



Keywords:

#EOSC, #PlantPhenomics, #DataSecurity, #CollaborativeResearchs, #PhenotypicData, #EGIServices, #cloudcomputing, #EGIFoundation

Empowering Plant Phenomics: Leveraging EOSC Services for Enhanced Data Security and Collaboration

Harnessing EOSC for Improved Data Management and Collaboration in Multi-Scale Plant Phenotyping

The Project Involved



The European Infrastructure for Multi-scale Plant Phenomics and Simulation (EMPHASIS) is a distributed research infrastructure with the aim of developing and providing access to facilities and services addressing multi-scale plant phenotyping in different agro-climatic scenarios. EMPHASIS addresses the technological and organizational limits of European phenotyping, for a full exploitation of genetic and genomic resources available for crop improvement in changing climate.

The Challenge

However, among the services available through the infrastructure, the data services need to be improved. There is a need to build a federated and interoperable e-infrastructure that allows researchers to share and analyse phenotyping data.

In recent years, in fact, technological progress has been made in plant phenomics, and high-throughput plant phenotyping platforms now produce massive datasets involving millions of plant images concerning hundreds of different genotypes at different phenological stages in both field and controlled environments.

Networks of sensors also measure environmental conditions in real-time. The ongoing robotisation of experimental processes foreshadows an explosion in the volume and complexity of the data produced by the different research facilities.

Isabelle Alic



Agronomical Engineer at INRAE

"EOSC offers access to high-performance computing resources, ensuring that the PHIS can handle complex computations efficiently."

EOSC service or tool used

EMPHASIS used several services from the EOSC marketplace: A storage layer based on the open-source Phenotyping Hybrid Information System [PHIS](#) is proposed to store, organise raw data and capture data provenance.

The [EGI Cloud Compute](#) is used to host three PHIS instances. The [EGI Check-In](#) has been integrated as the default authentication system. The connection to PHIS instances is thus proposed with this authentication system in addition to the usual connection.

Also, the [EGI Notebook](#) service was used to build training material on how to use the tool developed during the EGI-ACE project.

Useful tips & tricks

"We used the INRAE ForgeMia CI/CD pipeline to deploy PHIS instances (based on [OpenSILEX](#)) on the EGI [OpenStack](#) cloud compute.

We extended the number of models adapted to plants with [Apple tree blossom image segmentation and apple detection](#).

Also, we developed the possibility to improve a model: A user can now manually improve the given mask annotation by inferring a model with the [napari-pixel-correction plugin](#) of the [napari software](#).





Keywords:

#EOSC, #PlantPhenomics, #DataSecurity, #CollaborativeResearchs, #PhenotypicData, #EGI Services, #cloudcomputing, #EGIFoundation

The Research Community

EMPHASIS plays a key role by integrating the European plant phenotyping infrastructure and ensuring effective utilisation of synergies. In fact, plant phenotyping offers innovative tools to help understand the interaction between genotypes and the environment to better select new plant varieties for breeding by establishing an integrated plant phenotyping infrastructure.

Benefits and impact

EGI services, specifically through `deepaas-api`, ensure the secure handling of data for processing or analysis, maintaining confidentiality. This feature becomes crucial when compared to Google Cloud, where data stored may be subject to Google's access rights. EMPHASIS significantly benefits from the enhanced data security provided by the EOSC environment, utilising EGI's reliable infrastructure for confidential data processing.

Thanks to EGI Services, the National Plant Phenotyping Infrastructure in Finland (NaPPI) has successfully deployed a dedicated phenotypic database (PHIS) on a sustainable and user-friendly platform, enabling them to take a significant step forward in generating FAIR data. In the near future, these resources will be valuable to share multiscale data with collaborators. This use case supported a [Data Management Workshop](#) within the NordPlant community, and to promote best practices and tools.

Why do I need EOSC?

EOSC offers access to high-performance computing resources, ensuring that the PHIS can handle complex computations efficiently. EOSC provides a secure environment for handling sensitive data. Deploying PHIS within EOSC ensures that phenotypic data is managed with a focus on security and privacy, following established standards and protocols.

EOSC facilitates collaboration among researchers and institutions. PHIS can benefit from the collaborative environment, enabling researchers to share and work on phenotypic data collectively.

Useful material related to this story



EMPHASIS

Across disciplines

EOSC promotes the use of standards for data and metadata. Deployed within the EOSC framework, PHIS can adhere to these standards, ensuring consistency and interoperability in the representation of phenotypic information.

Phenotyping often involves collaboration between different scientific disciplines. EOSC, being a multidisciplinary platform, provides an ideal environment for integrating phenotypic data with other types of data, fostering interdisciplinary research.

Limitations and future improvement

It would be nice to make a GPU available on `deepaas-api` and EGI Notebook to run deep learning models: a deep learning model implemented in `deepaas-api` will make their manipulation easier for neophytes.



Liked this
#EOSCinPractice
story?

Follow
@EOSCFuture
for more!