

European e-Infrastructure for Extreme Data Analytics in Sustainable Development



EUXDAT

The e-Infrastructure for large data analytics in agriculture

F. Javier Nieto – ATOS Spain



Athens, 10th July 2019

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777549





Outline

- **1. EUXDAT Context & Objectives**
- 2. Global Picture
- 3. What Can EUXDAT Offer?
- 4. Overview of the Achievements
- 5. Challenges to be Addressed

EUXDAT Context (I)









Land Monitoring and Sustainable Mgt















EUXDAT Context (II)



Open Land Use Map Improvement Monitoring of Crop Status **Delimiting Agro-Climatic Zones** Looking for Climatic Patterns Changes Information Support for Field Use Recommendations Effective Utilization of Natural Resources



- "…EUXDAT will build up a Large Data Analytics-as-a-Service e-Infrastructure with several software layers supporting sustainable and productive agriculture…"
 - Manage data storage and movement + Support heterogeneous data sources + configurable policies
 - Adapt data processing tools for HPC + Users' Portal with advanced features + Hybrid HPC&Cloud resources management
 - Provide access to EUXDAT services + Pilots implementation
 - Facilitate long-term sustainability + Collaboration (i.e. PRACE, EOSC)

Example: Delimiting agro-climatic zones CEUXDAT





http://www.euxdat.eu/

Elevation as a factor influencing temperature (DEM)

- Slope orientation as a factor influencing temperature
- Hydrology as a factor influencing temperature (buffering) •







Global Picture (I)



Marketplace & Data Catalogue	Land Use	Crop Monitoring	Climatic zones		User Frontend layer User interfaces Data Visualization
Data catalo API	^{gue} OLU	Crop Mor	o Status hitoring	Climatic zones	dded value data API layer Standardized API (OGC, INSPIRE, IoT)
Jupyter Cloudify	Data Catalogue	Open Land Use /lap improvement OLU database	Infection index computation Stored index maps	Climatic map generation Climatic maps	Application layer Processing applications Data storage
UAV	EU-DEM	Sentinel	Meteorologica	·	Data sources layer
http://www.eu	(dat.eu/				7

Global Picture (II)





Global Picture (III) - Interfacing with DIAS



* Image from Budapest ICT Proposers Day (11/2017) – Space WP

What Can EUXDAT Offer?



- Management and analysis of large data
 - Create your own application
 - Use specific libraries for data analysis in Agriculture
 - Connect to different data sources and move large datasets easily
 - Visualization features and custom frontends
- □ Easy usage of HPC + Cloud resources
 - Do not care about complex access mechanisms for storage and computation
 - Optimize HPC+Cloud combination for the user
- Default set of applications ready for the user
 - 6 re-usable scenarios
 - 3 re-usable pilots
 - Standardized APIs



Try to minimize large data movement

- Datasets sources evaluation model → Decide best source when available in more than one location
- Use caching of datasets \rightarrow Local repos for most used datasets
- Try to keep computation close to the data → EUXDAT Cloud close to Mundi
- Optimize local storage resources
 - Be smart with caching
 - Try to predict datasets needs → Move earlier
- Example: existing tools for managing data







Data Analysis and Visualization Tools





cnes

CENTRE NATIONAL D'ÉTUDES SPATIALES



Requirements, Features and Pilots/Scenarios

- Requirements collected from pilots and stakeholders
- Two versions of the requirements, features and architecture
- Definition of the pilots and their scenarios
- First scenarios implemented and others ongoing
 - Open Land Use Map
 - Crop Monitoring
 - Agro-Climatic zones

Overview of the Achievements (II)



Technical Components

- Cloud + HPC environments
 - PaaS layer to host User's platform
 - Elastic resources management
 - HPC+Cloud orchestration
- User's platform
 - Data Connectors
 - Data Management
 - Development environment
 - Execution Orchestrator
 - Monitoring
 - Data Analytics tools (GRASS, Orfeo, etc...)

Overview of the Achievements (III)



Data Sources	Data Available
Satellite Images	>8PB
Sensors Data	4TB
Hyperspectral Images	Ongoing
Meteo Data	337+1TB (9PBs)
Machine Monitoring	-
VGI Data	40GB
Other Vector Data	1GB (100GB)

Data Services	Available
Data Connectors	6
Data Analytics Tools	5



Data sources...

- Analysis of Hyperspectral images from UAV
- Data-in-motion analysis (compute at the edge) → Farm stations & machinery
- □ Plugins for Rucio (i.e. with the Data Catalogue, monitoring...)
- Parallelization of analysis for HPC
- Profiling and optimization of tasks execution + SLAs
- Marketplace setup and access control/billing mechanisms to implement a sustainable business model
- □ Involve more end users (hackatons, give access, etc...)



European e-Infrastructure for Extreme Data Analytics in Sustainable Development

Thank you for your attention

F. Javier Nieto

francisco.Nieto@atos.net



















This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777549





@euxdat

www.EUXDAT.eu