EO4EU: AI-augmented ecosystem for Earth Observation data accessibility with Extended reality User Interfaces for Service and data exploitation

Stefano Natali (Sistema) & Marco Folegani (MEEO)

eurac research

Institute of Atmospheric Pollution Research
National Research Council of Italy

European Commission

EO4EU - brief intro

- EO4EU aims to further improve access to the EU EO data offered by a variety of platforms and data repositories. Data sources include Copernicus services and associate platforms like the DIAS, but also upcoming initiatives like Destination Earth (DestinE). Without prior knowledge about their structure and format, the platform shall be able to retrieve, process, fuse and deliver new datasets, supported by machine learning algorithms and advanced semantic annotations.
- An innovative ML-based learned compression algorithms shall enhance the accessibility of the data sources by reducing the respective data volumes needs to be transferred over the network, reducing the footprint of storage capacity and the network bandwidth requirements.
- The control and core plane components of the EO4EU platform adapts an event-driven and microservices-based architecture, hosted on the Platform as a Service (PaaS) Tier. PaaS platforms often integrate with Kubernetes to simplify microservices development, scaling, and management.
- User Experience are further enhanced with a set of visualization services and interfaces, including a multi-layered user interface (GUI) for visual analytics coupled with a Workflow Editor, a Command Line Interface (CLI), and a respective Application Programming Interface (API), and an extended reality (XR) interface to further boost the usability and the adoption of the platform, combining traditional access methods with cutting-edge technology stack
- All platform communications are handled through a message-based middleware (via a Message Bus). This provides a coherent communication model with distribution, replication, reliability, availability, redundancy, backup, consistency, and services across distributed heterogeneous systems.

EO4EU Partners

ENGINEERING · ECMWF · COMMpla Multichannel Communication Platforms and Online Solutions · Trust-IT Services Communicating ICT to markets · damans · IES Solutions · Latvijas Universitāte Anno 1919 University of Latvia · Hes·so Haute Ecole Spécialisée de Suisse occidentale University of Applied Sciences and Arts Western Switzerland · cmcc · novelcore Research & Innovation · CINECA · eBOS · FMI · MEEO Meteorological Environmental Earth Observation · Vilnius University · KEMEA · Fraunhofer IVI · SISTEMA Environmental Information Mining

EO4EU architectural bird's-eye view

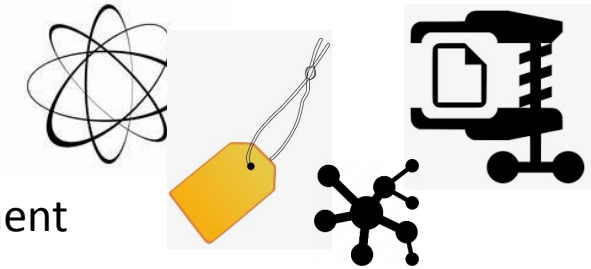
Orchestration Language

EO4EU applications

- Dashboards
- API
- Data processing workflows
- XR

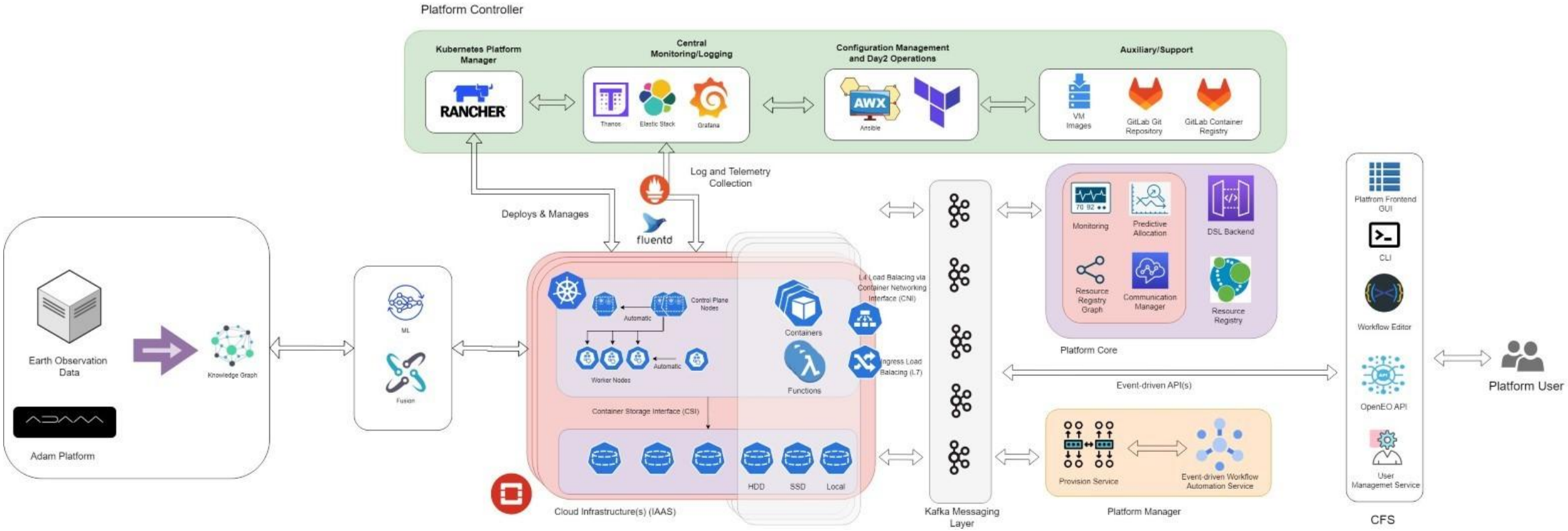
EO4EU processing infrastructure

- Fusion
- Semantic annotation
- Compression
- Knowledge management



EO sources

Desing architecture



Data Tier

A set of data sources is the input of the platform. Heterogeneous data that need pre-processing with the help of a Knowledge Graph.

- **Data Sources**
  - Interlink heterogeneous data sources (different type formats) with the EO4EU ecosystem through Open APIs (e.g. Climate Data Store API for historical occurrence of extreme weather events).
  - Access to historical and daily EO datasets.
  - Access to real time data collections streamlines (for live connections with devices and applications).
  - Access to open access cohorts of the EC through cloud-based platforms established to provide centralized access to Copernicus data, as well as to GEOSS, INSPIRE, DestinE, Galileo/ EGNOS programmes.
  - Access to open datasets and services provided by ECMWF.

Data Tier

Knowledge Graph-based Decision Making

- A Graph-Based Text Representation is introduced.
- This approach enables the extraction of informative features, structural or textual, for each entity related to the whole knowledge graph.
- For structure-related features, graph measures or indices such as common neighbors, preferential attachment and Adamic Adar indexing will be used.
- For text-related features, graph similarity techniques including graph neural networks and graph kernels will be used.
- By establishing a link prediction pipeline, EO4EU focuses on predicting possible relationship types between nodes of a knowledge graph.

ML Tier - Fusion engine

Provides all machine learning models in a toolbox for the post processing of the retrieved or fused data.

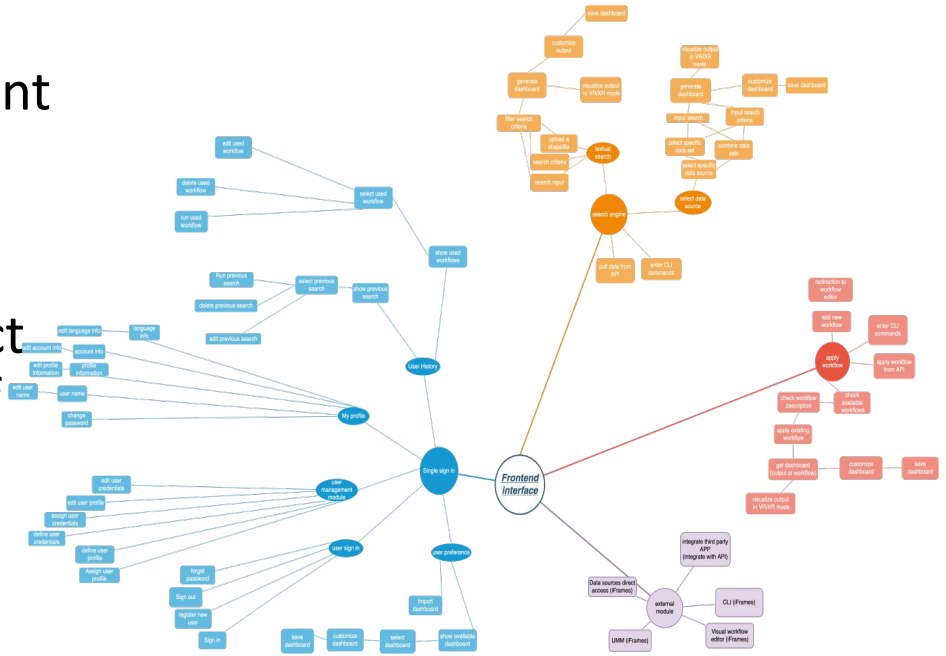
Generic Machine Learning pipeline for semantic annotation

- Minimize the requirement of use-case specific labelled data.
- The Generic ML pipeline of EO4EU will enable the learning of a robust and transferable representation of the input data in a latent space, in an unsupervised way.
- Rely on approaches such as SimCLR, which is a simple framework for contrastive learning of visual representations.
- Explore other approaches such (e.g., Mocov2 and any potential new work in this rapidly evolving area of research).
- Such a representation will provide the input representation over which the downstream tasks will operate for the learning of task specific models.

Front-end Tier - Dashboard - Data Analytics Visualization

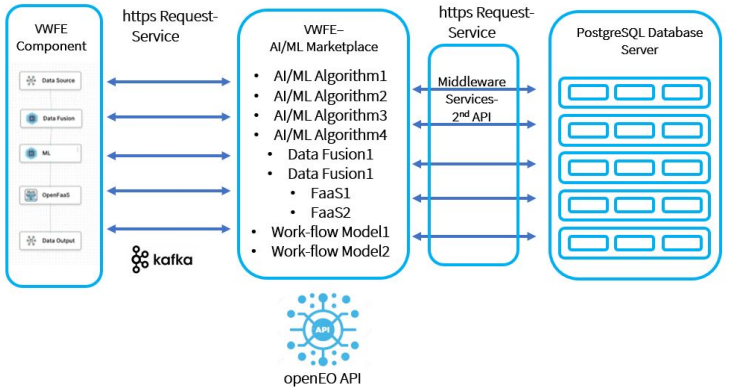
Provides multi-dimensional User Interface-UI (Web, XR, CLI, API) that enables the user to interact and control the platform.

- Decision-making and policy-maker
- Real-time data analytics and interpretation of environment observations
- Real-time mapping and interactions
- Smart Search Engine based on Text or Annotation – Select Data - Smart Category-Type-Parameters Search Engine of the searched item
- Dashboard creation
- Web XR/VR Visualization methods
- Data Analytics based on statistical metrics

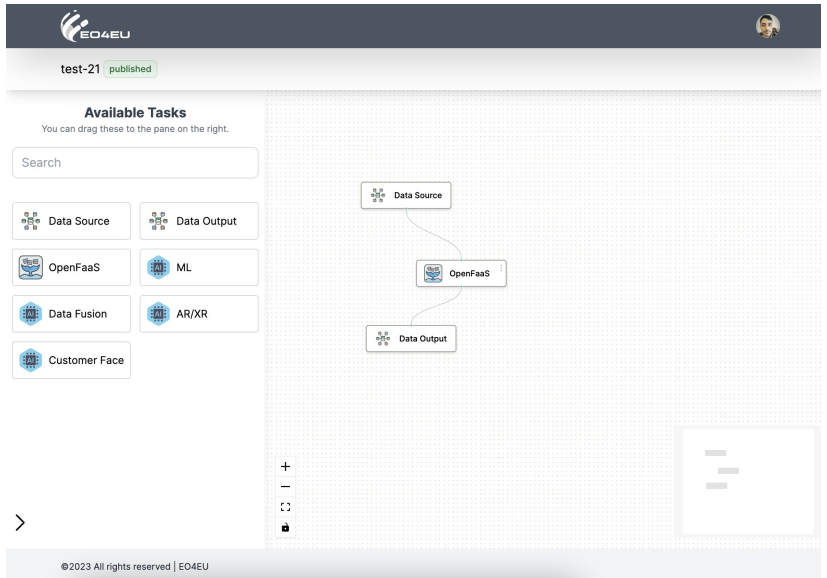


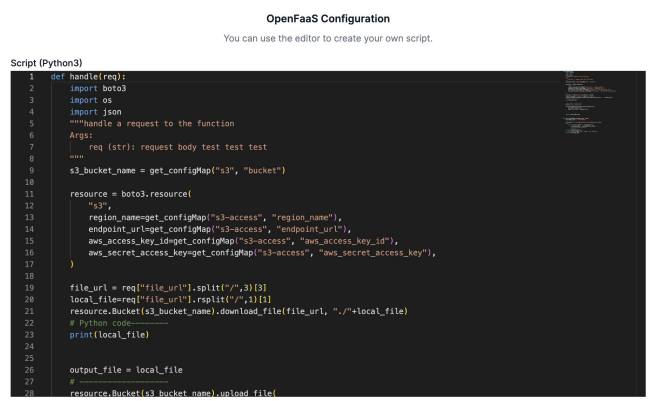
Front-end Tier - AI/ML Marketplace

- AI/ML Models-Algorithms-Techniques
- Metadata
- Data Models for Processing and Communication from Block to Block
- Programming Code
- Configuration Files
- Documentation



Building processing workflows





EO4EU Use Cases (UCs)

#1 - EO for personalized health care services EU to Globe, focus in Latvia, Lithuania

#2 - Ocean monitoring Far East-N. Europe, N. Europe-S. America

#3 - Food security Italy & other

#4 - Forest ecosystems Austria

#5 - Soil erosion Italy

#6 - Environmental pests West and East Africa, Middle East, India

#7 - Improving Civ. Prot. activities Italy (Sicily)

UC#1: EO for Personalized Health care Services

Expanding the capacity of the PASYFO model, the first-ever operational symptom forecasting model that includes a mobile application. The PASYFO components will be developed into full-scale working systems (TRL 8).

Input data	Spatial extent	Temporal extent
Citizen self-reported allergic symptoms for each region in Europe and the globe	Europe, a collection of individual points of citizen observations	Daily updated 1-day citizen reports
CAMS and SILAM AQ predictions	Europe: 10 km, Globe, ~20km	Daily updated operational 1-hour forecasts

UC#2: Ocean monitoring

Enhance the quality of information and the optimization method for containerships voyages, so as to produce better analysis and plan more efficient routes in terms of fuel consumption, safety and arrival time precision.

The Use Case will demonstrate the capability of handling extreme volumes of data by fusing the meteorological data collected from EO data sources and a vessel to perform route optimization during the voyage of the ship.

EO4EU will offer enhanced visualization capacity of the data obtained both by sources as well as the in situ data, offering a multi-layer interaction with the user onboard, while augment user-friendliness and responsiveness.

UC#3: Food Security

- Impact analysis, mainly based on observation data (from ground, satellite, production and climatic time series)
- Risk of loss or damage estimation, through the development of predictive algorithms, based on recent time-series, forecast data and the characterization of the impact estimates of the previous component
- Identification of new areas with favorable climate conditions for specific crops
- Identification of crops suitable for new climate conditions

A specific pilot service concerning the agricultural product traceability will also be included (Sicily)

Input data	Spatial extent	Temporal extent
Sentinel 2 L1C – L2A	Southern, central northern Italian regions	From 2016 to present
Sentinel 1 GRD	Southern, central northern Italian regions	From 2016 to present
Temperature, rainfall, soil moisture from climate re-analysis	Southern, central northern Italian regions	From 2016 to present
Temperature, rainfall, soil moisture from climate seasonal forecast	Southern, central northern Italian regions	From present to end of project
Crop production from statistical offices	Southern, central northern Italian regions	From 2016 to present
Crop production from produces associations	Southern, central northern Italian regions	From 2016 to present

UC#4: Forest ecosystems

Apply a forest ecosystem model to simulate forest ecosystem services, e.g. carbon/water use efficiency, under ensemble of climate projections and potential management, so to support smart forest management

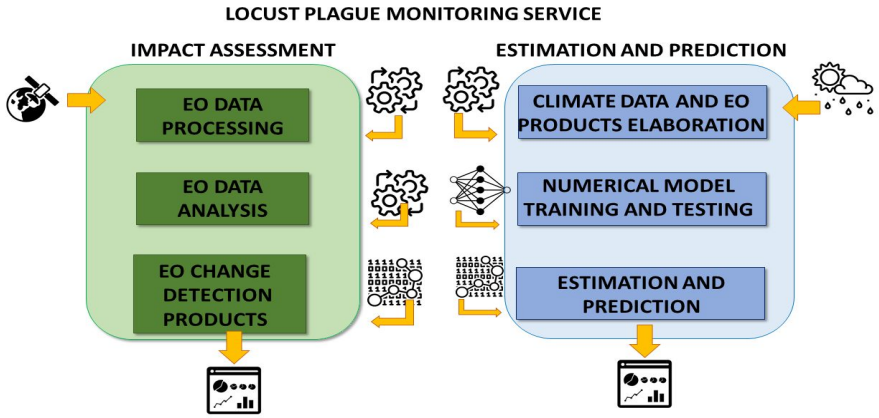
Input data	Spatial extent	Temporal extent
LAI or NDVI indices derived from Copernicus Land Monitoring Service (CLMS) products.	Part of Austria forest land, 300m grid spacing	2014 to 2020 (sensor PROBA-V) + 2020-present (sensor S3/OLCI), 10-days synthesis
Meteorological observations, point-based, interpolated or reanalysis (ERA5-Land or UERRA from C3S or downscaled CMCC runs) for rainfall temperature, vapour pressure deficit, radiation etc.	≈2-9 km grid spacing for interpolated data.	Aligning to satellite data (2014-present)
EURO-CORDEX for future projections	0.11° grid spacing	2021-2100
Soil properties data (SOILGRIDS; ESDB raster or vector)	250m – 1km grid spacing	Latest dataset/version available
Elevation Data (EU-DEM from Copernicus Land Monitoring Service)	25 m grid spacing	Latest dataset/version available
Forest management plans/inventory data	Point/areal survey/field campaign data	For the latest dates available

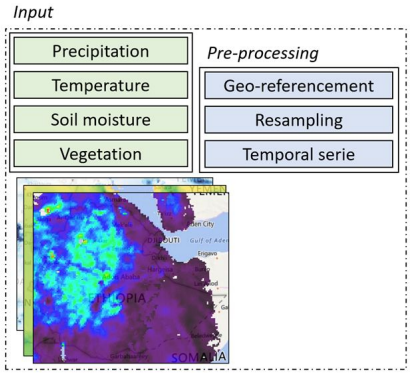
UC#5: Soil Erosion

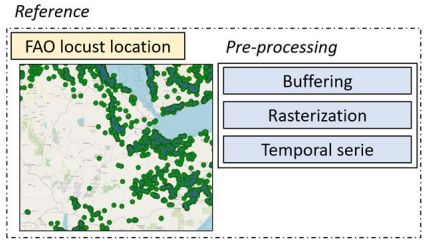
Combine EO and model data for robust estimation of rainfall erosivity and soil loss by water-induced erosion, by advancing empirical approaches with ML methods, under ensemble of climate projections, so to support smart land management

UC#6: Environmental Pests

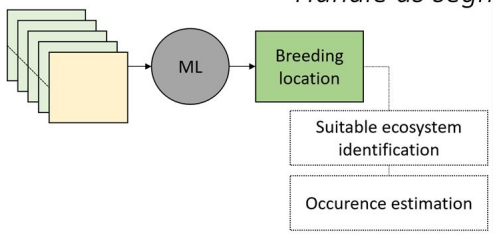
- Deliver an information service of locust plague impact assessment and prediction processing, combining together EO (esp. S1/S2) and climate data by means of AI/ML techniques

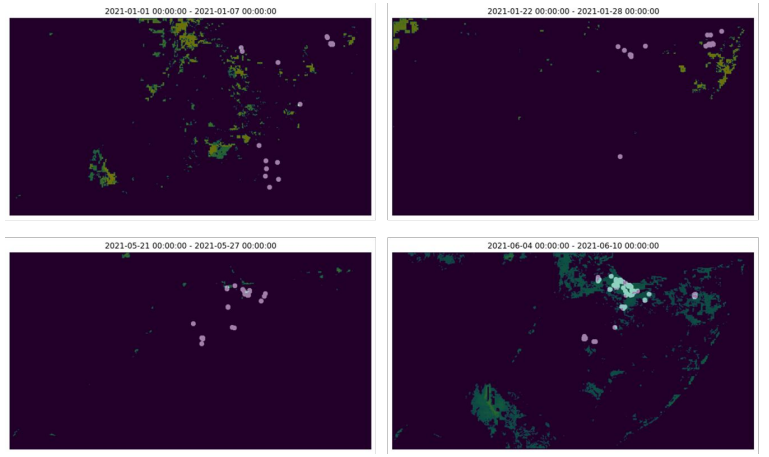






Handle as segmentation problem





LSTM results (pixel per pixel)

Bands class : Precision = 0.75

UC#7: Improving Civil Protection activities

Better exploitation and use of EU observed datasets to be used in all phases of Civil Protection operations, with focus in for wildfires and earthquakes

Input data	Spatial extent/ Resolution	Temporal extent/ Refresh
Copernicus Sentinel-2A/-2B L1C L2A	30,000 km^2, 12 tiles/10m-20m	M16-M35/average 10-day repeat
Copernicus CSCDA VHR1 VHR2	1,600 km^2/1¸4m	upon alert alert/only short term (1-3 days)
Planet Dove (Flocks) 2-4	10,000 km^2/3.7m	upon earthquake alert /vs. monthly mosaic
Eumetsat METOP LSA-SAF ENDV110v2	same as above/1.0 km	M16-M35/10-Day integration
Eumetsat MSG LSA-SAF FRM	same as above/3.3 km	M16-M35/ Daily + 24/48/72h
Sicily CTR (Regional Technical Chart)	same as above/10m	Static, latest update (curr. 2018, exp. 2023)
Sicily Fire Danger (Regional Forest Corps)	same as above/1.0km	Static, latest update (curr. 2020, yearly update on ex-post basis)