

Multi-source data integration for parallel COVID-19 incidence forecast

Miguel Guzmán-Merino, Christian Durán, Maria-Cristina Marinescu, Concepción Delgado-Sanz, Diana Gomez-Barroso, Jesus Carretero and **David E. Singh**

dexposit@inf.uc3m.es

CABAHLA meeting 15/6/2022

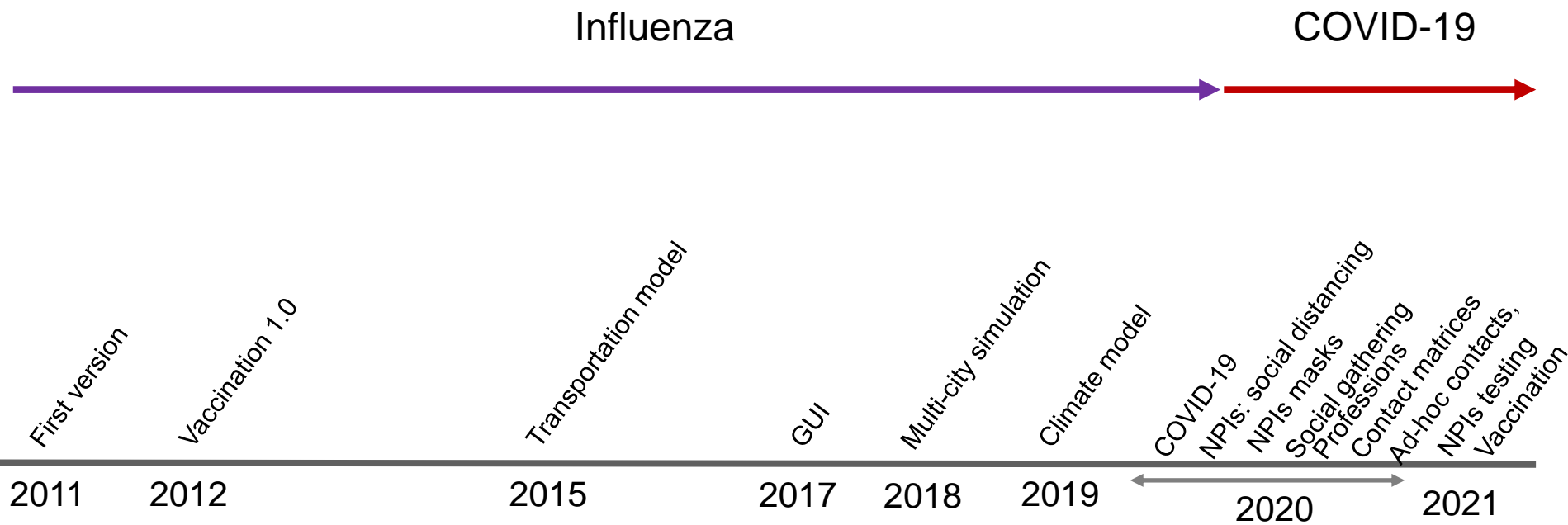


Overview

- ▶ Epigraph provides COVID-19 forecast for Spain (including all provinces)
 - ▶ Ongoing: Germany and Finland
- ▶ Forecasted information:
 - ▶ Long-term and short-term incidence and hospitalization cases
 - ▶ Evaluation of different vaccination strategies
 - ▶ Evaluation of waning immunity scenarios
 - ▶ New COVID-19 variants
- ▶ Partners (that receive the forecasted information):
 - ▶ Panel responsible of the design and implementation of the COVID-19 vaccination program for Spain (David E. Singh is member of the panel)
 - ▶ European COVID-19 Forecast Hub
 - ▶ Centro Nacional de Epidemiología
 - ▶ Spanish Ministry of Health (Vaccination Area)
 - ▶ European Centre for Disease Prevention and Control (ECDC)
 - ▶ Comunidad de Madrid
- ▶ Decision-making support
 - ▶ European Commission, Ministry of Health

Overview

- ▶ EpiGraph
 - ▶ Parallel program
 - ▶ Developed in C language with MPI
 - ▶ 20,000 code lines

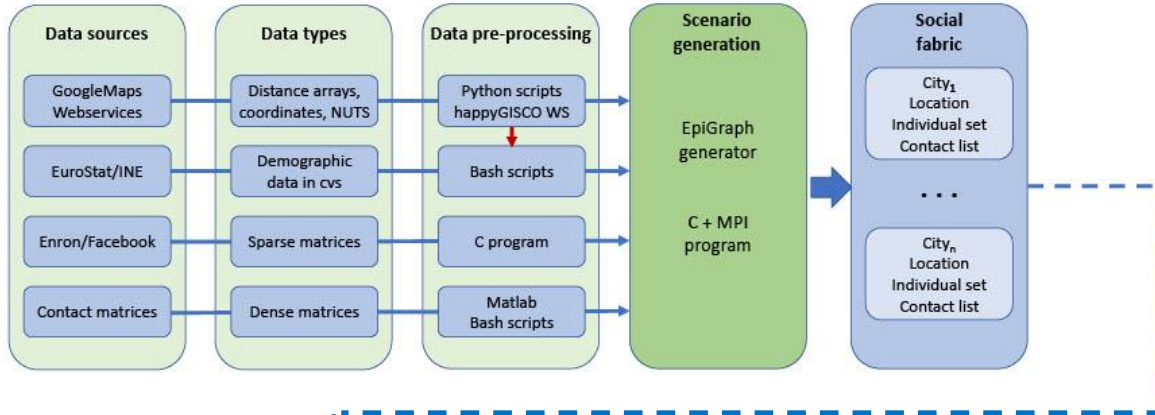


Overview

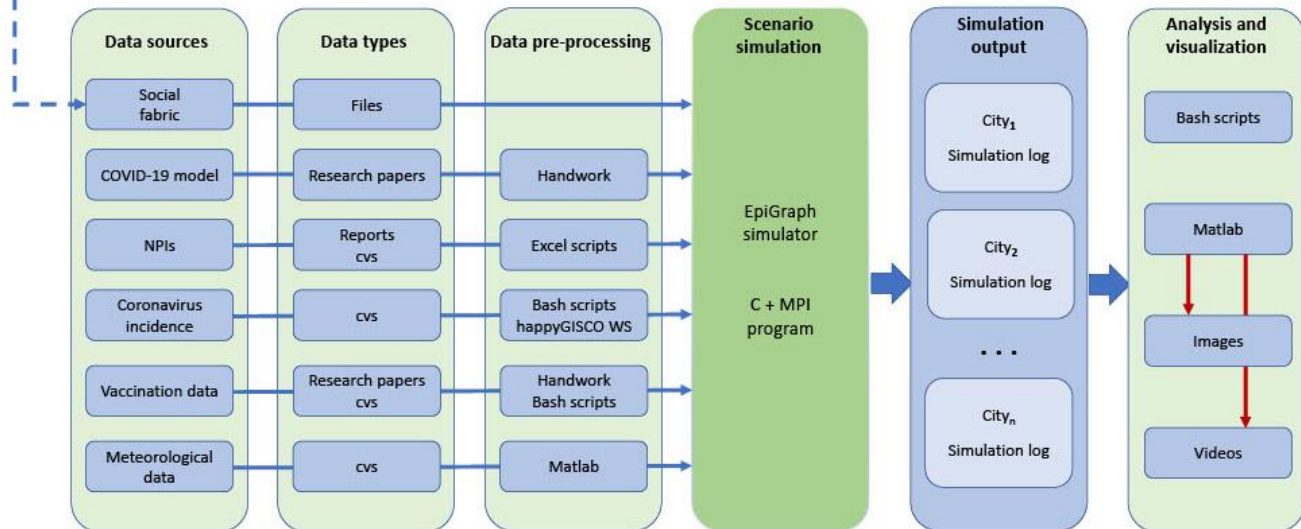
- ▶ EpiGraph
 - ▶ Parallel program
 - ▶ Developed in C language with MPI
 - ▶ 20,000 code lines
- ▶ Integrates models
 - ▶ The pathogen (influenza, COVID-19, including variants)
 - ▶ Social interactions
 - ▶ Vaccine effectiveness
 - ▶ Transportation model
 - ▶ Climate conditions
- ▶ Implements politics
 - ▶ Non pharmaceutical interventions
 - ▶ Vaccination

Data management

Scenario generation

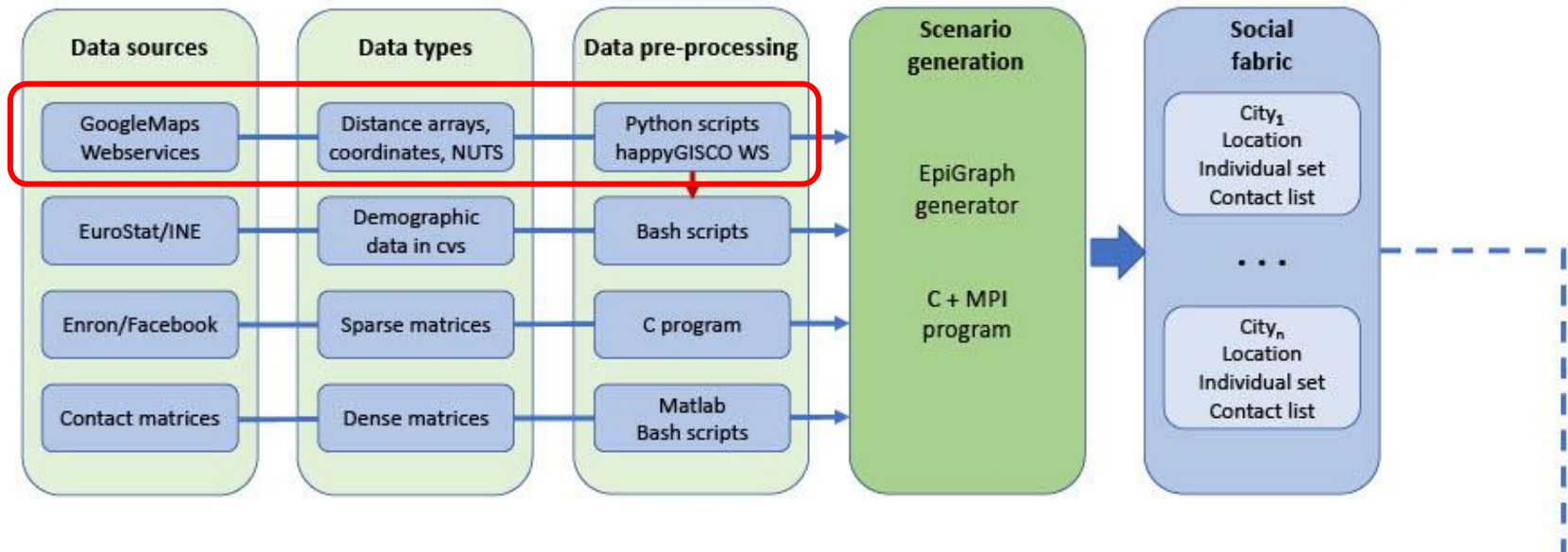


Scenario simulation



Data management

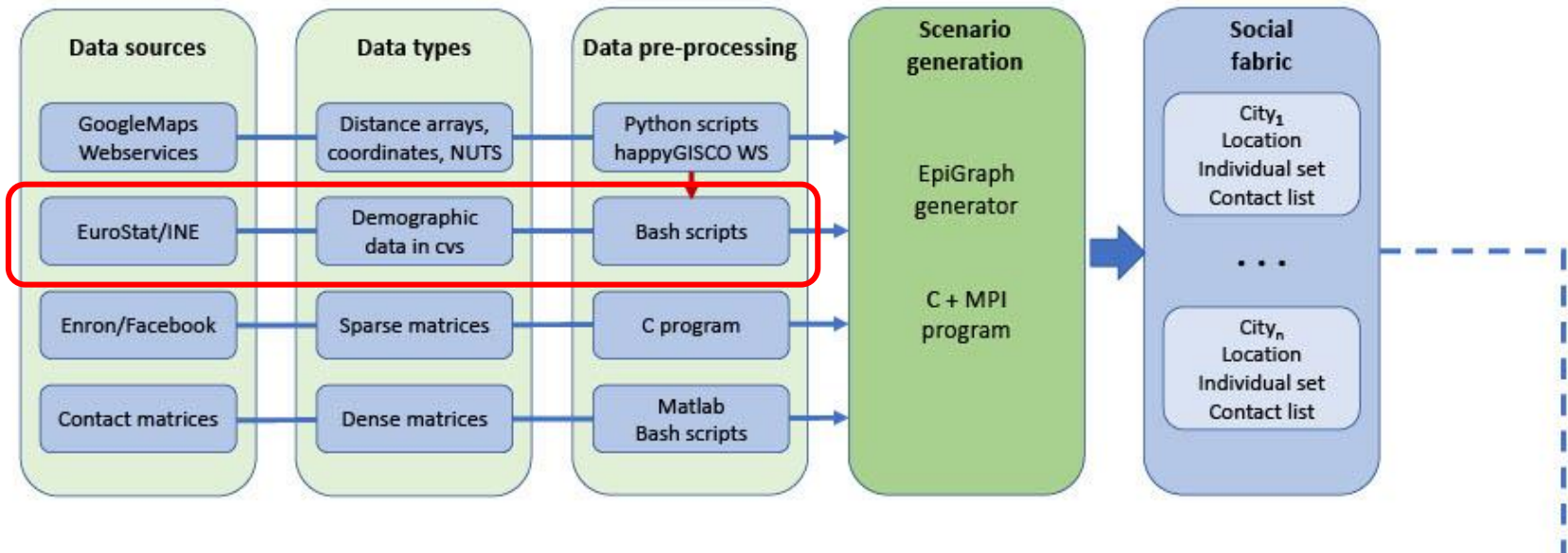
Scenario generation



- ▶ Social model built from different data sources
- ▶ City geolocation used by the transportation model and for accessing the demographic data.

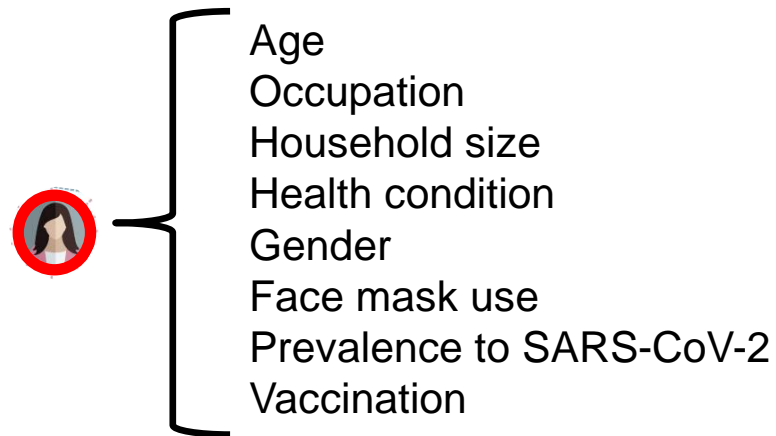
Data

Scenario generation



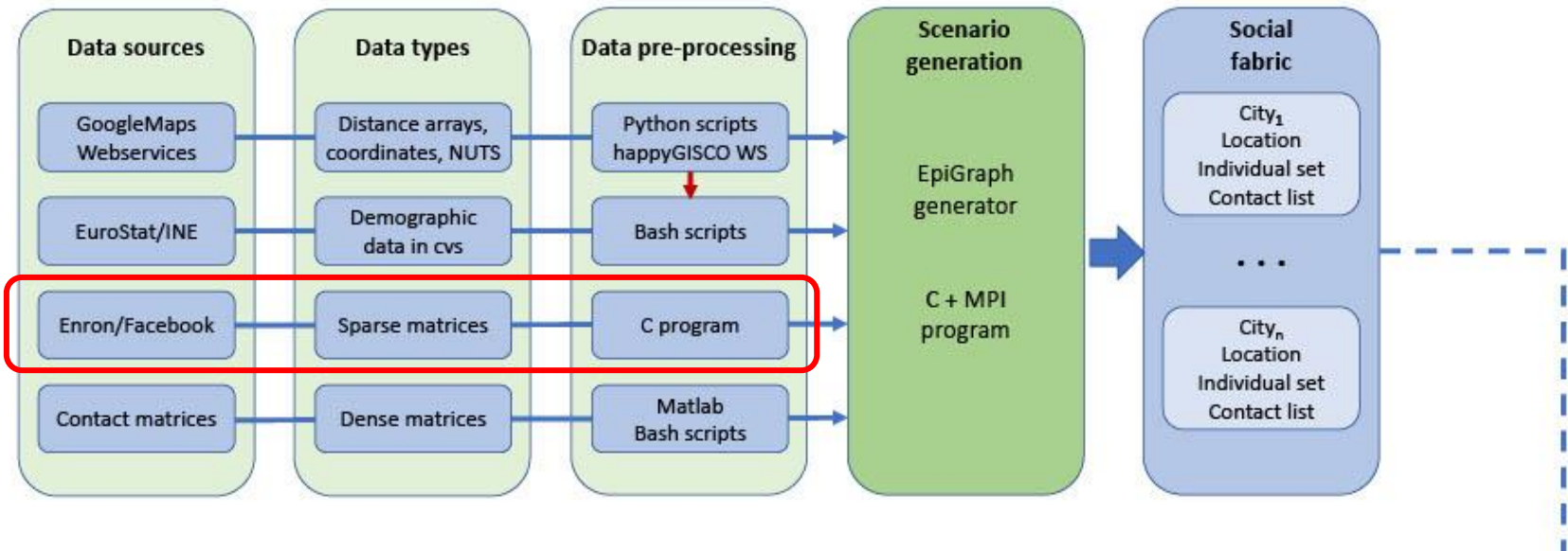
- ▶ EuroStat and Spanish Statistic Institute for modelling the population
 - ▶ Population Pyramid
 - ▶ Family size distribution
 - ▶ Job sectors

Social model



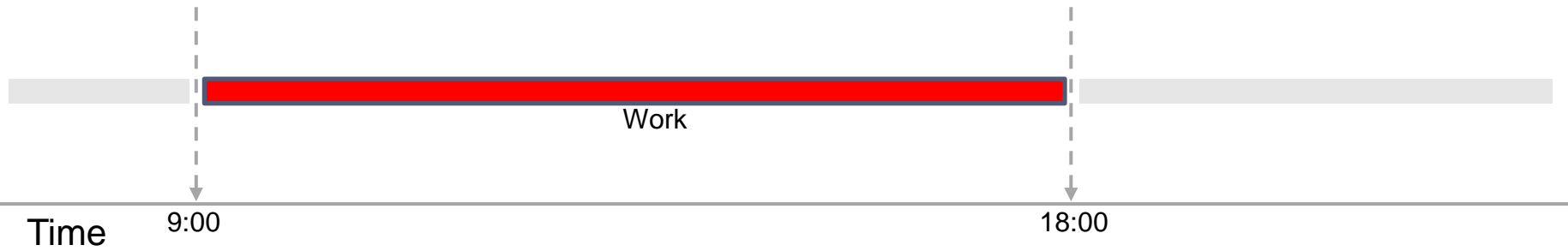
Data

Scenario generation



- ▶ Social networks are used to model the social contact patterns of each collective
 - ▶ Enron email corpus for work contacts
 - ▶ Facebook for student, elderly people and unemployed contacts

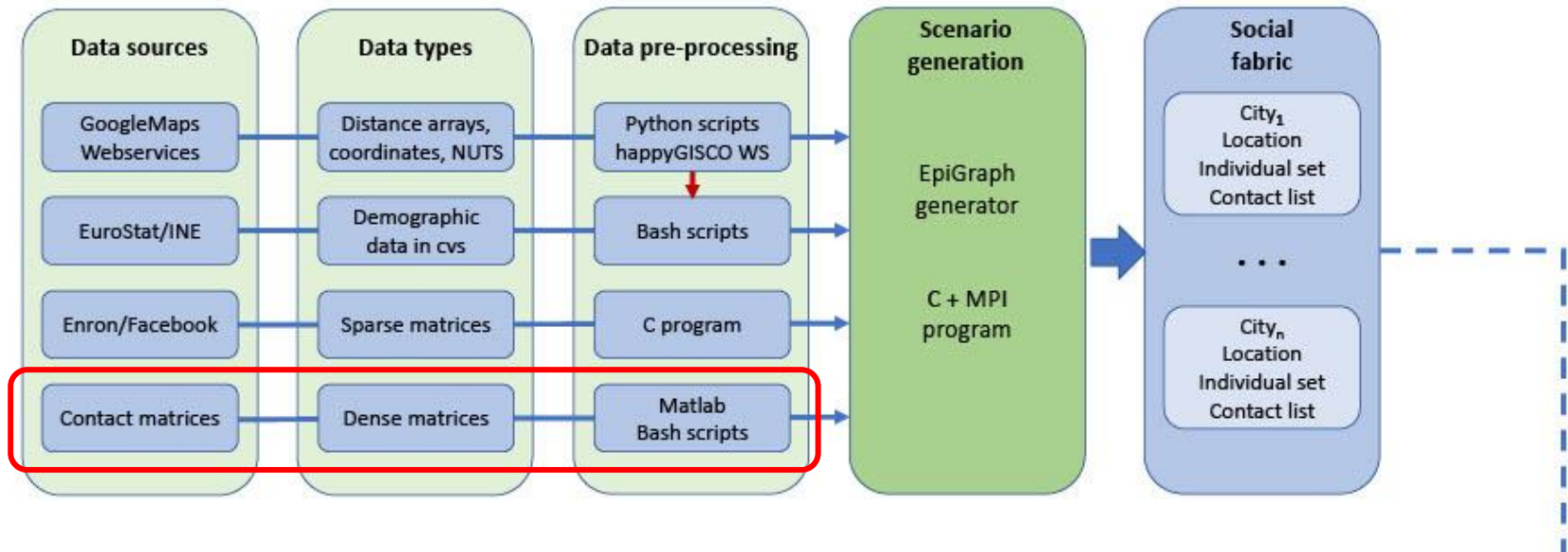
Social model



Work group

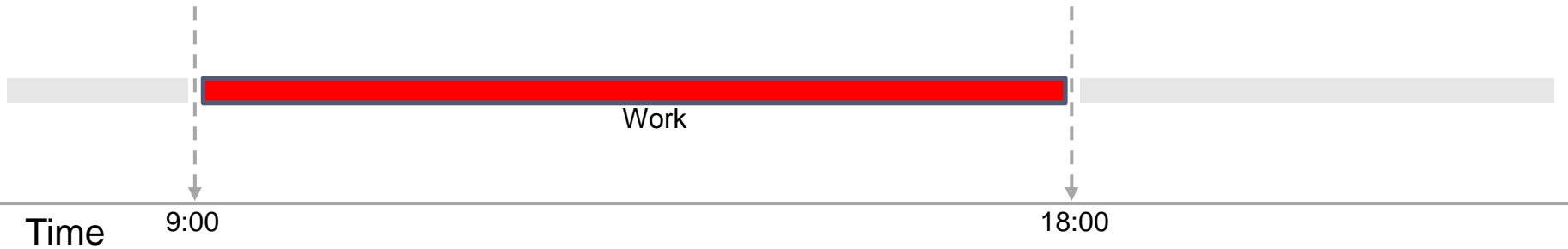
Data

Scenario generation

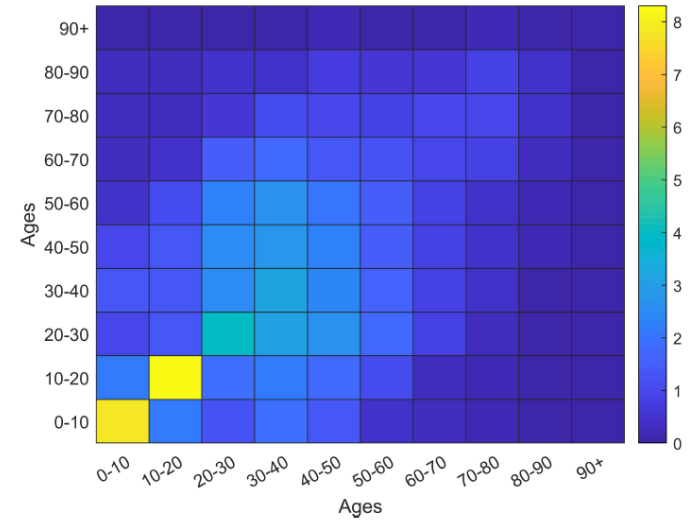


- ▶ Contact matrices represents average number of daily interactions between individuals
 - ▶ Based on age intervals
 - ▶ Separated by groups: students, workers, etc.
- ▶ Different contact matrices for each country

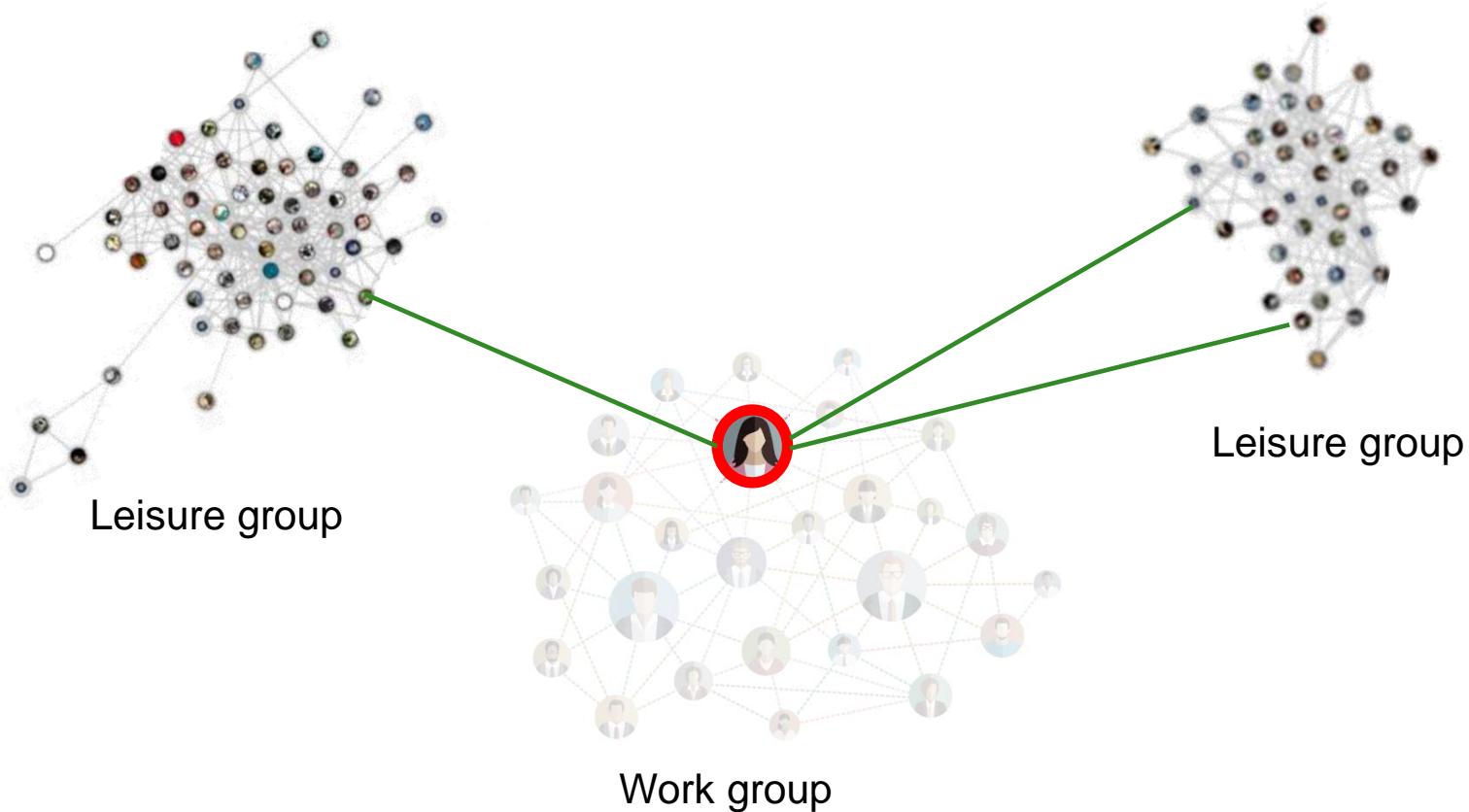
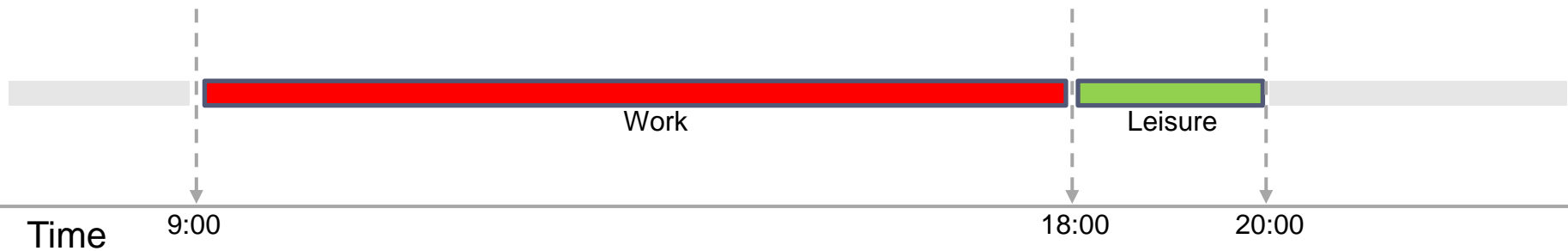
Social model



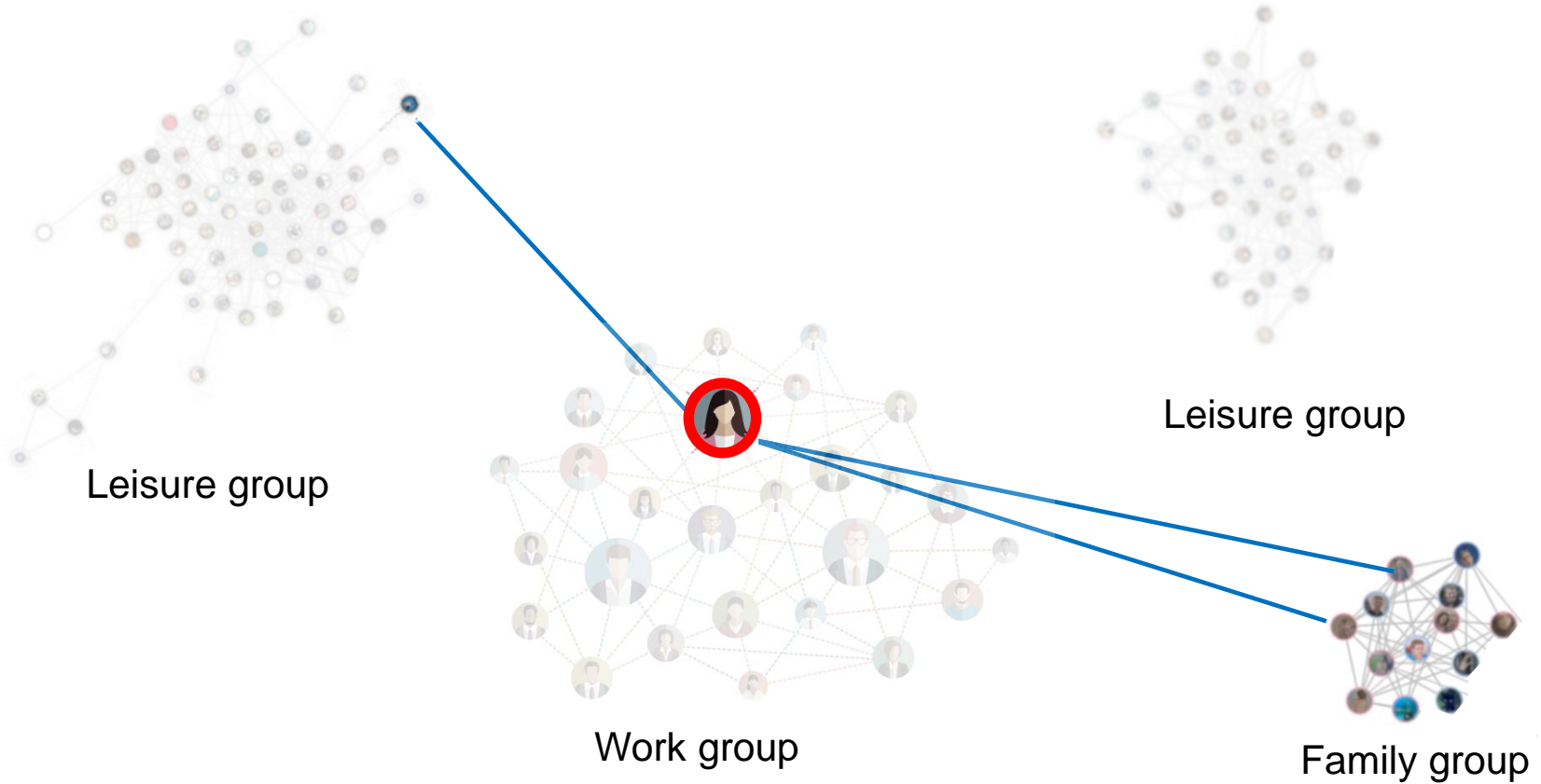
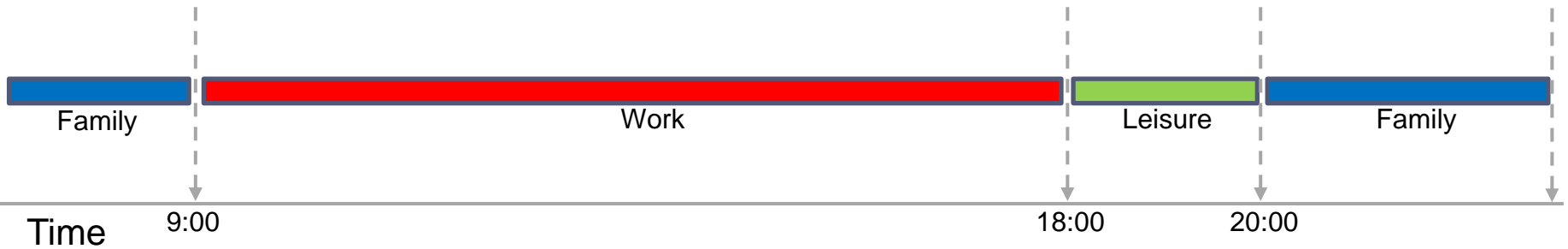
Work group



Social model



Social model

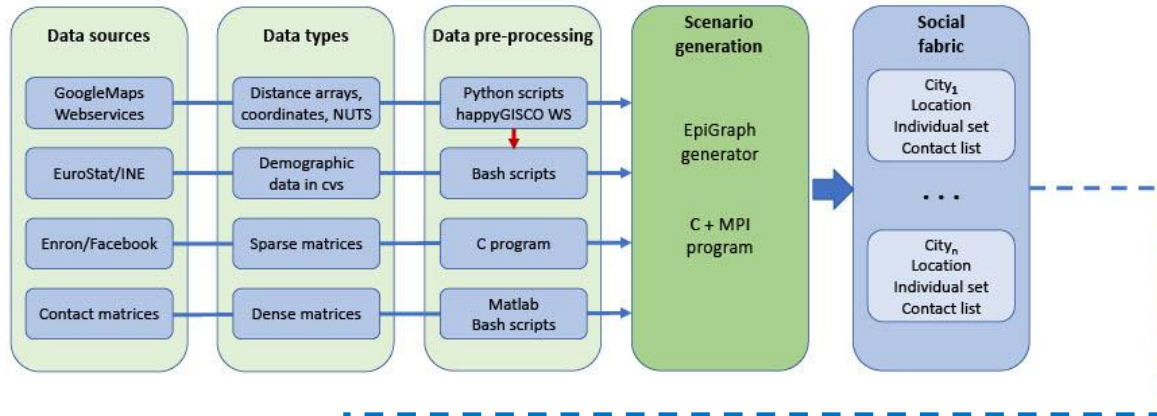


Social model

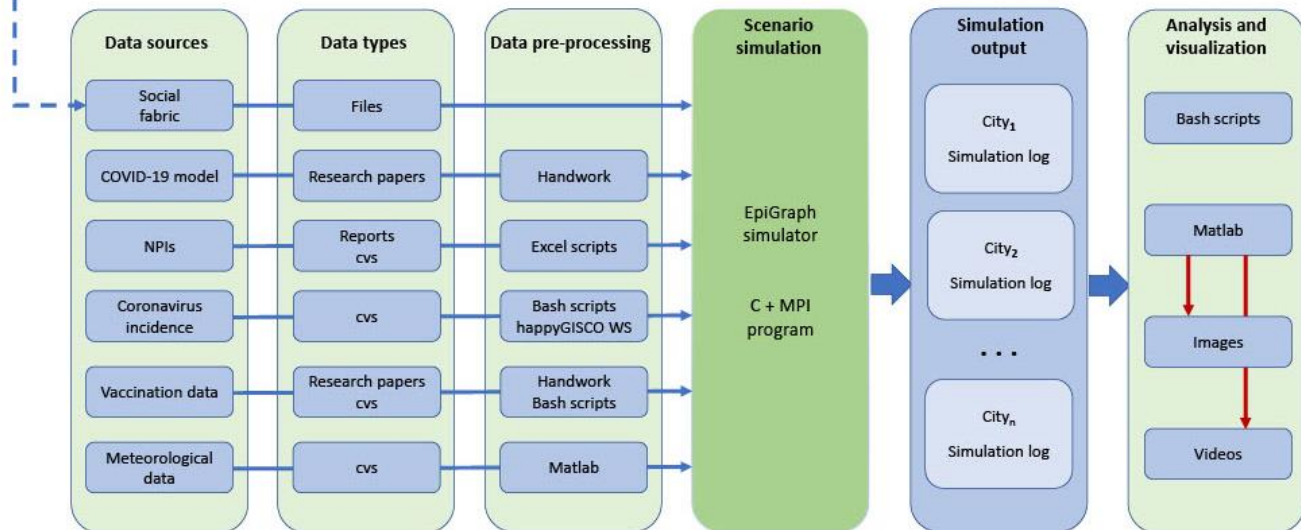
- ▶ We model different collectives:
 - ▶ Static contacts for some collectives
 - ▶ Teachers with students
 - ▶ Socio-sanitary workers with elderly people in nursing homes
 - ▶ Dynamic contacts for some collectives
 - ▶ Health professionals
 - ▶ Security forces
 - ▶ Catering workers

Data management

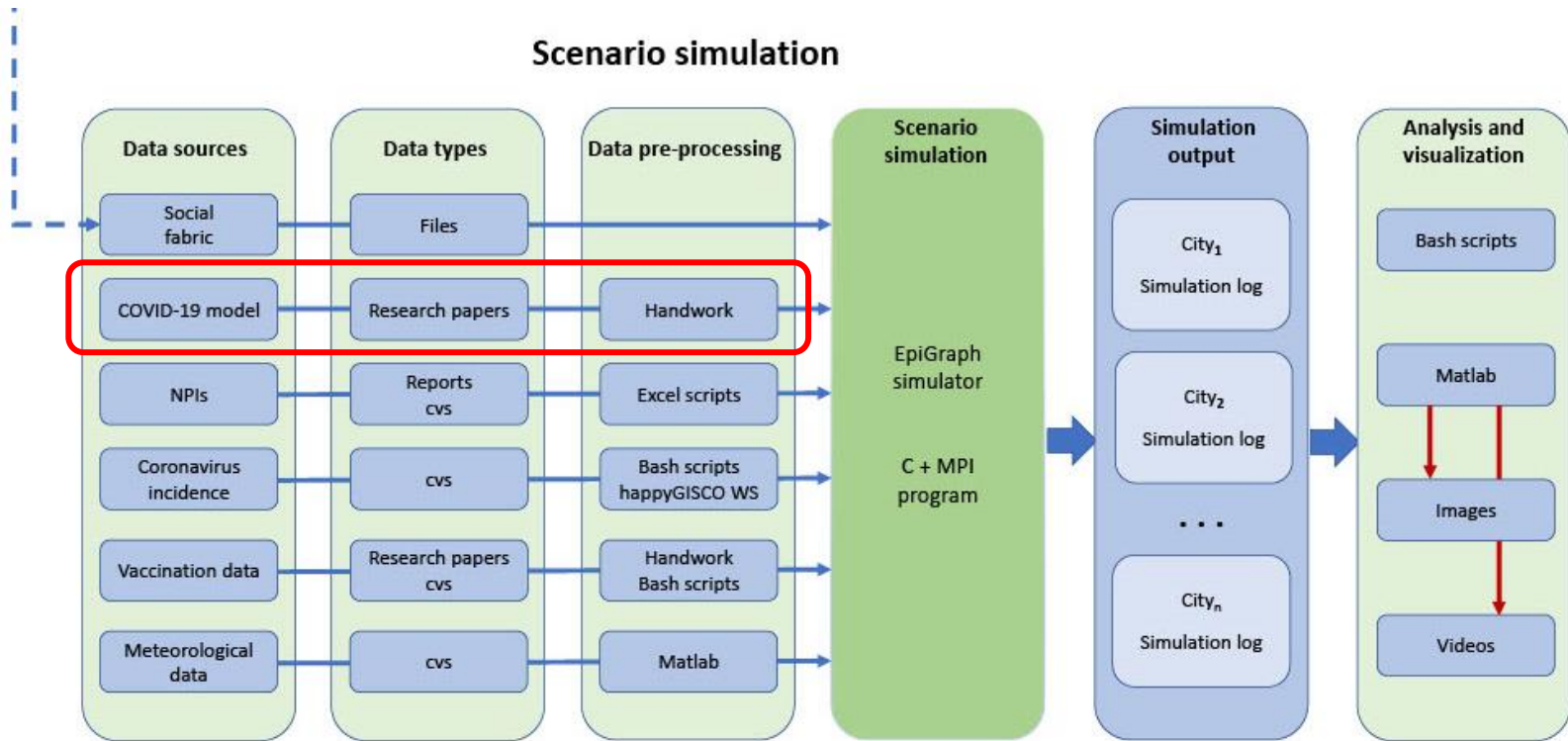
Scenario generation



Scenario simulation

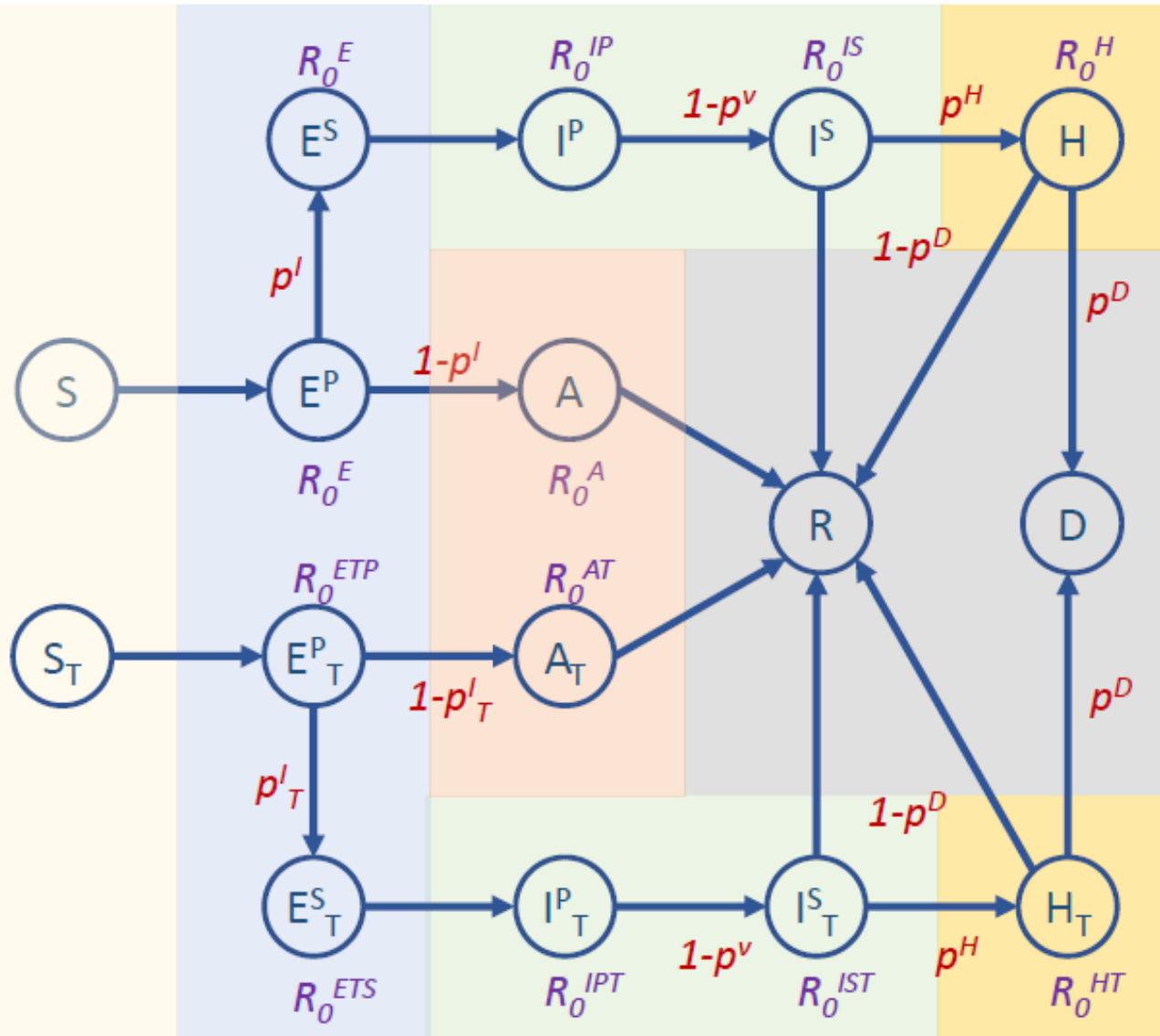


Data

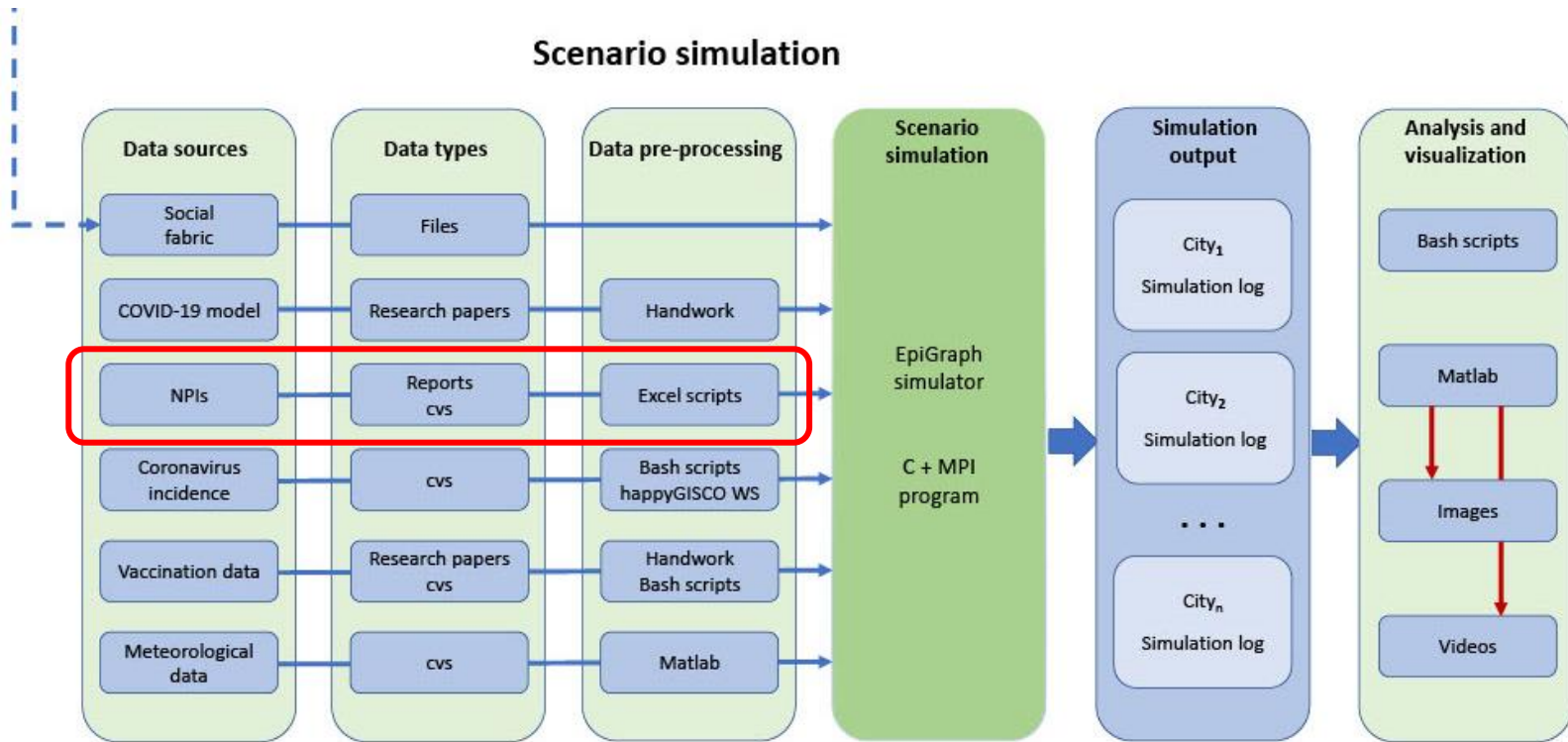


- ▶ Epidemiologic model characterizes the COVID-19 propagation.
- ▶ Calibrated based on research papers.

Epidemiologic model

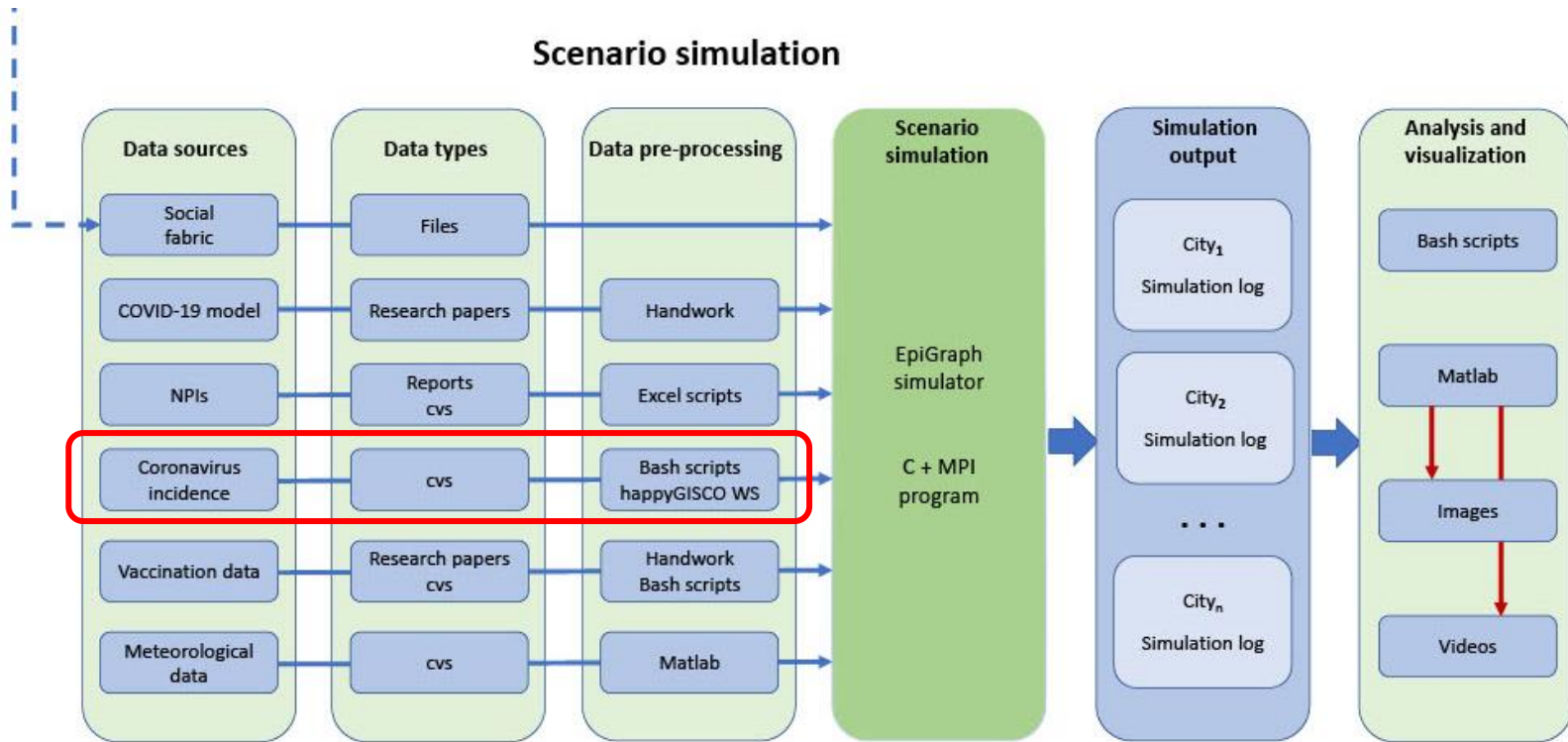


Data



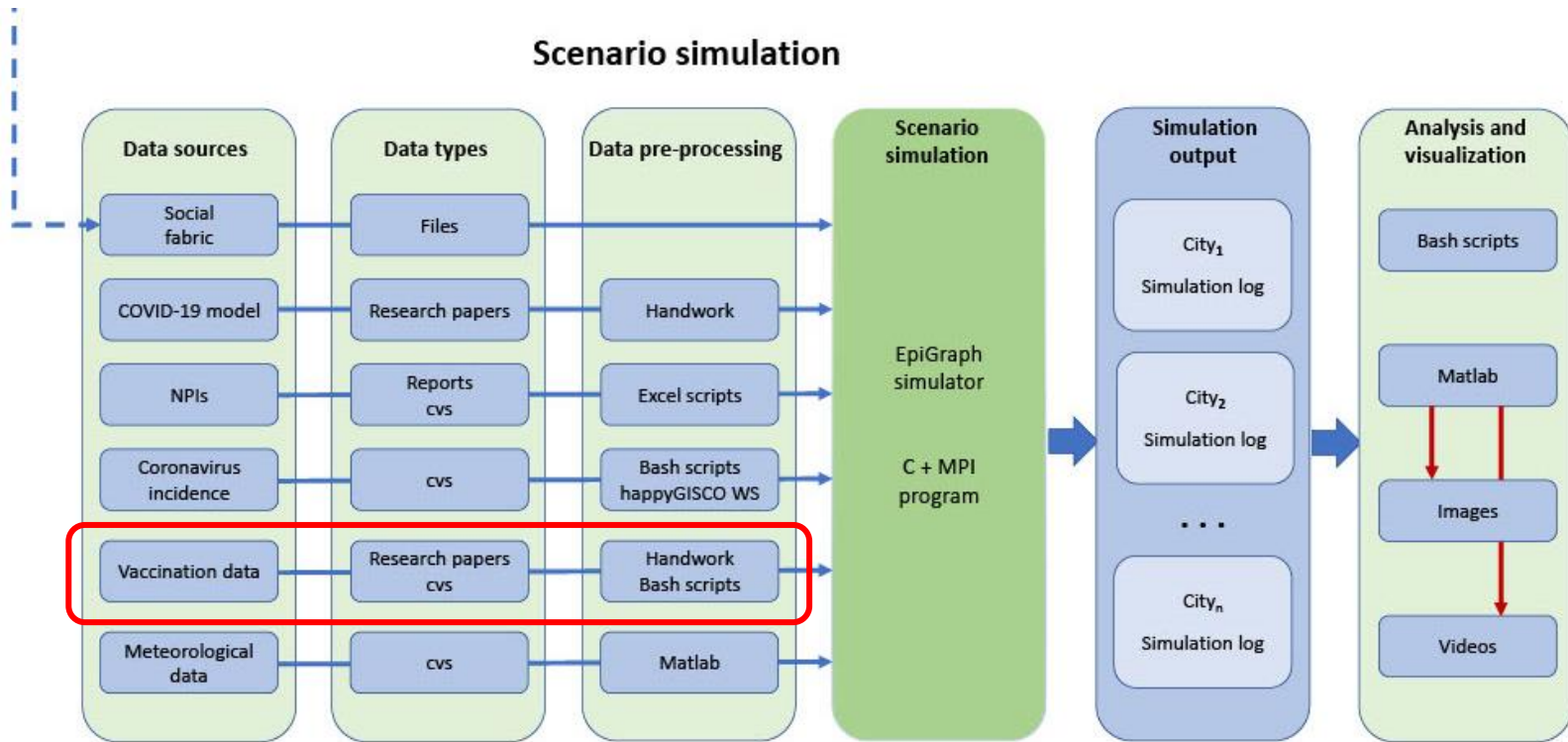
- ▶ Non-pharmaceutical intervention was collected from reports from Spanish Health Ministry and European Centre for Disease Prevention and Control

Data



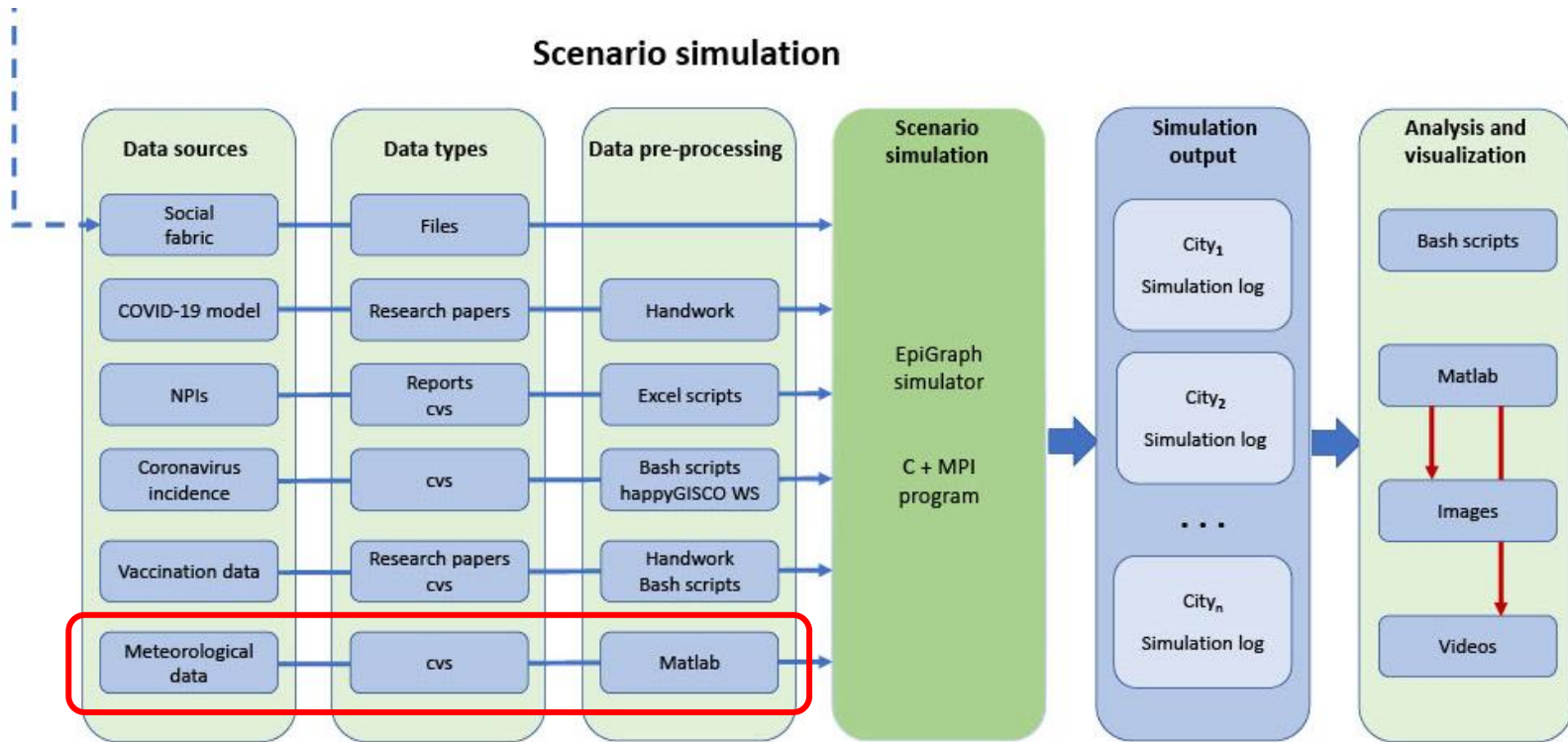
- ▶ Coronavirus incidence is used to set the initial percentage of the infected population of each city.

Data



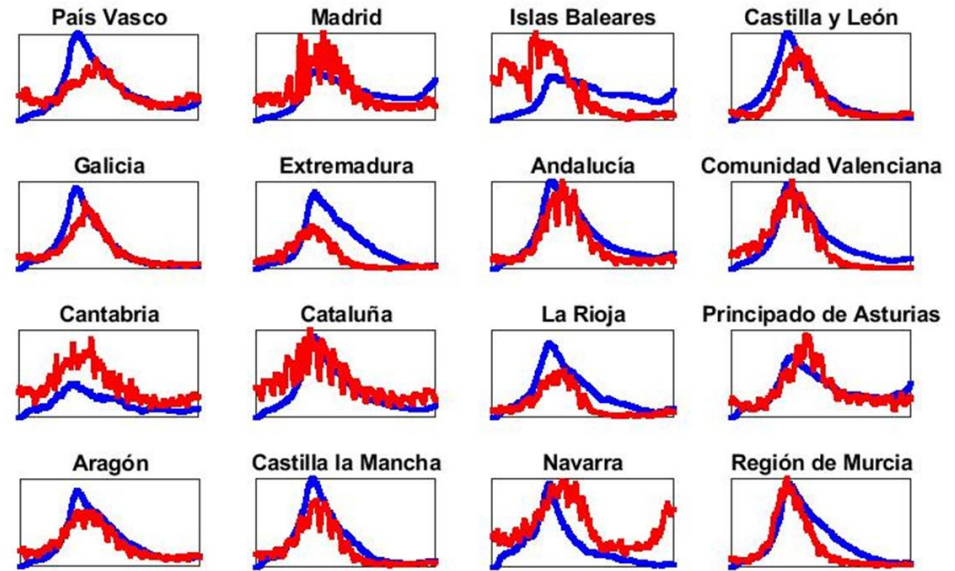
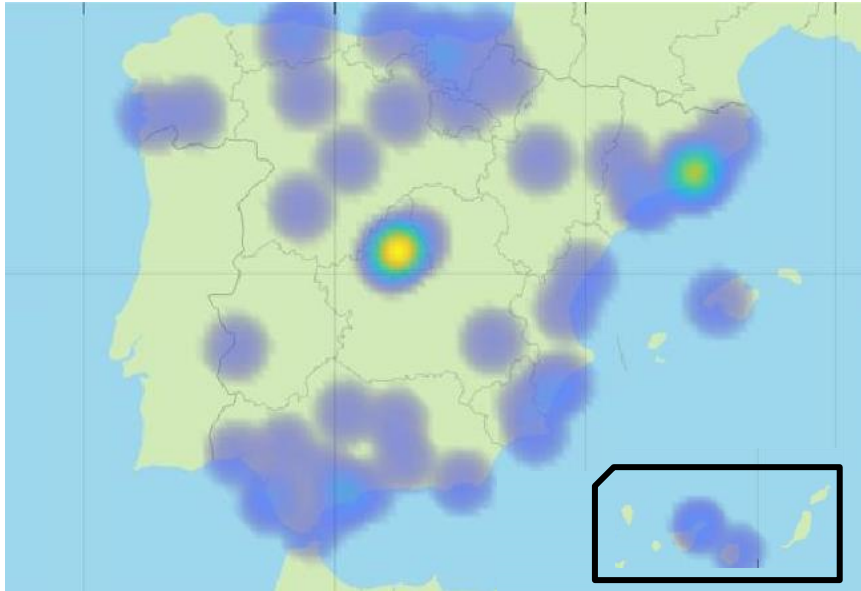
- ▶ Vaccination data includes:
 - ▶ The vaccine effectiveness model
 - ▶ The vaccination strategy

Data



- ▶ Meteorological data consisting samples taken by meteorological stations in Spain
 - ▶ Pressure, humidity and temperature.
 - ▶ 10-minute samples

Results



64 most populated Spain cities

- ▶ Median of cities of the same province
- ▶ Aggregated province data

Simulation (blue line) fit to real cases (red line) by province

Current research directions

- ▶ Vaccination strategies
 - ▶ New COVID-19 variants
- ▶ Mask use relaxation
- ▶ Evaluating the impact of social gathering events.
- ▶ European Covid-19 Forecast Hub
 - ▶ <https://covid19forecasthub.eu/>