



# European efforts for phenomics data management

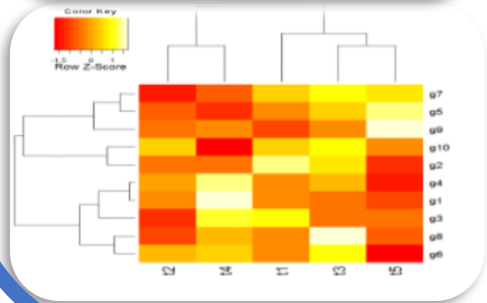
Sylvain Poque, 21.11.2024



# Levels of phenotyping => Trans scale analysis

## Controlled environments

Lab phenotyping



Phenotyping in controlled environment

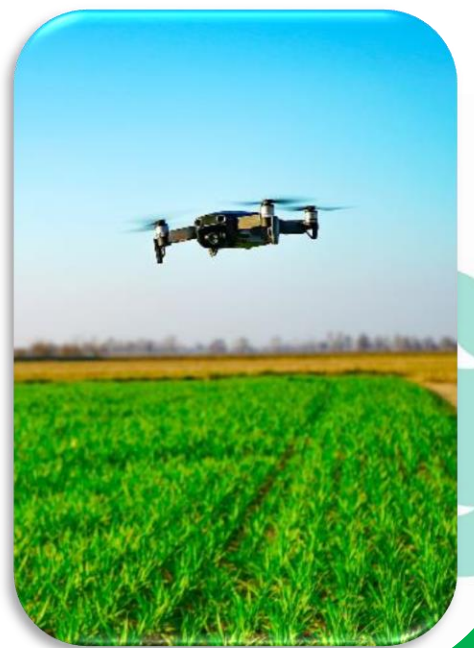


## Natural environments

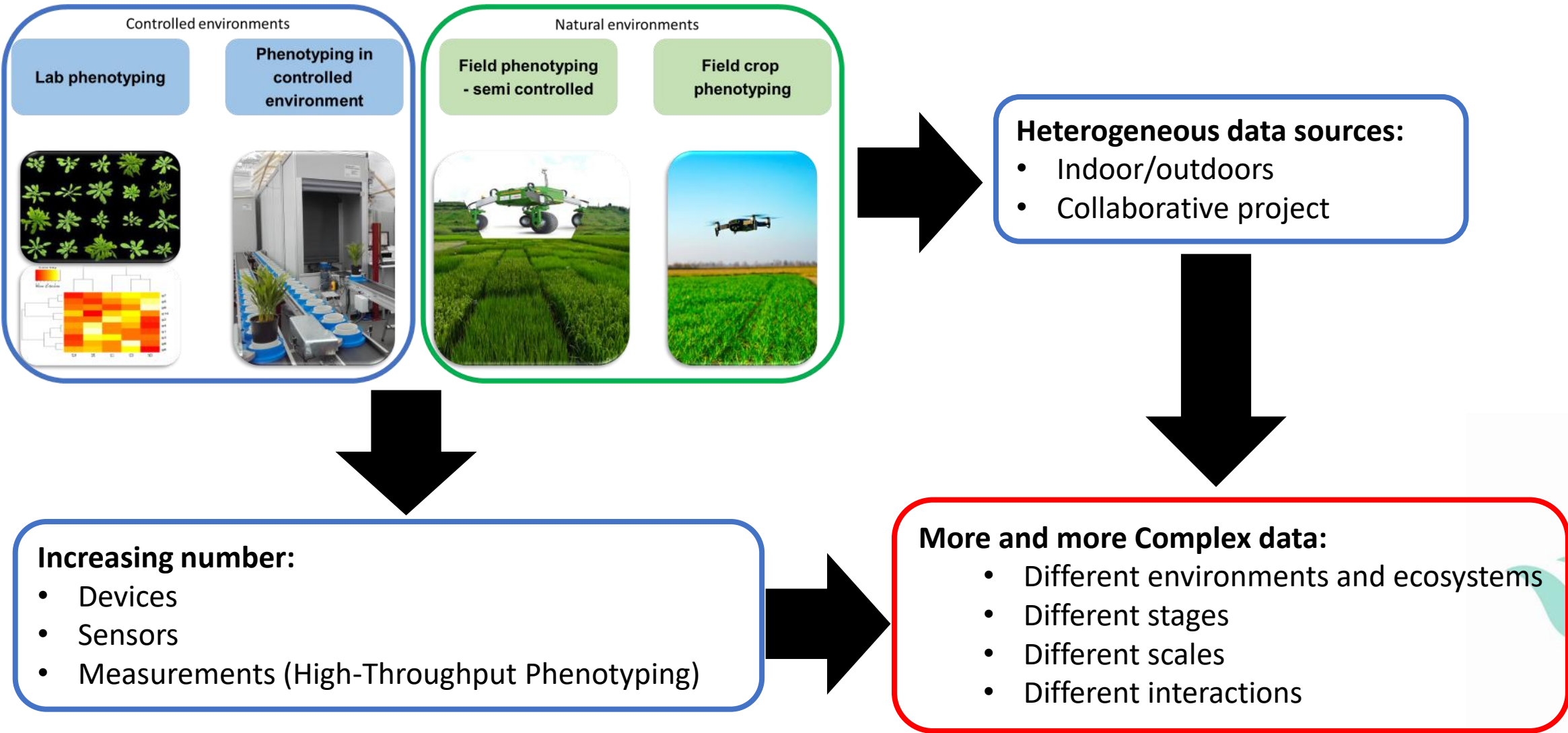
Field phenotyping - semi controlled



Field crop phenotyping



# Levels of phenotyping => Trans scale analysis



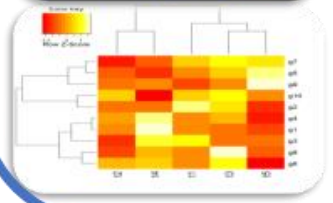


# Big data management difficulties

Controlled environments

Lab phenotyping

Phenotyping in controlled environment



Natural environments

Field phenotyping - semi controlled

Field crop phenotyping



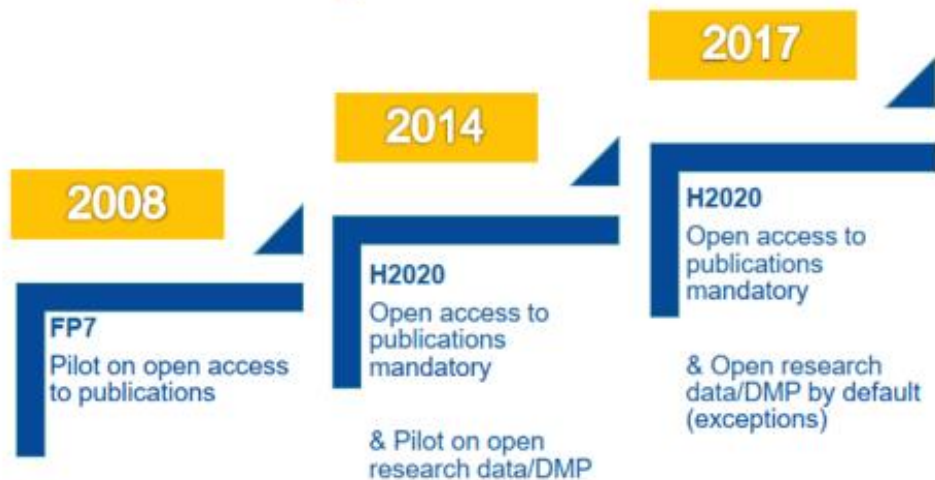
## Big Data



# EU requires for more and more Open-Science

## From open access to open science

Evolution of Open Science policies across Framework Programmes



### Under Horizon Europe (2021)

- Open Science **embedded** across Horizon Europe
- **Strengthening of the open access obligations** and focus on responsible research data management in line with the FAIR principles

Open Science @EU:  
an overview and current developments

Victoria Tsoukala, PhD

European Commission, Directorate-General for Research & Innovation, Unit 'Open Science'

EuroCRIS Spring Membership Meeting  
Brussels, 31 May 2023

### Open-Science

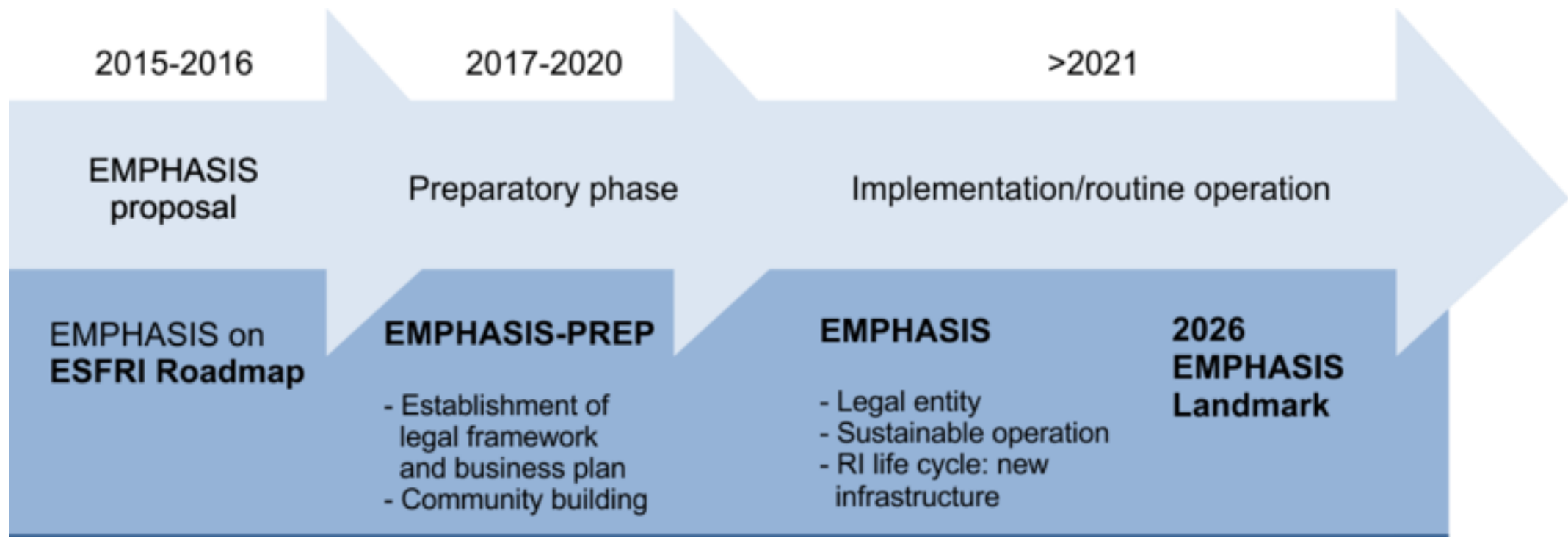
- **Requires Data accessibility:**
  - By facilitating open access of research data
- **To enhancing:**
  - Collaboration among researchers worldwide
  - Transparency and Reproducibility in research
  - **Accelerated Innovation and scientific discoveries**
  - **Maximize value of past, ongoing and future data**



# EMPHASIS

## EMPHASIS: European Infrastructure for Multi-scale Plant Phenomics and Simulation

- **Listed on:** ESFRI Roadmap 2016 for Research Infrastructures
- **Objective:**
  - Establish a Europe-wide plant phenotyping infrastructure. **[Trans-Scale Analysis]**
  - Connect data collection with European information systems and modeling. **[FAIR Transition]**
  - Develop and share advanced knowledge and technologies. **[Open-Science]**



# Who Are Their Targets?

## **Researchers**

Seeking quantitative plant assessment tools.

## **Public Sector Investors**

Complementing plant phenotyping infrastructure across Europe.

## **Scientific Institutions**

Utilizing synergies in operating plant phenotyping infrastructure within Europe.

## **Industry**

Driving innovation in technology development and its application in breeding.

## **Society in General**

Supporting sustainable increases in food quality and quantity under changing climate conditions.

# Infrastructure Categories

Plant Phenotyping requires integration of both **facilities** and **activities**

## Controlled Conditions



## Intensive Field



## Lean Field



## Modelling



## Data & Computational Services







### I) Controlled Conditions

- Greenhouses and growth chambers
- Monitoring of (semi-)controlled environmental conditions
- Often automated, e.g., plant to sensor or sensor to plant
- Throughput typically between 100-1000s plants



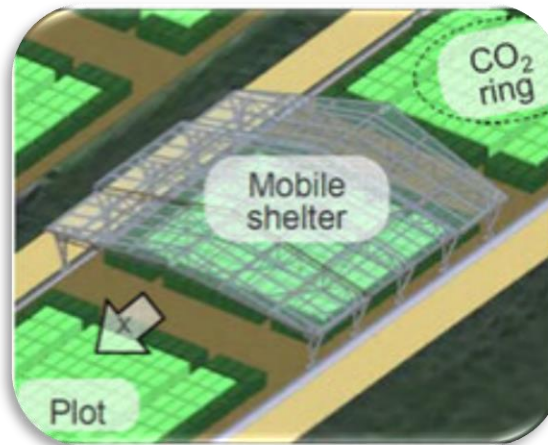
### 124 Affiliated Facilities





## II) Intensive Field

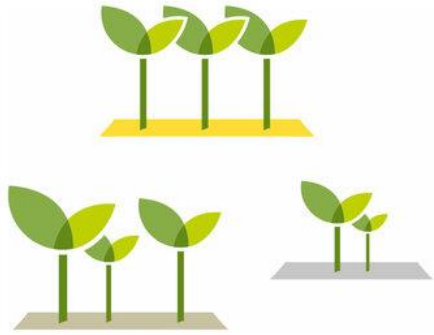
- Fully equipped field trials
- Detailed environmental monitoring
- High-quality/throughput phenotyping measurements
- Semi-controlled intensive field sites: alter environmental conditions



67 Field Installations







### III) Lean Field

- Field trials with environmental monitoring
- Phenotyping equipment for yield and moisture content
- Potentially ground based or airborne sensing systems
- Usually in networks of fields



# Infrastructure Categories

Plant Phenotyping requires integration of both **facilities** and **activities**

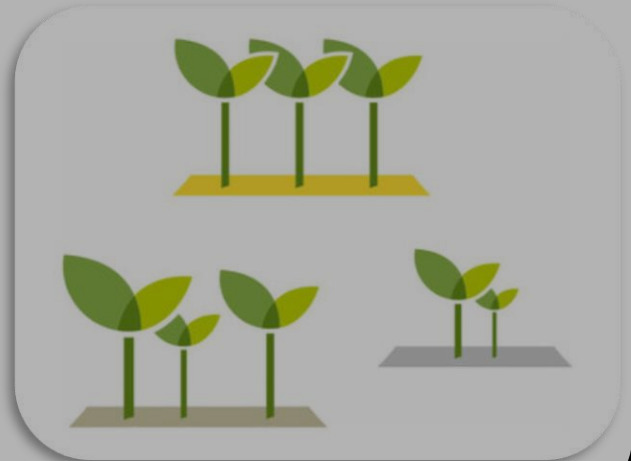
Controlled Conditions



Intensive Field



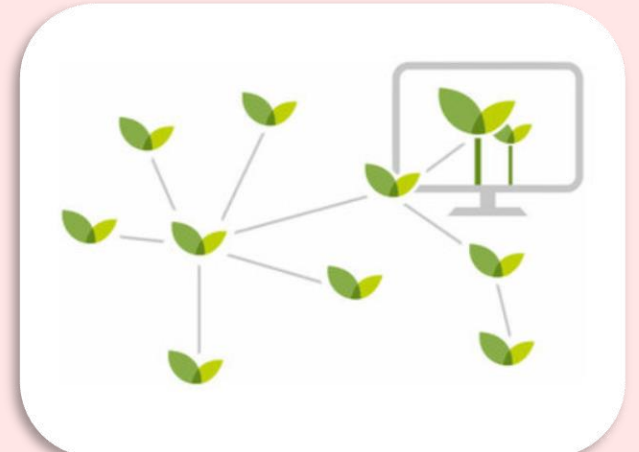
Lean Field



Modelling



Data & Computational Services

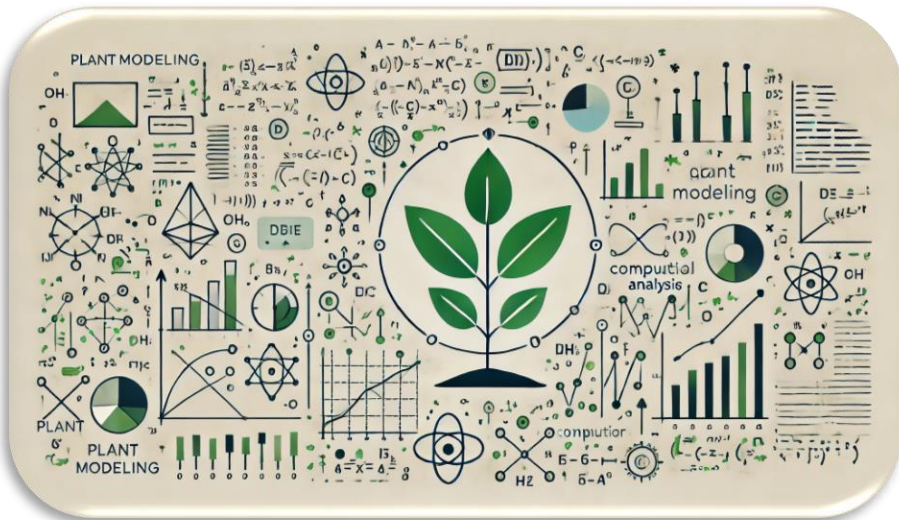




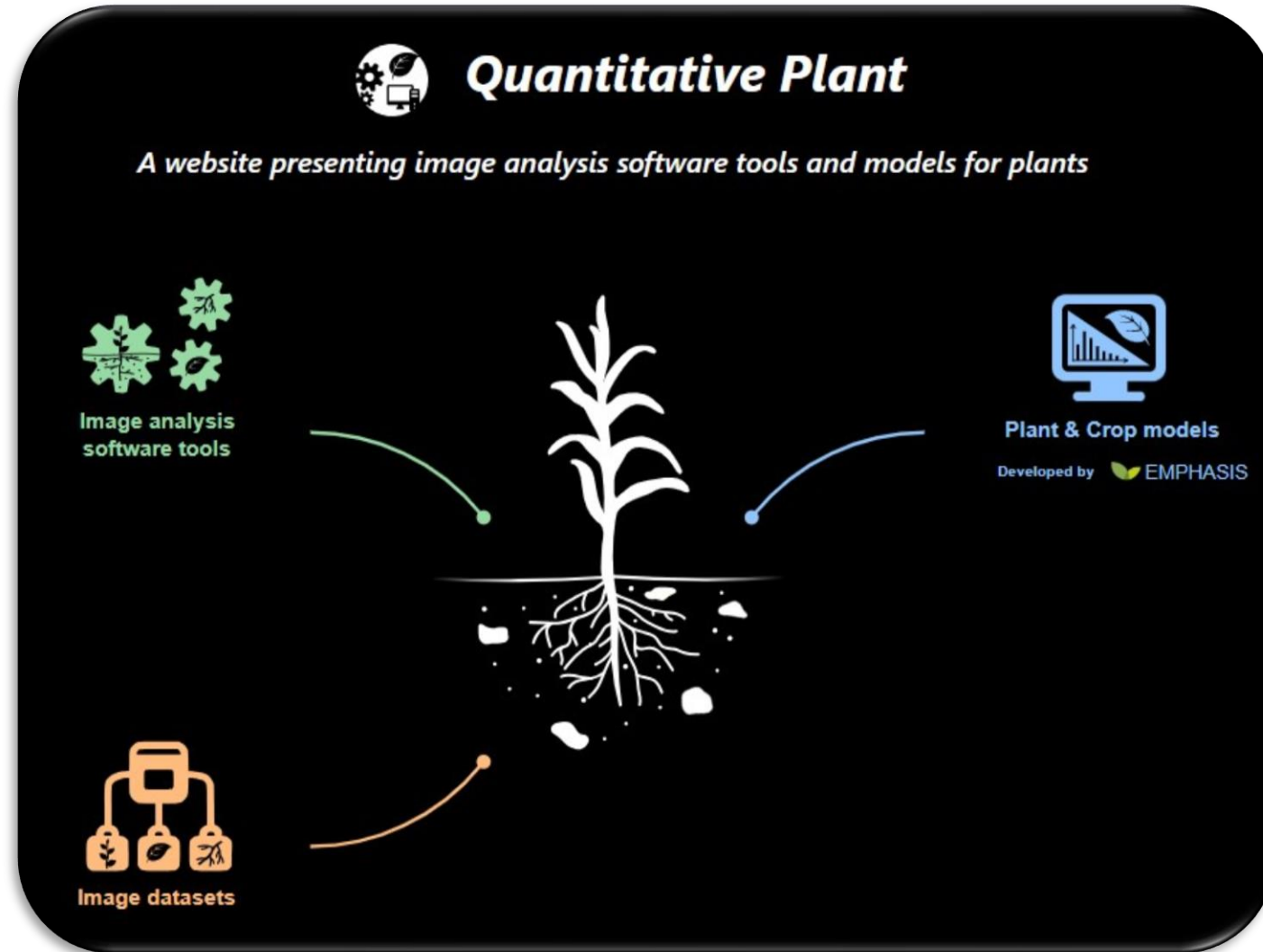


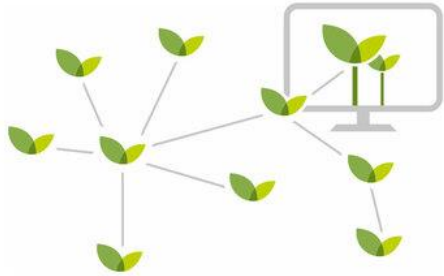
#### IV) Modelling

- Virtual platforms
- Integrated or interfacing with phenotyping data
- Different types of models, e.g., Crop Models, FSPM



Modelling Pilot:  
<https://www.quantitative-plant.org/>





### V) Data & Computational Services

- FAIR Information systems plant phenotyping data
- Access to data
- Support for local installation data management
- Integration of information system into pan-European system



**F**indable  
**A**ccessible  
**I**nteroperable  
**R**e-usable



Essential for meta-analysis and AI tool development, enabling data integration, pattern discovery, and innovation.

# PHIS: an Ontology driven Information System for Plant Phenomics



PHIS is an applications of OpenSILEX, developed to

- Store
  - Organize
  - Manage
- Multi-scale and multi-source data

Allowing

- Day-to-day monitoring of experiments
- Data sharing with collaborators and scientific community

## Identification

- Everything can be identified: plants, experiments, sensors, events, etc.
- Persistent, unambiguous, resolvable

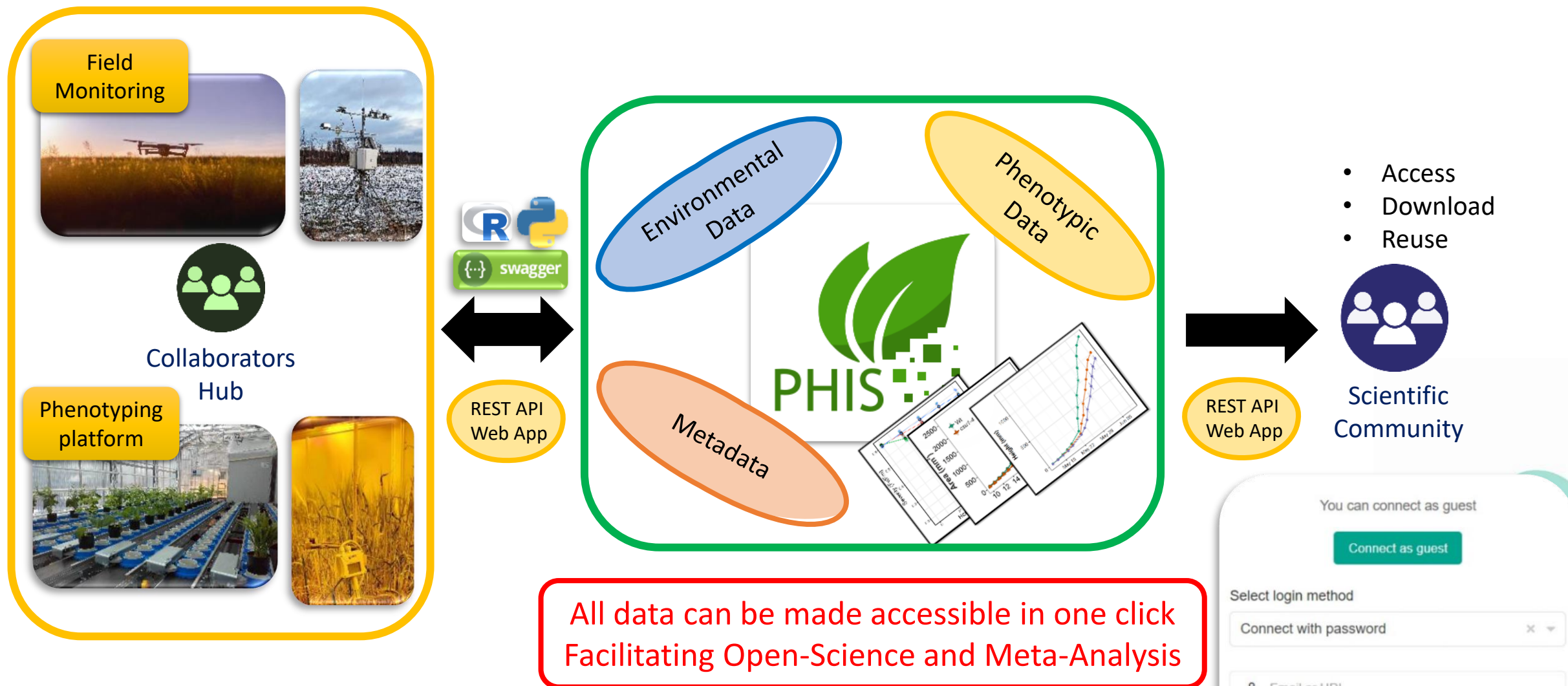
## Semantics

- Naming Conventions
- Controlled vocabulary
- Formalized relationships between entities
- Data annotation and enrichment

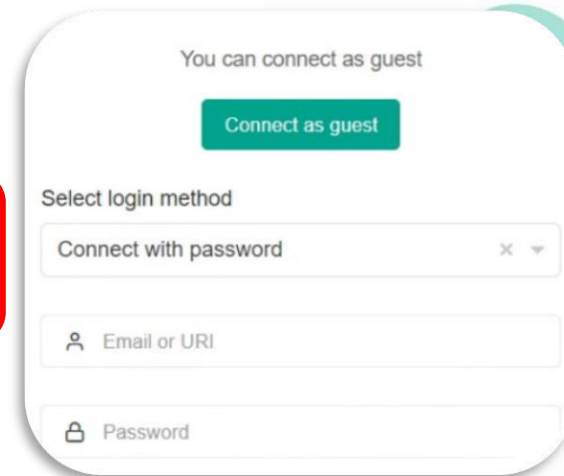
## Highly structured data

- Making data **“FAIR”**
- Enhancing flexibility for research
- Improving understanding and reproducibility of data
- Easing the transition to **Data Management Plans and Open Science**

# PHIS, making data FAIR & open to the scientific community



All data can be made accessible in one click  
Facilitating Open-Science and Meta-Analysis







## In collaboration with OpenSilex (INRAE), NaPPI team has been actively involved in:

- **Initiate the implementation of PHIS for Data Management in Nordic Countries:**
  - Creation of a UH PHIS instance on EGI Grid to manage experimental data from collaborative projects.
- **Developing Scripts for Data Import:**
  - Creating and customizing scripts to streamline NaPPI data import process into PHIS.
  - [<https://forgemia.inra.fr/opensilex-scripts>].
- **Organizing Training Sessions:**
  - Workshop in 2023 within the EGI-ACE project, focusing on FAIR data management and PHIS usage.
    - Video recording of the workshop can be found on NaPPI website [[EGI-ACE workshop 2023](#)]
  - Conducting a series of "Collaborators Training" session, providing hands-on learning opportunities on PHIS.
    - Generation of Step-By-Step video material

PHIS

## Covered Subject:

- **FAIR Data & Metadata**
- **Data Standards**
  - For Biologists
  - For Computer Scientists
- **DMP (Data Management Plans)**
  - Pragmatic Approaches
  - Useful Applications for RDM

## Providing Hands-On:

- **MIAPPE Template**
  - Variable Building
- **PHIS Usage**
  - Introduction and manage your data
  - Script-based Data Importation
- **Data Submission Platforms**
  - Dataverse
  - Zenodo
- **Sequence Data Management**
  - EBI (European Bioinformatics Institute)
- **ISA-Based Solutions**
- **AI Tools**
  - Image Information Extraction
- **Cloud Infrastructure**
  - Deploying Analysis Workflows (EGI)

## Teachers:

- Salvatore Cuomo
- Sebastian Beier
- Yin Chen
- Sylvain Poque
- Sven Warris
- Peter Selby



# NBPPN, Hybrid Workshop on data standards in Plant phenotyping

On-site training event in Helsinki - 26, 27 & 28 February 2025 (Jesper Cairo Westergaard, Sylvain Poque, Isabelle Alic)

## Covered Subjects (Online)

### •Concept of Multi-Domain Data

- Integrating diverse datasets from various domains to enhance research outcomes

### •Importance of Meta-Analysis

- Leveraging combined data for broader insights and robust conclusions

### •Benefit of BrAPI

- Standardized Breeding API for seamless data exchange

### •Data Management

- **Benefits:** Enhanced efficiency, data reusability, and compliance with standards
- **Common Practices:** FAIR principles, proper documentation, and consistent workflows
- **Potential Tools:** Focus on PHIS

## On site Hands-On to get to know PHIS:

### •How PHIS can benefit your Meta-Analysis

- Collecting data from existing trial and visualize it

### •How to Import Data to existing trial

- web browser interface, Python scripts

### •Design and reference your own Trial

- Ensuring proper metadata inclusion for comprehensive data

## Teachers:

- Francois Tardieu
- Cyril Pommier
- Peter Selby

Target audience both **Breeders** and **Researchers**.



UNIVERSITY OF HELSINKI





# Thanks for your attention



UNIVERSITY OF HELSINKI

PhD **Sylvain Poque**, NaPPI, [sylvain.poque@helsinki.fi](mailto:sylvain.poque@helsinki.fi)  
Project planner at the National Plant Phenotyping Infrastructure  
Department of Agricultural Sciences, University of Helsinki, Finland



PhD **Kristiina Himanen**, Docent in developmental plant biology  
Research coordinator at the National Plant Phenotyping Infrastructure,  
Organismal and Evolutionary Biology Research Program and Dept. of Agricultural  
Sciences, University of Helsinki, Finland  
[kristiina.himanen@helsinki.fi](mailto:kristiina.himanen@helsinki.fi)

