

PTI+ Salud Global: WP3

FAIR Computational Epidemiology: FACE





Fondos de recuperación del programa NextGen EU

Total: 1M de euros.

Personal + Infraestructura

Start: July 2021

Researchers from four CSIC institutes

José Javier
Ramasco



Álvaro
López



Diego
Ramiro



Frederic
Bartumeus



Coordinators

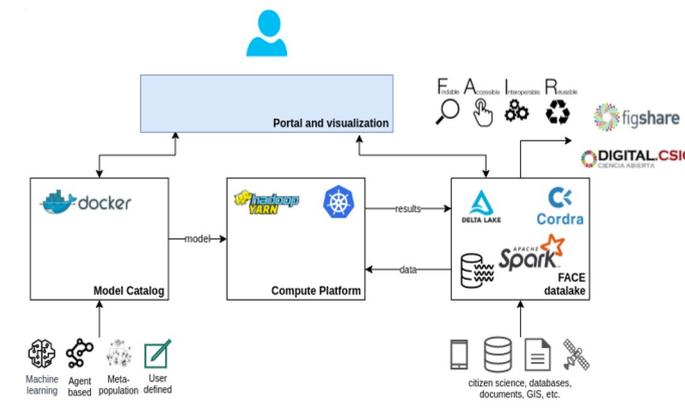
Objective 1: Build a data-driven open simulation and modeling platform that would serve as the reference system for epidemiological modelling in Spain and the EU.

Objective 2: Implementation of a data-lake access model, where data owners can store and share data assets with different accessibility and anonymity levels, with metadata-rich and linked data approach.

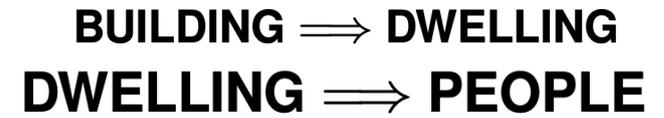
Objective 3: Provide feature-rich visualization tools, as well as reproducibility paths for the experiments results.

Objective 4: Foster an open science approach following the FAIR principles, in line with EU and national scientific policies.

- **Completado:**
 - Data-lake (DeltaLake, Spark, Hadoop y Ceph RBD)
 - Plataforma de computación basada en YARN
 - Sistema de gestión de workflows (Prefect)
- **En proceso:**
 - Migración de YARN a Kubernetes (Q3)
 - Integración de Jupyter Notebooks sobre plataforma (Q3)
 - Ingestión automática de datos (Q4)
- **Planificado:**
 - Autenticación basada en OpenID Connect (Q3)
 - Integración con catálogo de modelos y portal web (subcontratación) (Q3, Q4)
 - Visualización de datos (Q4)
 - Integración de servicio en el EOSC (2023?)



Modelos EPI, población sintética



1. INE DATA:

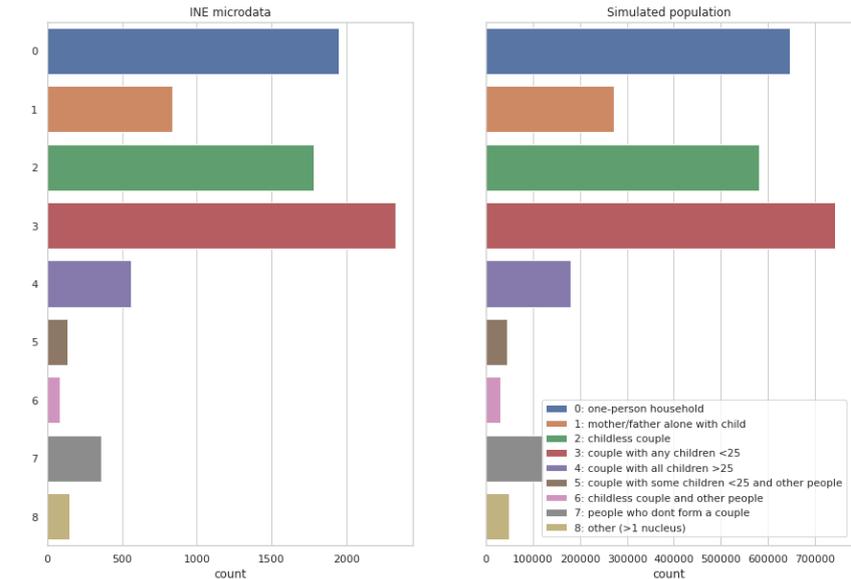
- Population of Spain (1 January 2021).
- Census 2011: Number of empty houses and population.

2. INE microdata (homes, people and EPA).

3. INSPIRE CADASTRAL (buildings).

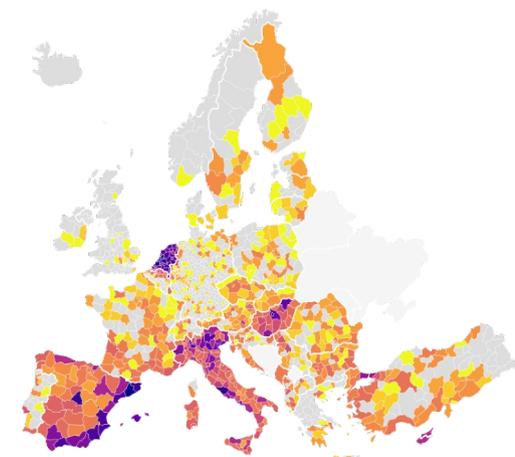


Madrid - household_type



```
/disk/dataset/Other-species/richiardii-Coqui
/disk/dataset/Not-sure/not-best/a8acdec4-ed2
/disk/dataset/Probably-Culex-sp./best/8d18ea
```

1 50 5000





THANK YOU

for your attention