

# GO-PEG project

## Generation Of cross border Pan-European Geospatial datasets and services

Stefania Morrone

64<sup>th</sup> MIG-T meeting - Friday, 22 January 2021

# GO-PEG project



3 years' project (Oct 2019 – Sept 2022) - co-funded by the Connecting Europe Facility (CEF) of the European Commission

## Consortium



ICT SMEs from Italy, Spain, Germany and Denmark, coordinated by the Katholieke Universiteit Leuven, Belgium.



in the project's Advisory Board

# GO-PEG project

## Creation of high-value, harmonised thematic open datasets

- related to environment, emergency and disaster management thematic areas
- to be provided via standard-based APIs
- to be published in the European Data Portal as **Open Data**



**Strong involvement of the stakeholders** since the very drafting of the use cases

# GO-PEG project

**Development of automated and highly reusable tools** for creation of harmonised data and relevant web services

Integration of INSPIRE and non-INSPIRE sources into multi-country datasets (focusing on cross-border sources and INSPIRE Priority Datasets) harmonised where possible to **simplified /extended INSPIRE data models**



Project expected outputs **in line with the evolution of environmental data sharing in Europe** and requirements of the Open Data Directive on availability of **High Value Datasets**.

# GO-PEG stakeholders' key role

## Contribute by:

- Drafting the use cases
- Providing data
- Giving insight on project's harmonisation workflows
- Testing and feedback on methods and datasets

## Benefit from:

- feedback on own data and harmonization workflows
- harmonised data related to the use cases of interest
- documentation on innovative harmonisation methods and tools
- working together with other stakeholders
- Getting to know **OGC APIs** as **INSPIRE download services**

**Still in time to get on board !**



**ISPRA**

Istituto Superiore per la Protezione  
e la Ricerca Ambientale



**Sistema Nazionale  
per la Protezione  
dell'Ambiente**



Agency for  
Data Supply and  
Efficiency



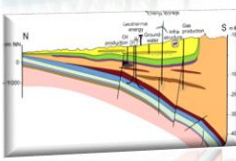
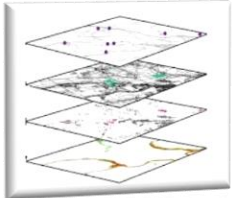
**IH cantabria**

INSTITUTO DE HIDRÁULICA AMBIENTAL  
UNIVERSIDAD DE CANTABRIA



# Overview of GO-PEG use cases

- Combine population statistics and buildings data to calculate the impact on the population of an emergency event (POPIMPACT)
- Preserve and secure the European drinking and fresh water supplies (DRWAD)
- Support adaption of forest management to climate change (FutureForst)
- Organize and deliver subsurface information for a sustainable and safe use of natural resources (GO\_DEPTH)
- Leverage geospatial data and cutting-edge technologies for simpler data access and better documented COVID-19 data (geo-COVID Watch)





# Overview of GO-PEG use cases



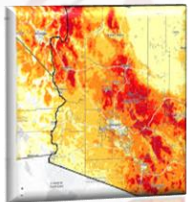
- Develop accurate intermodal routing tool improving reliability of Estimated Time of Arrival and reducing CO2 and other emissions (IntermodalRouting)



- Trace the flow of contaminated water to define possible affected areas and to try to stop the progress of contamination (Tracing)

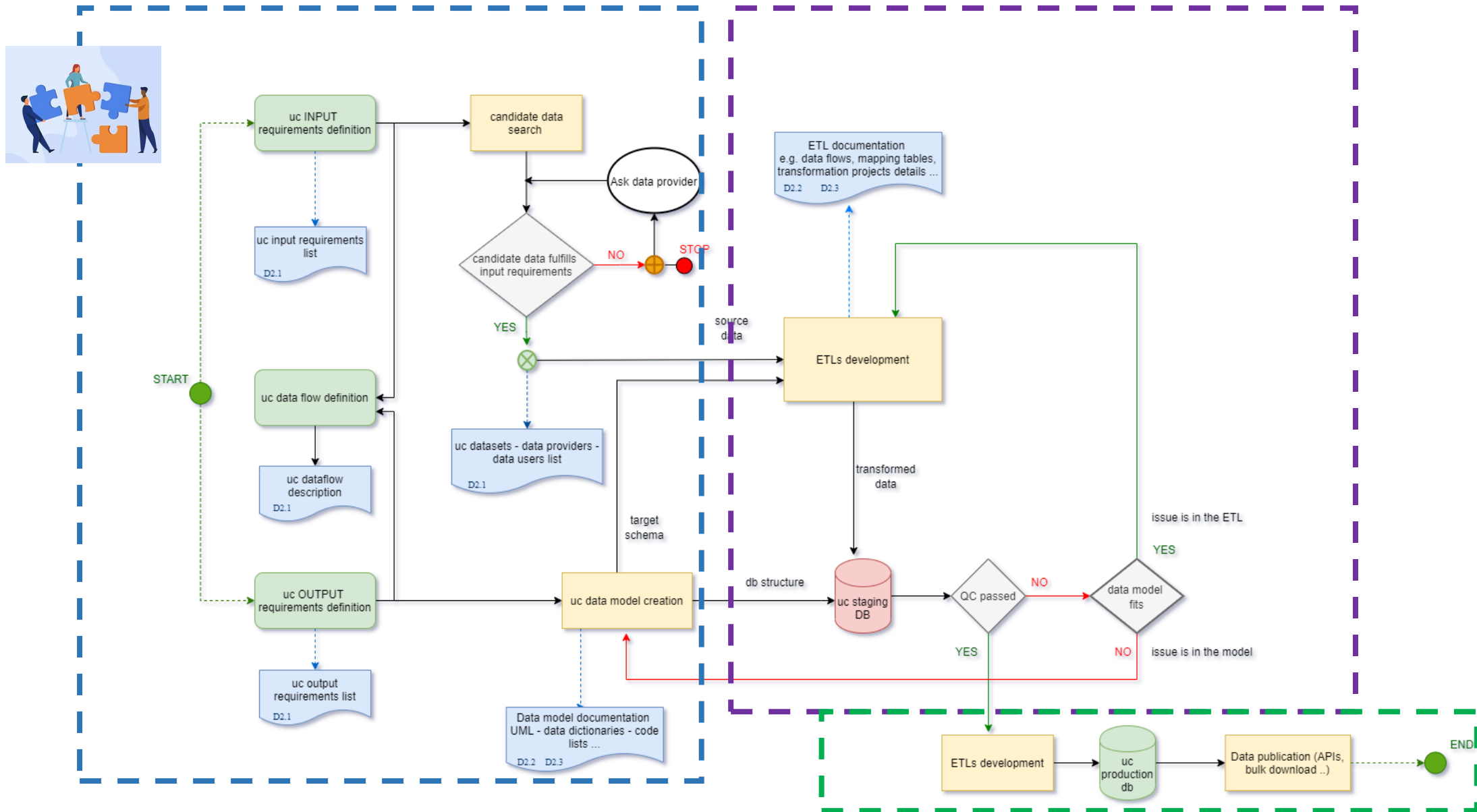


- Provide high-quality no fly zones for drones based on harmonised spatial reference data augmented with AI-based detection of objects in aerial photos (FAIRPORT)



- Calculation of the vulnerability due to potential risk of wildfire (POP\_WILDFIRE)

# Use Case Workflow





# The geoCOVID-Watch Use Case

The aim:

Contribute to a common understanding of the pandemic and its impacts leveraging on

- Geospatial data standards
- State-of-the-Art technologies

Provide **simpler access**

- to **better documented** COVID-19 data
- using the mechanisms for **direct access to measurement data** developed in the environmental sector in recent years

# The geoCOVID-Watch Use Case

The how:

Create a repository of known initiatives providing data relevant to COVID-19:

- Expose information on initiatives available through **OGC API -Features** (*OGC standard to create, modify and query features on the Web*)

Create an aggregator collecting COVID-19 related data across Europe

- Expose COVID-19 data via **SensorThingsAPI (STA)** (*OGC standard to interconnect IoT (Internet of Things) sensing devices, data, and applications over the Web*)

Create a registry service for COVID-19 concepts and metrics

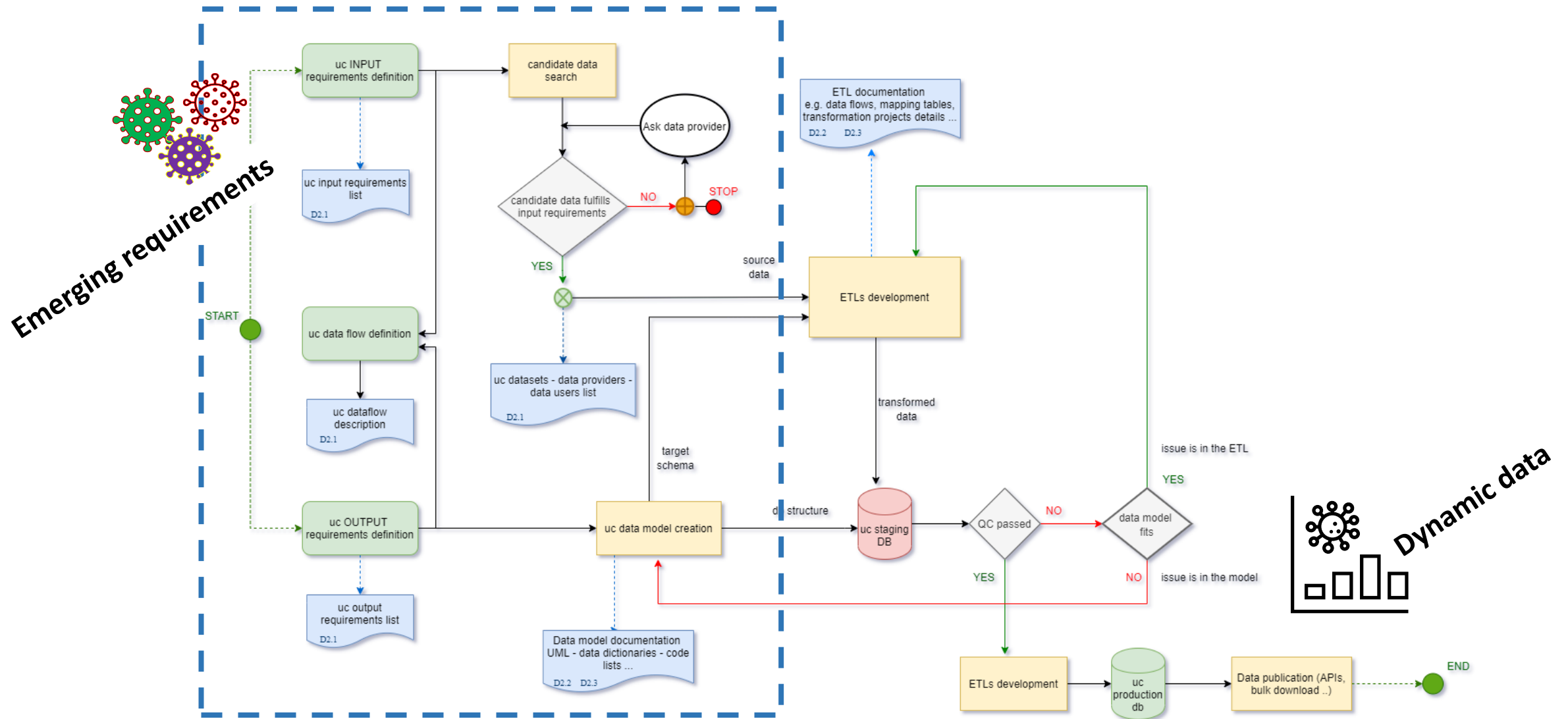
- Expose COVID-19 related definitions via **Re3gistry** instance

# Collaboration with the **API4INSPIRE** project

## Build on what is already available



# COVID-19 data landscape continuously evolving !



# COVID-19 data sharing: current Status

We live in a world of **big data**, with **artificial intelligence** and machine learning permeating all aspects of our lives. We have **smart factories** and **smart cities**; we have self-driving cars and machines trained to exhibit human intelligence... **BUT**

**Excel / CSV** are still most used for COVID-19 **data collection & sharing !!!!**

**Excel: Why using Microsoft's tool caused Covid-19 results to be lost**

**Covid: how Excel may have caused loss of 16,000 test results in England**

**Im Propagandaministerium | Die Corona-Hauptstadt**

**Actual cases Vienna: ~600**

**Nationally reported cases Vienna: ~800**

Gesundheitsstadtrat Peter Hacker (SPÖ) (Bild: APA/HERBERT NEUBAUER)

# Current Status

## Lack of agreement:

### - *statistical metrics*

...died *OF* Covid-19 **vs** died *WITH* Covid-19?

### - *data aggregation process*

**little provenance data** + often **unclear methodology** for data derivation

### - *country response measures*

**comparability** of national testing strategies  
...test all first level contacts or just put them in quarantine?

## Lack of transparency!



EUROPEAN COMMISSION  
DIRECTORATE-GENERAL FOR HEALTH AND FOOD SAFETY  
Public health, country knowledge, crisis management  
Health Security and Vaccination

There is a great diversity between countries concerning the implementation of testing strategies for **asymptomatic cases**<sup>8</sup>.

### EU health preparedness:

#### Recommendations for a common EU testing approach for COVID-19

Agreed by the Health Security Committee on 17 September 2020

Concerning the testing of people admitted to hospitals or requiring specific treatments, nine countries (AT, DE, DK, FR, IE, LV, PT, SK, and BIH) test all **admitted patients** to hospitals, regardless of whether they are displaying symptoms or not.



# Potential Solutions

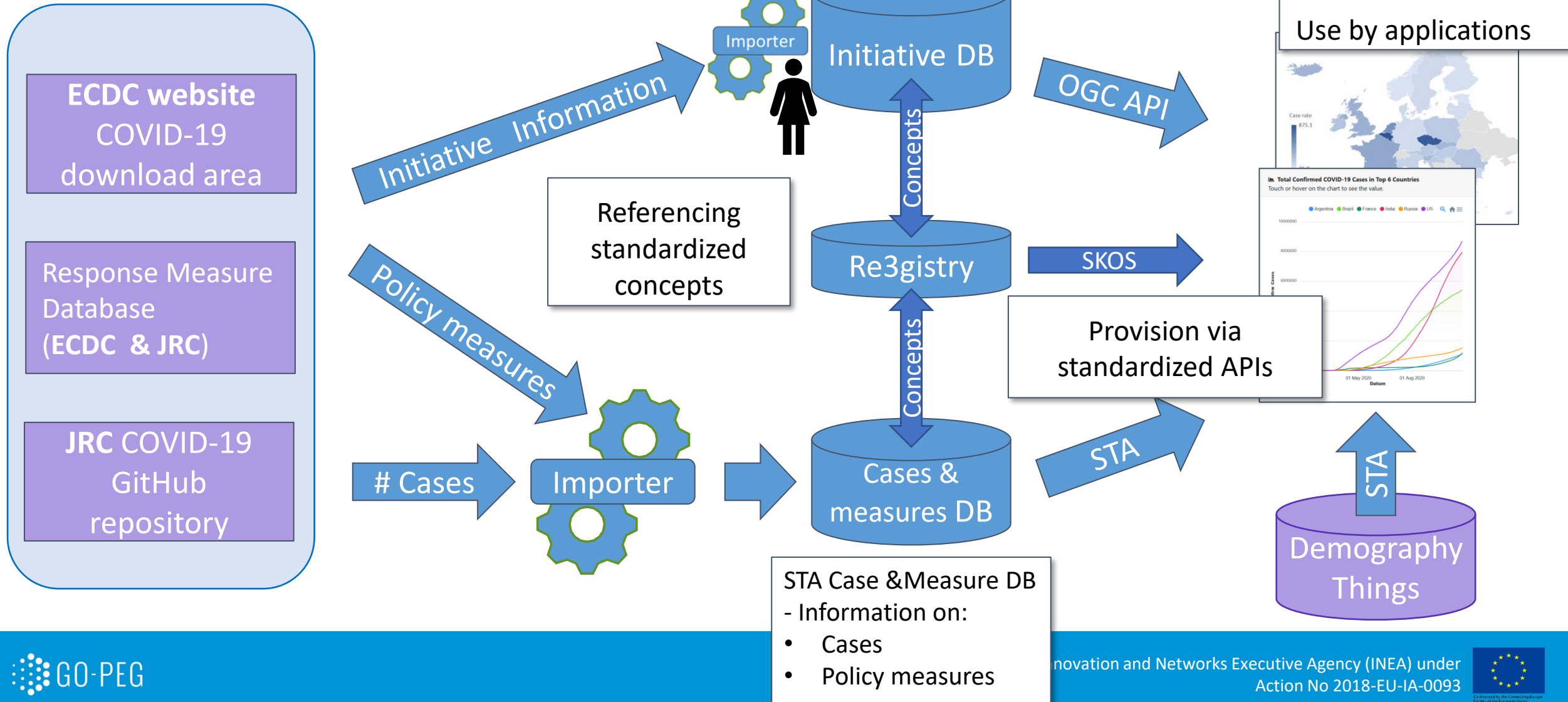
## COVID-19 Brief: The role of data science, technology, and AI in infectious disease tracking

September 29, 2020 | COVID-19



- Define clean & automated data flows
  - Clear governance
  - SOTA technology (NOT Excel!)
  - Leverage standards
- Clear concepts and metrics
  - Central codelist of definitions
  - Agreed calculation of metrics
- Enrich with additional data sources to make data comparable
  - Demography /Population
  - Hospital Beds
- Expose data using cutting edge technology
  - Open data provision via APIs
  - Enable alternative portals based on authoritative information
  - Leverage AI / BI

# geoCOVID Watch Data Flows



# Containing the Complexity

- Identify newly emerging metrics:  
**geoCOVID-Watch Initiative DB**  
*OGC API – Features*
- Provide and publicize precise definitions per metric:  
**geoCOVID-Watch Concept Registry**  
*Re3gistry*
- Provide access to real-time data on cases and measures:

**Linked**

```
graph TD; A[Identify newly emerging metrics: geoCOVID-Watch Initiative DB] --> B[Provide and publicize precise definitions per metric: geoCOVID-Watch Concept Registry]; B --> C[Provide access to real-time data on cases and measures:];
```

# Data mirrored via OGC SensorThings API

**Case data (daily update)** imported from JRC COVID-19 GitHub repository:

*National statistics*

<https://github.com/ec-jrc/COVID-19/blob/master/data-by-country/jrc-covid-19-countries-latest.csv>

*(TBD) Regional statistics:*

<https://github.com/ec-jrc/COVID-19/blob/master/data-by-region/jrc-covid-19-regions-latest.csv>

## **National Containment measures (weekly update)**

*currently imported from “download COVID-19 dataset” area of the ECDC website*

[https://www.ecdc.europa.eu/sites/default/files/documents/response\\_graphs\\_data\\_2.csv](https://www.ecdc.europa.eu/sites/default/files/documents/response_graphs_data_2.csv)

Under investigation the possibility to add data from the **Response Measures Database (RMD)** of the ECDC and the JRC :

<https://covid-statistics.jrc.ec.europa.eu/Home/DownloadCsv>

# Re3gistry instance for statistical metrics and response measures

- Using a registry service (instead of current pdf files to be searched for in different websites) will **increase coherence, transparency and interoperability** of the relevant data.
- Definitions will be assigned **persistent and unique URL**, and existing resources such as electronically available pdf documents will be linked
- **Re3gistry** instance will consider the [Response Measures Database](#) three-level hierarchical coding system (Levels 1, 2 and 3) and will reference different sources

# Initiatives in the different countries as OGC API - Features

Collect data on initiatives related to COVID-19 pandemic by main EU health institutions /organisations, and on the different statistical metrics and containment measures addressed.

Initiative DB will assure that:

- the concepts being provided via the registry will be aligned with the data provision requirements of the underlying data,
- **emerging concepts will be identified and standardized** (concept alignment)

Examples from [ECDC Covid-19 website](#).



# geoCovid Watch

- Initiatives with all their details: <https://service.datacatalog.ca/Initiative/items?f=ar>
- Vocabularies with the details: <https://service.datacatalog.ca/Vocabulary/items?f=ar>
- Individual Indicators: <https://service.datacatalog.ca/Indicator/items?f=ar>

```
type: "FeatureCollection"
  features:
    0:
      type: "Feature"
      id: "COVID-19 Data Hub"
      geometry: null
      properties:
        @featureType: "Initiative"
        prj_code: 6
        name: "COVID-19 Data Hub"
        website:
          @href: "https://covid19datahub.io"
          status: "ongoing"
          project_leader: "Institute for Data Valorization IVADO, Canada"
          startdate: "2020-04-19Z"
          description: "Data Aggregator with many Variables"
          geographicalscope: "World"
        conditions_access_use: "GPL-3: Cite Guidotti and Ardia (2020) https://doi.org/10.21105/joss.02376; Add the URL https://covid19datahub.io"
        endpoints:
          0:
            id: 11
            url:
              @href: "https://covid19datahub.io"
            type: "downloadable file"
            description: "Data Aggregator with many Variables"
            prj_code: 6
            dataType:
              0:
                dataType: "Primary Cases"
              1:
                dataType: "Policy"
            accessMode:
              0:
                accessMode: "Download_File"
            outputFormat:
              0:
                outputFormat: "CSV"
            indicator:
              0:
                name: "internal_movement_restrictions"
```

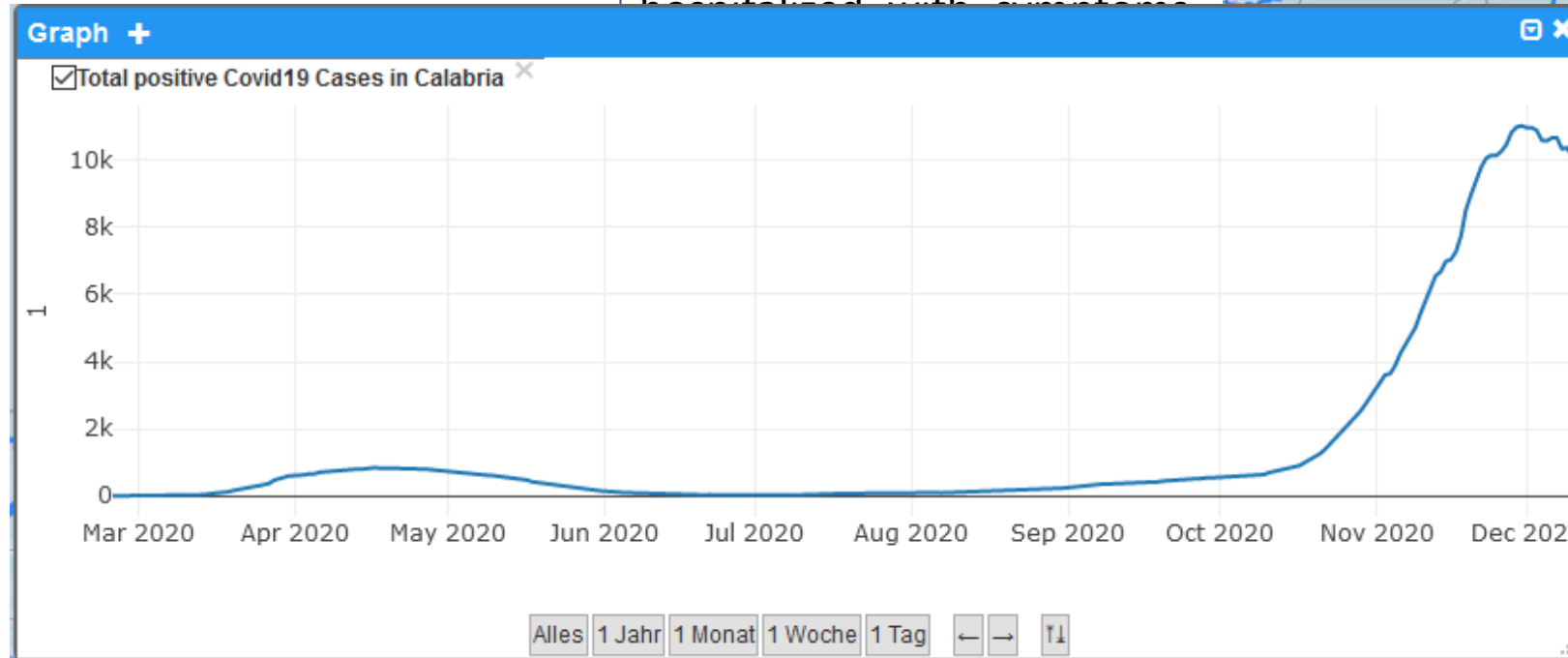
# Example: Italian Case Data - SensorThings API

- Data from Italian Civil Protection
- NUTS Levels 0, 2 & 3
- Various Indicators:

swabs

total\_cases

hospitalized\_with\_symptoms



- Population Count age 15-64, gender female
- Population Count age 15-64, gender male
- Population Count age 15-64, gender total
- Population Count age <15, gender female
- Population Count age <15, gender male
- Population Count age <15, gender total
- Population Count age >65, gender female
- Population Count age >65, gender male
- Population Count age >65, gender total
- Population Count age total, gender female
- Population Count age total, gender male
- Population Count age total, gender total
- Population Count age unknown, gender female
- Population Count age unknown, gender male
- Population Count age unknown, gender total
- Population Density
- cases\_from\_screening
- cases\_from\_suspected\_diagnostic
- cases\_tested
- deceased
- home\_isolation
- hospitalized\_with\_symptoms
- intensive\_care
- new\_positives
- recovered\_healed
- swabs
- total\_cases
- total\_positive
- total\_positive\_variation
- XX Plot all Datastreams

# Example: New Cases & Po

[|  |   |
|--|---|
| type:  | "FeatureCollection"                               |
| ▼ features:                                  |   |
| ▼ 0:   |   |
| type:  | "Feature"   |
| ▼ properties:                                |   |
| name:  | "Toscana"   |
| Datastreams/0/name:                          | "total population of age total in Toscana \(ITI1\)" |
| Datastreams/0/Observations/0/phenomenonTime: | "2019-01-01T00:00:00.000Z"                        |
| Datastreams/0/Observations/0/result:         | 3729641   |
| Datastreams/0/ObservedProperty/name:         | "Population Count age total, gender total"        |
| Datastreams/1/name:                          | "Total positive Covid19 Cases in Toscana"         |
| Datastreams/1/Observations/0/phenomenonTime: | "2021-01-17T16:00:00.000Z"                        |
| Datastreams/1/Observations/0/result:         | 8204  |
| Datastreams/1/ObservedProperty/name:         | "total\_positive"                                  |
| ▶ geometry:                                  | {...}   |
| ▼ 1:   |   |
| type:  | "Feature"   |
| ▼ properties:                                |   |
| name:  | "Emilia-Romagna"                                  |
| ▼ Datastreams/0/name:                        |   |
| Datastreams/0/Observations/0/phenomenonTime: | "2019-01-01T00:00:00.000Z"                        |
| Datastreams/0/Observations/0/result:         | 4459477   |
| Datastreams/0/ObservedProperty/name:         | "Population Count age total, gender total"        |
| Datastreams/1/name:                          | "Total positive Covid19 Cases in Emilia-Romagna"  |
| Datastreams/1/Observations/0/phenomenonTime: | "2021-01-17T16:00:00.000Z"                        |
| Datastreams/1/Observations/0/result:         | 54831   |
| Datastreams/1/ObservedProperty/name:         | "total\_positive"                                  |
| ▶ geometry:                                  | {...}   |
| ▼ 2:   |   |
| type:  | "Feature"   |
| ▼ properties:                                |   |
| name:  | "Umbria"  |](https://demography.k8s.$filter=(properties/count)$select=name&$expand=ObservedProperty/name'Population Count age to$expand=Observations($orderBy=phenomenonObservedProperty($select$resultFormat=GeoJSON</a></p></div><div data-bbox=)

# Outlook



- Exemplary requests for standard use cases will be provided
- Can be easily modified by users for specific purposes

More information available at:

<https://datacoveeu.github.io/API4INSPIRE/sensorthingsapi/1Home.html>



# THANK YOU

<https://www.go-peg.eu/>

