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)
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
EO4EU architectural bird's-eye view

- **EO4EU applications**
  - Dashboards
  - API
  - Data processing workflows
  - XR

- **EO4EU processing infrastructure**
  - Fusion
  - Semantic annotation
  - Compression
  - Knowledge management

- **EO sources**

**Orchestration Language**
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
EUROGEO WORKSHOP 2023

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).

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

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)

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

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

**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.
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

<table>
<thead>
<tr>
<th>Input data</th>
<th>Spatial extent/ Resolution</th>
<th>Temporal extent/ Refresh</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Copernicus Sentinel-2A/-2B L1C L2A</strong></td>
<td>30,000 km², 12 tiles/10m-20m</td>
<td>M16-M35/average 10-day repeat</td>
</tr>
<tr>
<td><strong>Copernicus CSCDA VHR1 VHR2</strong></td>
<td>1,600 km²/1.4m</td>
<td>upon alert/only short term (1-3 days)</td>
</tr>
<tr>
<td><strong>Planet Dove (Flocks) 2-4</strong></td>
<td>10,000 km²/3.7m</td>
<td>upon earthquake alert /vs. monthly mosaic</td>
</tr>
<tr>
<td><strong>Eumetsat METOP LSA-SAF ENDV110v2</strong></td>
<td>same as above/1.0 km</td>
<td>M16-M35/10-Day integration</td>
</tr>
<tr>
<td><strong>Eumetsat MSG LSA-SAF FRM</strong></td>
<td>same as above/3.3 km</td>
<td>M16-M35/ Daily + 24/48/72h</td>
</tr>
<tr>
<td><strong>Sicily CTR (Regional Technical Chart)</strong></td>
<td>same as above/10m</td>
<td>Static, latest update (curr. 2018, exp. 2023)</td>
</tr>
<tr>
<td><strong>Sicily Fire Danger (Regional Forest Corps)</strong></td>
<td>same as above/1.0km</td>
<td>Static, latest update (curr. 2020, yearly update on ex-post basis)</td>
</tr>
</tbody>
</table>