and more useful for contact tracing.



### 3 Graph diagram

The designed Graph diagram is shown below.



## 4 Dataset description

We built our dataset by making a script in Python, using some useful packages like for example random-italian-people that automatically generates people with random attributes.

## 5 Queries and Commands

The commands we have designed are shown below.

The queries we have designed are shown below.

## 6 References and Sources

We have ...

## 7 Conclusion

??