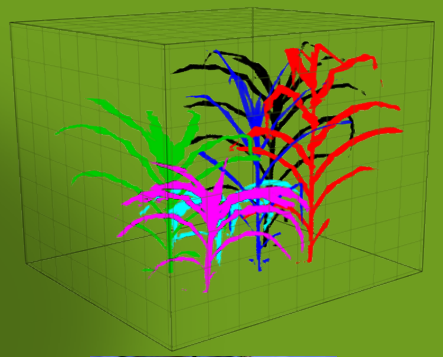
Dealing with multi-source and multi-scale information in Plant Phenomics: the PHIS ontology-driven Information System



Vincent Nègre

Llorenç Cabrera-Bosquet

INRA-LEPSE Montpellier, France

✉: *llorenc.cabrera-bosquet@inra.fr*

Data Management Challenge

Plant phenomics is multi-source and multi-scale, with and added complexity of time

« omics » Platforms

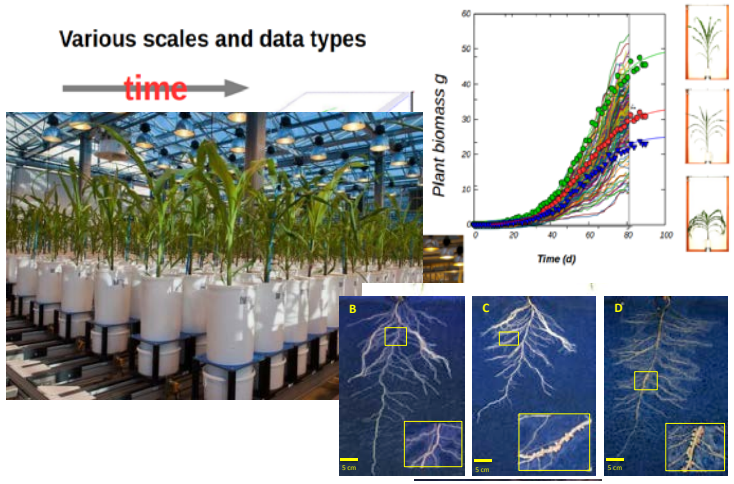
Various data complex types

Genomics

Composition and the structure of biopolymers

Quantification of metabolites and enzyme activities

Green house Platforms



Managing all these data is a real challenge an probably one the most limiting steps

Field Platforms

Various scales and data types

- Cell, organ, plant, population
- Images, hyperspectral, spectral, sensors, human readings...

Farm Platforms

Various scales and data types from thousands of farms

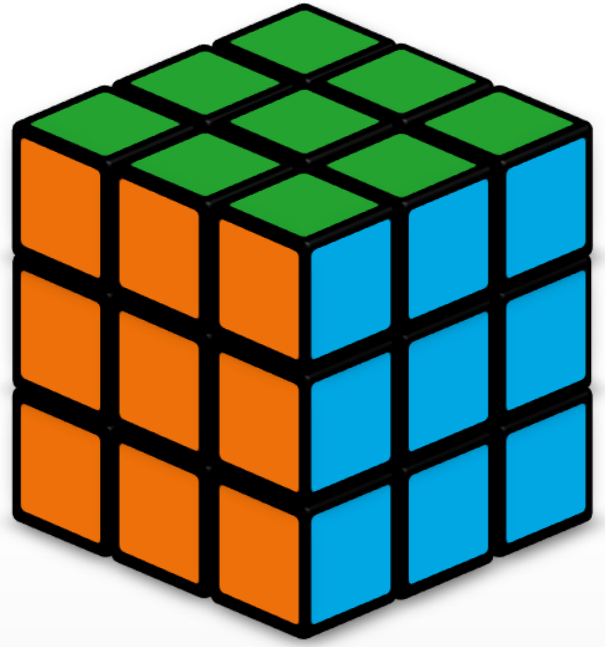
- organ, plant, population, site
- Images, sensors, human readings...

PHIS is an Information System for plant phenomics able to... (Phenotyping Hybric Information System)

Data management

Store, organise and manage

- Highly heterogeneous data (e.g. images, spectra, dynamic traits, environmental data)
- Multi-spatial and temporal scale data (leaf to canopy level)
- Multi-source (field, platform)



www.phis.inra.fr

New Phytologist — Research

Methods

Dealing with multi-source and multi-scale information in plant phenomics: the ontology-driven Phenotyping Hybrid Information System

Pascal Neveu[1], Anne Tireau[1], Nadine Hilgert[1], Vincent Nègre[2], Jonathan Mineau-Cesari[1,2], Nicolas Brichet[2], Romain Chapuis[3], Isabelle Sanchez[1], Cyril Pommier[4], Brigitte Charnomordic[1], François Tardieu[2] and Llorenç Cabrera-Bosquet[2]

[1]MISTEA, INRA, Montpellier SupAgro, Université de Montpellier, Montpellier 34060, France; [2]LEPSE, INRA, Montpellier SupAgro, Université de Montpellier, Montpellier 34060, France; [3]UE DIASCOPE, INRA, Montpellier SupAgro, Université de Montpellier, Montpellier 34060, France; [4]INRA, UR1164 URGI – Research Unit in Genomics-Info, INRA de Versailles-Grignon, Route de Saint-Cyr, Versailles 78026, France

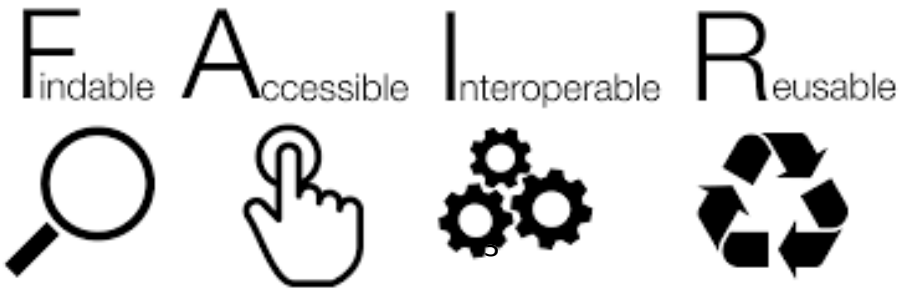
Neveu et al. *New Phytologist* 2019

Interoperability

Interoperate and integrate data into/from external resources (e.g. modelling platforms or external databases)

Metadata enrichment

Not only storing data but… **Enrich datasets** with the necessary **knowledge** (enable reuse of data and meta-analyses)



=> Provide FAIR data (Wilkinson *et al*. 2016)

Object identification (URIs)

In PHIS all objects are identified using URIs (Uniform Resource Identifiers)

=> standardized, unique, unambiguous identification

Prefix m3p: *<http://phenome-fppn.fr/m3p>* (a)

URI of plant
<m3p:arch/2017/c17000118>

URI of pot:
<m3p:arch/2013/pc13001542>

URI of cart:
<m3p:arch/2013/ct1300123>

URI of cabin:
<m3p:arch/2018/ac180015>

URI of camera:
<m3p:arch/2018/ac180019>

URI of image:
<m3p:arch/2017/ic17002295855>

Prefix diaphen: *<http://phenome-fppn.fr/diaphen>* (b)

URI of plot
<diaphen:2017/o1700029>

URI of plant:
<diaphen:2017/17000147>

URI of leaf:
<diaphen:2017/l17000590>

URI of camera:
<diaphen:2018/ac180002>

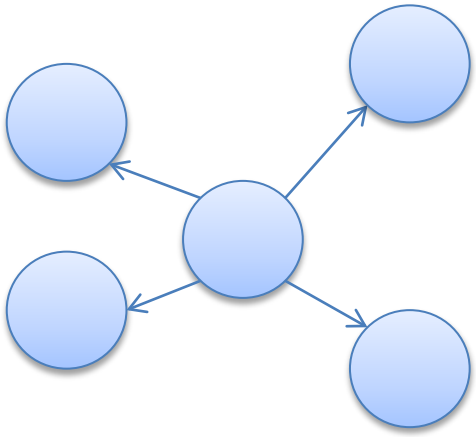
URI of image:
<diaphen:2017/ic14001480237>

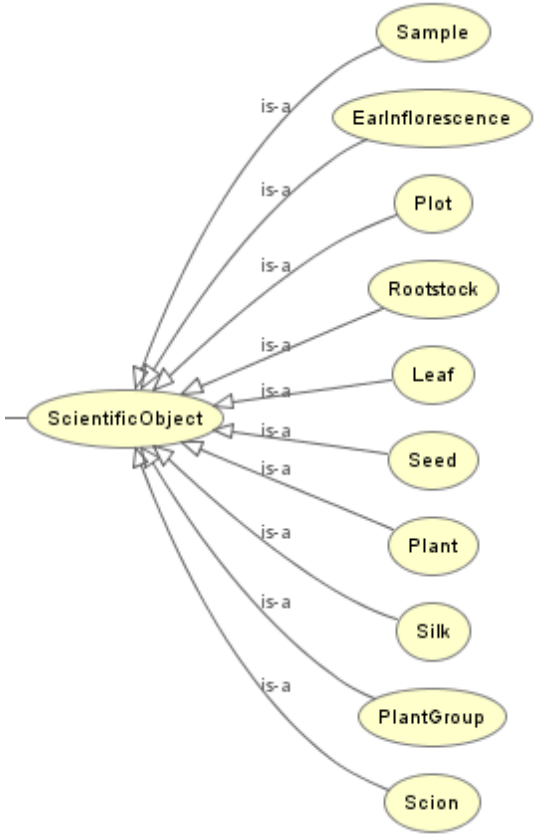
Neveu et al. *New Phyt* 2018

The same applies to sensors, people, events, infrastructure, **variables**…

Ontologies

- Une **ontologie** est l'ensemble structuré des termes et concepts dans un domaine de connaissance.
- L'ontologie constitue en soi un modèle de données représentatif d'un ensemble de concepts, ainsi que des relations entre ces concepts.
- L'ontologie est aux données ce que la grammaire est au langage.





OEPO (Ontology for Experimental Phenotypic Objects)

How information is organised? Object identification and vocabulary (URIs and ontologies)

Object identification

In PHIS all objects are identified using URIs (Uniform Resource Identifiers)
=> standardized and unambiguous identification

Prefix diaphen: *<http://phenome-fppn.fr/diaphen>*

(b)

URI of plot
<diaphen:2017/o1700029>

URI of plant:
<diaphen:2017/17000147>

URI of leaf:
<diaphen:2017/l17000590>

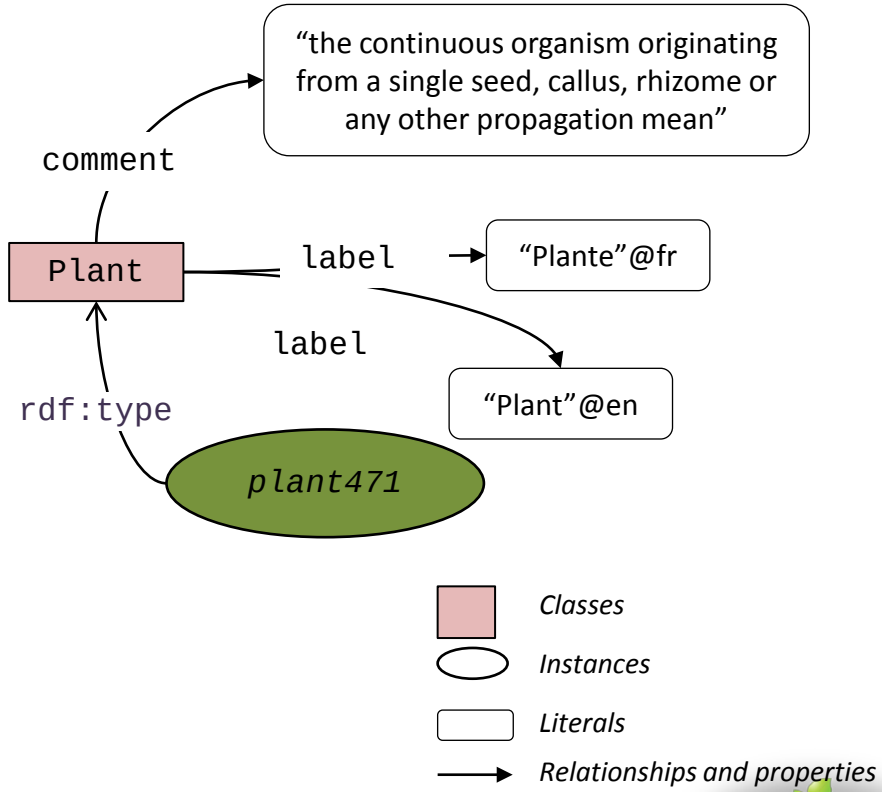
URI of camera:
<diaphen:2018/ac180002>

URI of image:
<diaphen:2017/ic14001480237>

Neveu et al. *New Phyt* 2019

Semantics (controlled vocabulary)

Ontologies allow to define terms and formalise relationships between them



How information is organised? Object identification and vocabulary (URIs and ontologies)

Object identification

In PHIS all objects are identified using URIs (Uniform Resource Identifiers)
=> standardized and unambiguous identification

Prefix diaphen: *<http://phenome-fppn.fr/diaphen>*

(b)

URI of plot
<diaphen:2017/o1700029>

URI of plant:
<diaphen:2017/17000147>

URI of leaf:
<diaphen:2017/l17000590>

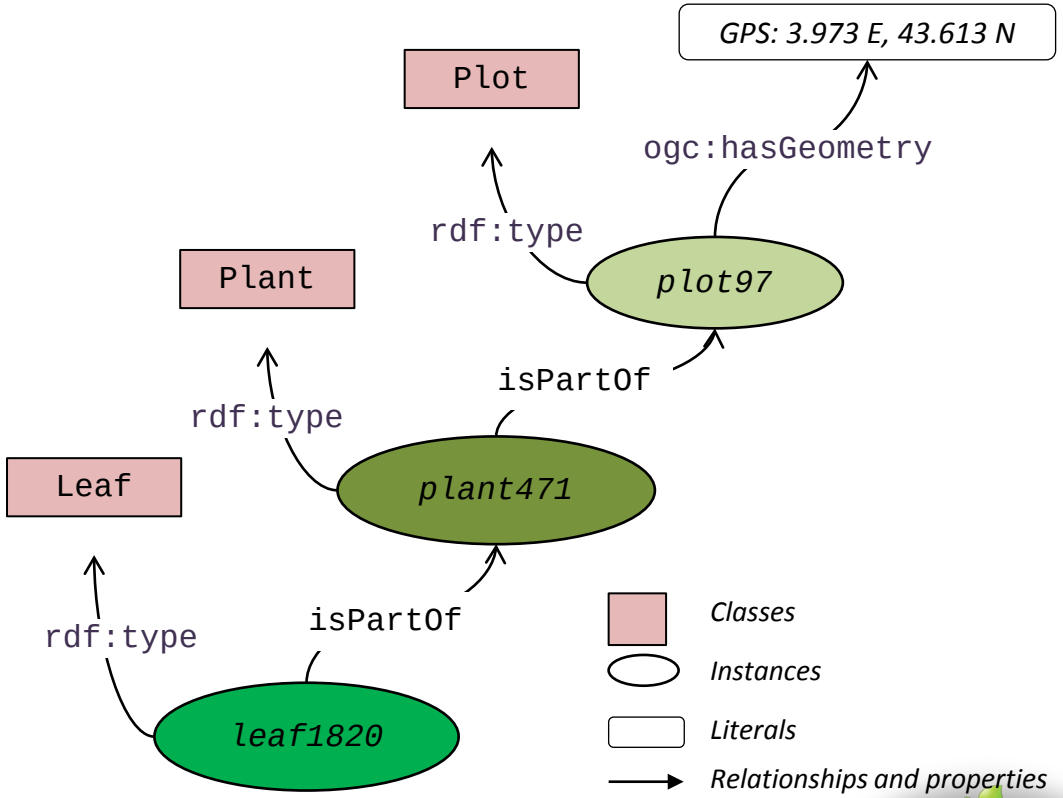
URI of camera:
<diaphen:2018/ac180002>

URI of image:
<diaphen:2017/ic14001480237>

Neveu et al. *New Phyt* 2019

Semantics (controlled vocabulary)

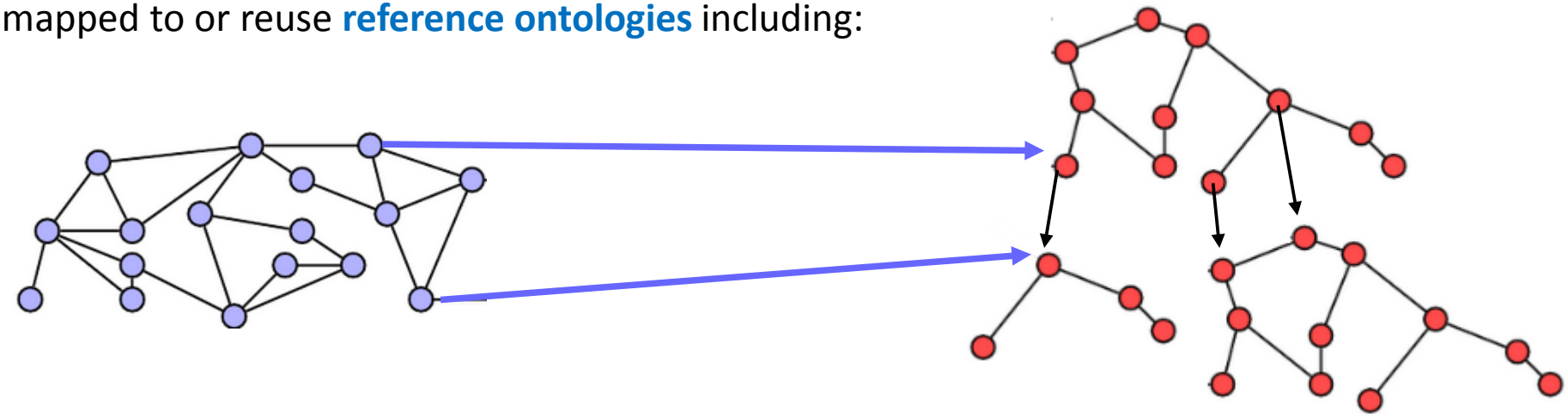
Ontologies allow to define terms and formalise relationships between them



The same applies to sensors, people, events, infrastructure, variables…

Application vs reference ontologies

OEPO and OEEv **application ontologies** can be specific to either field or controlled conditions, but a number of terms and definitions are mapped to or reuse **reference ontologies** including:



Application ontologies
(cross domains of knowledge)

- **OEPO:** Ontology specialized for describing objects participating in phenotyping experiments
- **OEEv:** Ontology aimed to represent events that occur during an experiment

Reference ontologies (domain specific)

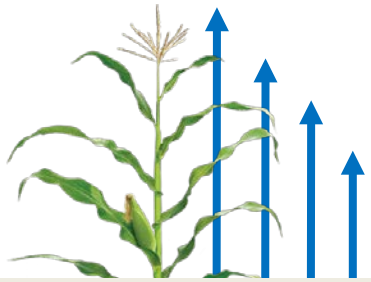
- Crop ontology (CO)
- Plant Ontology (PO)
- FAO/Bioversity Multi Crop Passport Descriptors
- Semantic Sensor Network Ontology (SSN)
- PATO
- PPEO (MIAPPE)
- other semantic resources such as the AGROVOC, ICASA, EMPHASIS

Variable identification: Plant height example

Crop Ontology
for agricultural data

In PHIS each measured variable is described using the ***Trait + Method + Unit*** structure

Trait + Method + Unit



M1: Total height

M2: First tassel branch

M3: Last expanded leaf

M4: Youngest growing leaf

…There is an uncountable number of combinations…
Each trait, method and unit has to be identified if we want to share and reuse data

T1: Plant Height

M5: Highest pixel corresponding to plant

U3: pixel

PHIS

A qui et à quoi sert PHIS?

1. Platform users and staff
    - Management of phenotyping experiments / infrastructure
    - Data exploration and pre-analysis

2. External users (Interoperability with external platforms)
    - Re-analysis and meta-analyses
    - Modelling

3. Implemenation into other platforms

Experiments: contextual information

ARCH2017-03-30

Add annotation | Add event | QrCode | Graphic Visualization | Return to the list

Alias	ZA17
URI	http://www.phenome-fppn.fr/m3p/ARCH2017-03-30
Species	Maize
Project	PHENOME-FPPN, Progress_Genetique, XYZ
Start	30 Mar 2017
End	30 Jun 2017
Installation(s)	http://www.phenome-fppn.fr/m3p/phenoarch, http://www.phenome-fppn.fr/m3p/phenodyn
Scientific Supervisor	Cabrera@supagro.inra.fr, Francois.tardieu@supagro.inra.fr, Olivier.turc@supagro.inra.fr, Welcker@supagro.inra.fr
Technical Supervisor	Nicolas.brichet@supagro.inra.fr, Nathalie.luchaire@inra.fr, Benoit.suard@supagro.inra.fr
Comment	Cette manip vise à analyser les traits associés avec le progrès de rendement dans une série historique de maïs composée d'hybrides représentatifs des différents zones et leur réponse au déficit hydrique, en combinant des approches innovateurs basées sur le phénotypage à haut-débit en plateforme en complément avec des mesures au champ. Parmi les traits explorés: la variabilité de la croissance et la transpiration et leur réponse au déficit hydrique / haut température ainsi qu'un suivi sur le développement reproducteur via analyse d'image (caméra XYZ) et mesures fines. Le panel est composé de 60 hybrides français. Dans cette manip on a fait une pré-calibration de la manip sorgho qui aura lieu en septembre 2017.
Keyword	Genetic gain, maize, grain abortion, silk growth, greenhouse, phenoarch
Document	Plme_d_17_00129.pdf

#Object	349 Ear inflorescences, 359 Leafs, 1496 Plants
#Variety	62

Projects

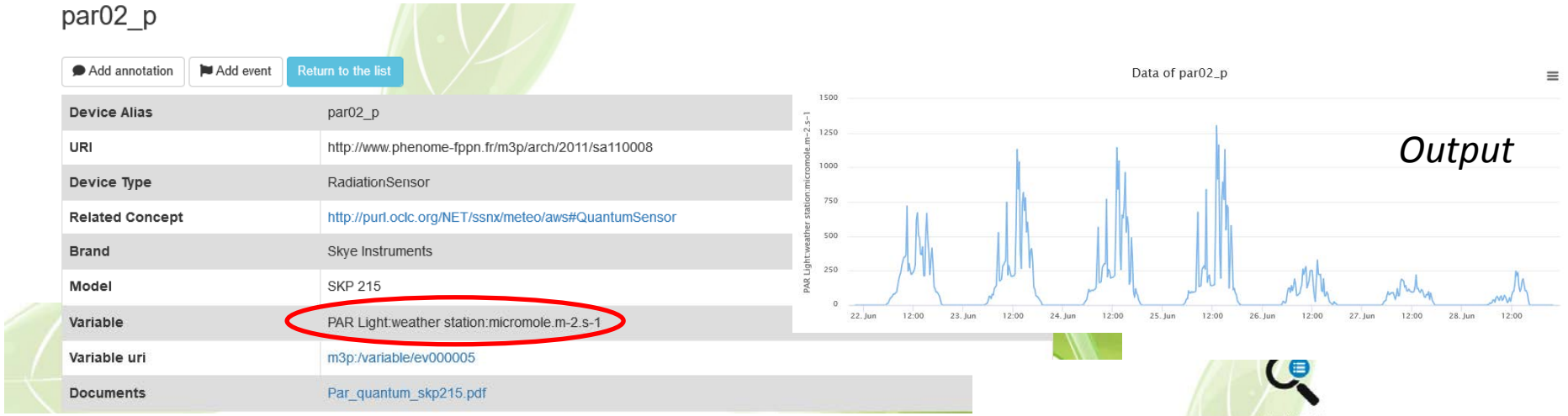
Installations

Staff & teams

Objects

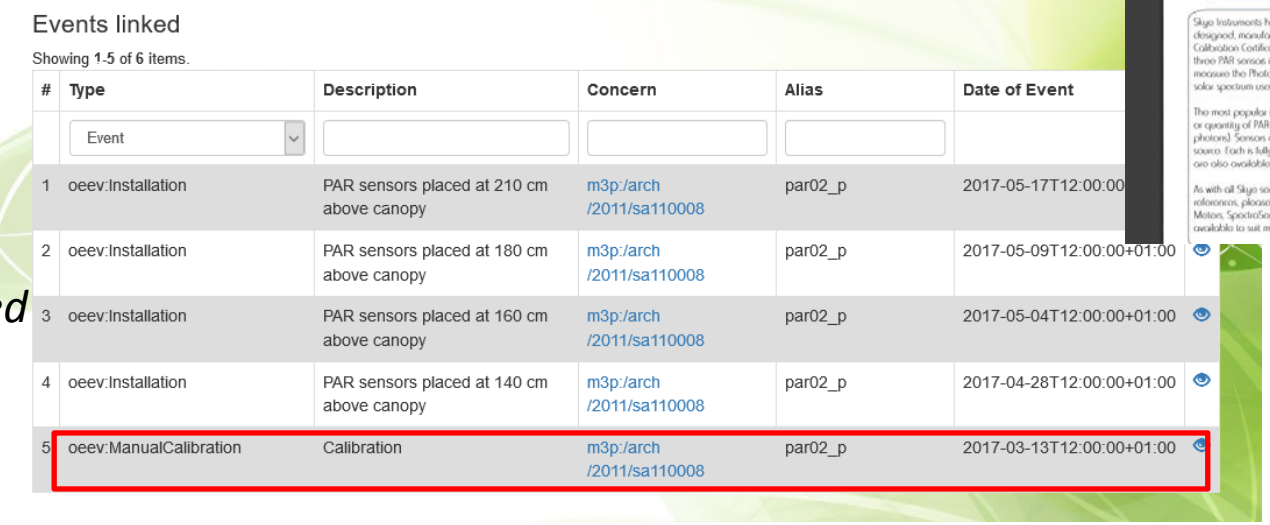
Sensors and environmental variables

Sensor characteristics, linked events and outputs are easily visualized



Sensor characteristics

Linked documents



Events linked

Object tracking

Phenotyping Hybrid Information System | Experimental Organization | Data | Tools | ••• | Llorenç Cabrera-Bosquet

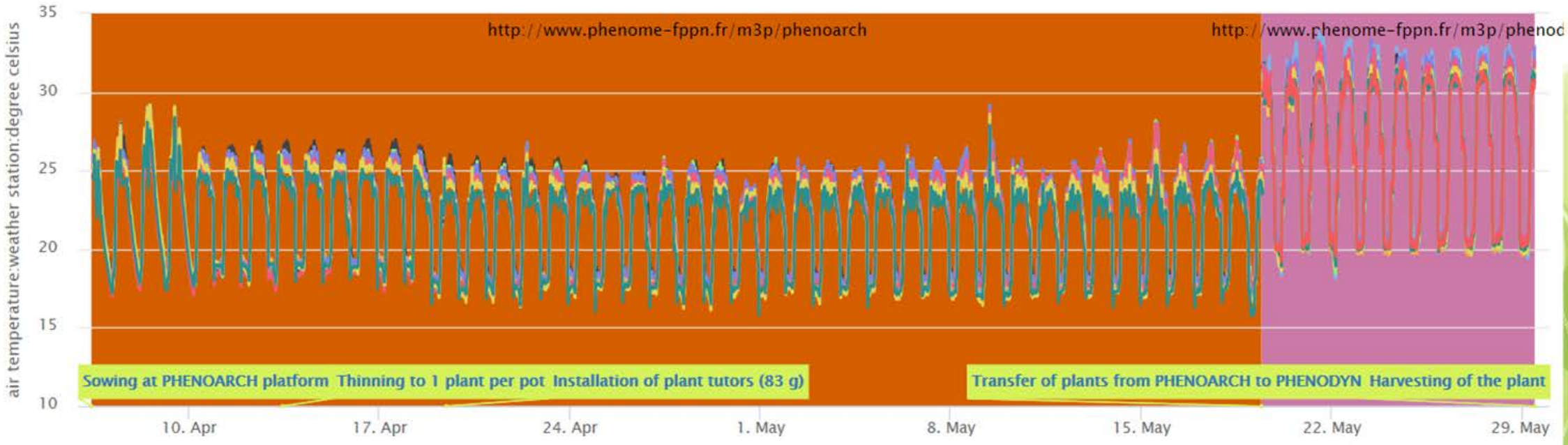
http://www.phenome-fppn.fr/m3p/arch/2017/c17000241

Code Variable

air temperature:weather station:degree celsius

Show

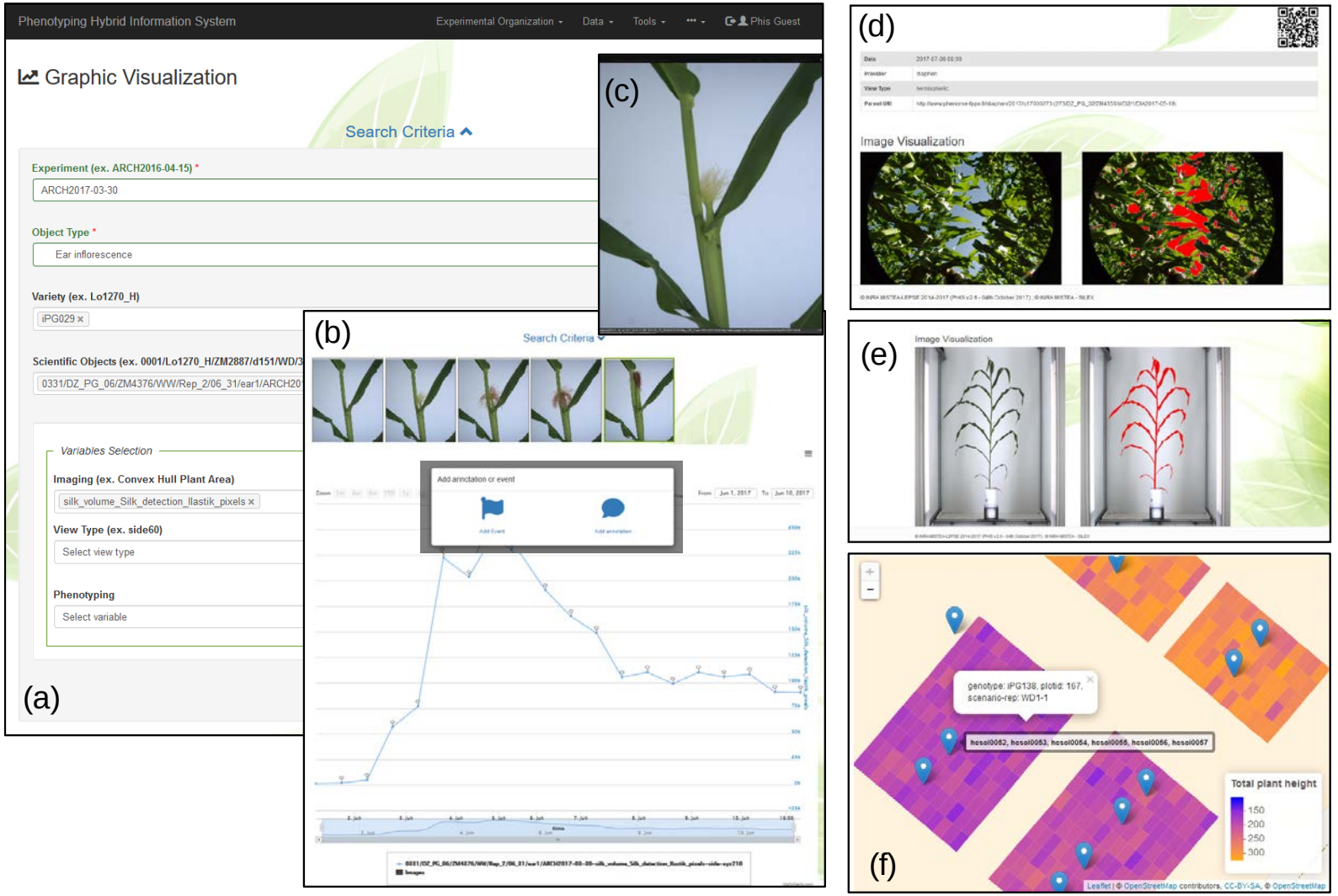
Data of air temperature:weather station:degree celsius



PHIS

Graphic visualisation

- Display of images, and time courses of phenotypic and environmental variables
- Raw images, segmented images and metadata can be displayed in both field and greenhouse experiments
- The interaction with the knowledge linking data with events and environmental sensor outputs.



A qui et à quoi sert PHIS?

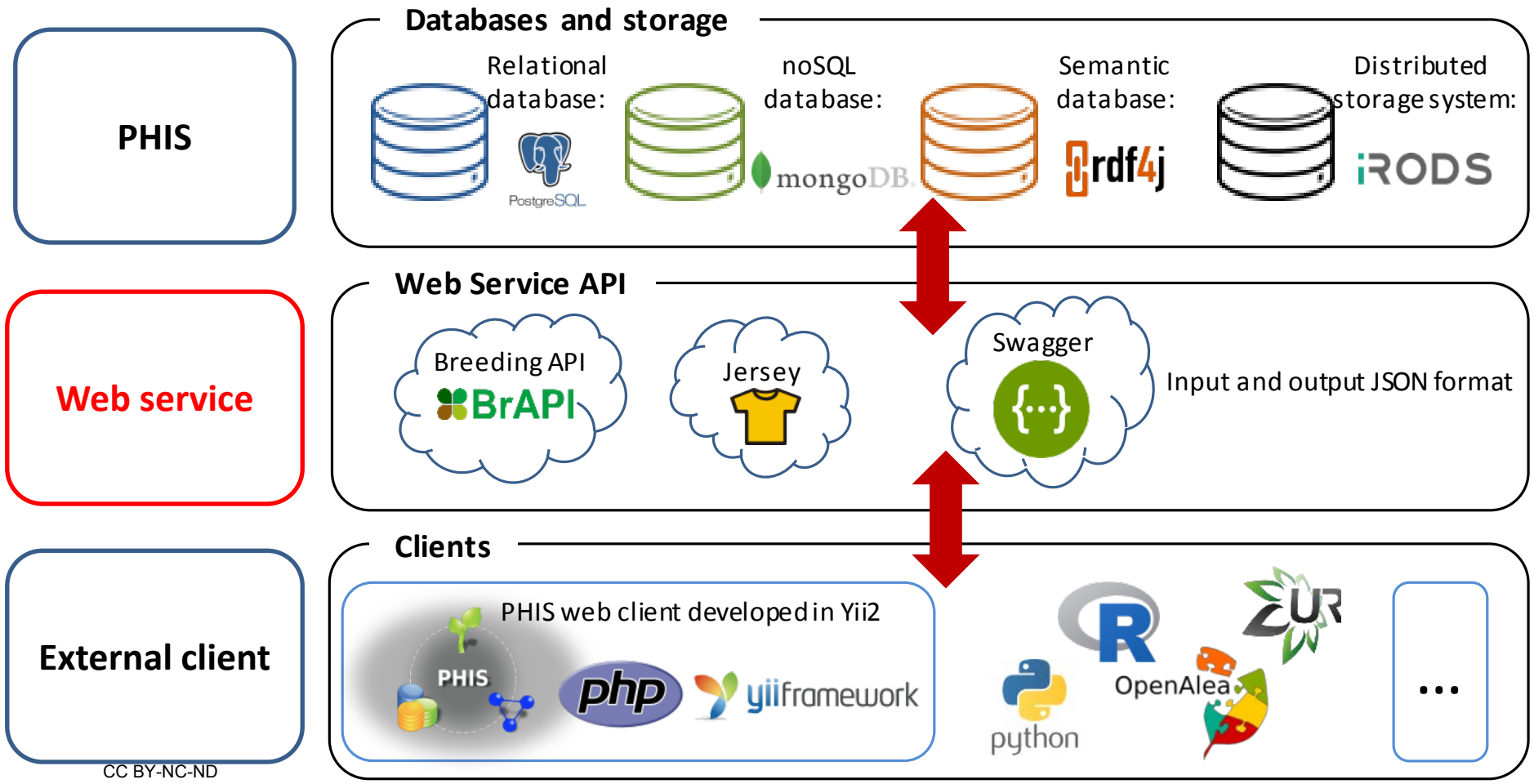
1. Platform users and staff
    - Management of phenotyping experiments / infrastructure
    - Data exploration and pre-analysis

2. External users (Interoperability with external platforms)
    - Re-analysis and meta-analyses
    - Modelling

3. Implemenation into other platforms

Web service API

Web service API allows one to virtually interoperate with any external client



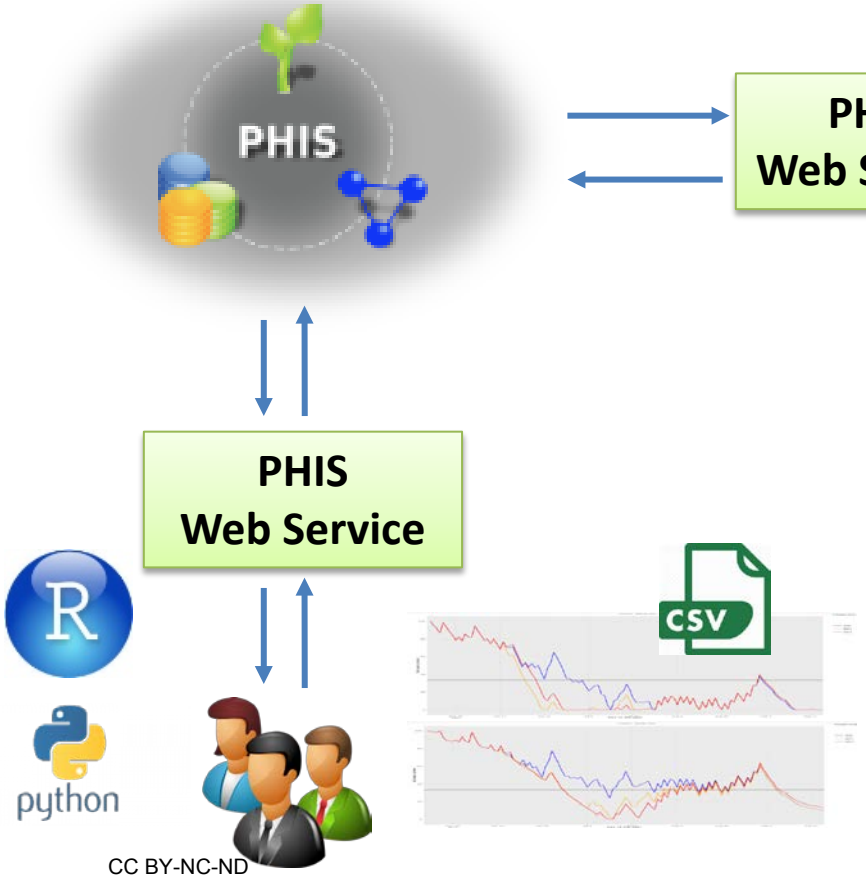
Interoperability with external DBs and resources

Web services allow to virtually integrate and import data to any external client

Integrate environmental data from a diversity of sources

Agriscope (Agribases)

Agroclim stations CIMEL



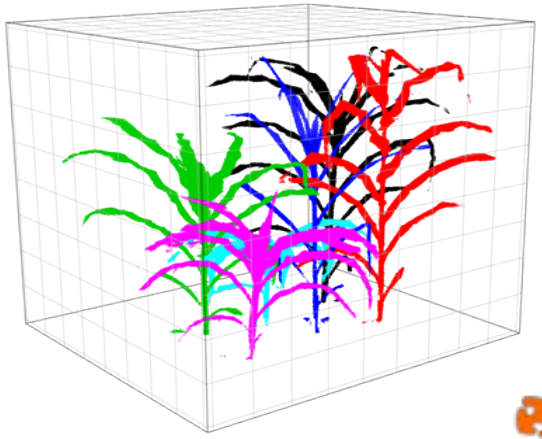
Data and graph query (via web user interface or R/Python client)

Interoperability: 3D reconstruction pipeline

PHIS may provide data and contextual information necessary to perform image analysis and re-analysis of datsets

- Images
- Camera settings
- Events
- Camera Calibration

PHIS Web Service



OpenAlea

Phenomenal Pipeline (3D reconstruction)

Artzet *et al.* in prep

A qui et à quoi sert PHIS?

1. Platform users and staff
   - Management of phenotyping experiments / infrastructure
   - Data analysis

2. External users (Interoperability with external platforms)
   - Integration of external data
   - Re-analysis and meta-analyses
   - Genetic and Modelling platforms

3. Implementation into other platforms

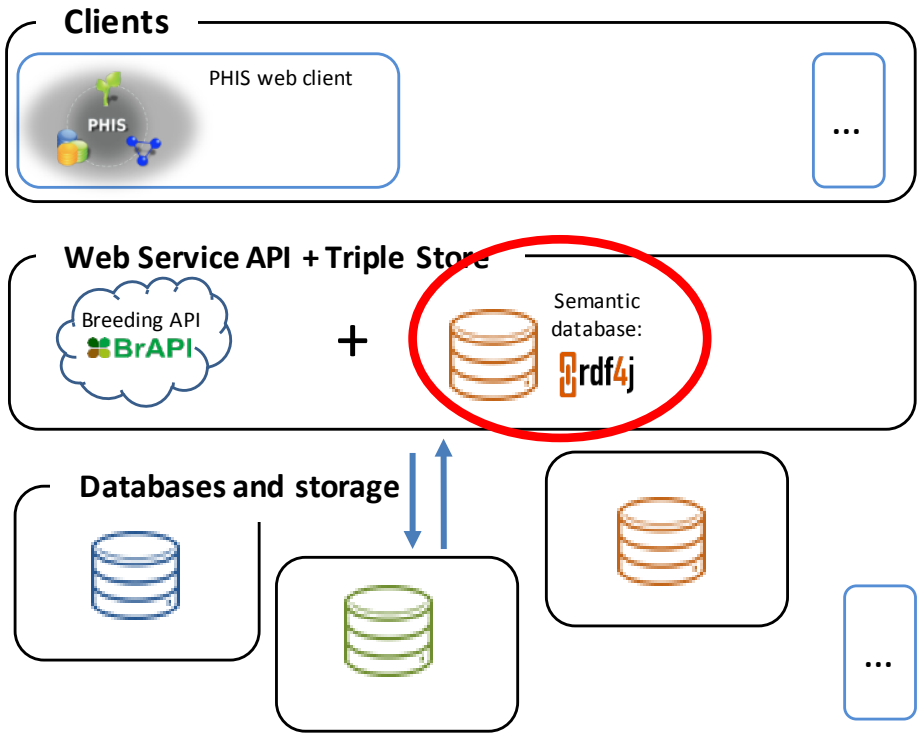
Implementation into other platforms ?

Can PHIS be implemented in other installations?

Ontology adaptation

Interoperability with local DBs

PHIS



Local system (proprietary) DBs, sensors, …

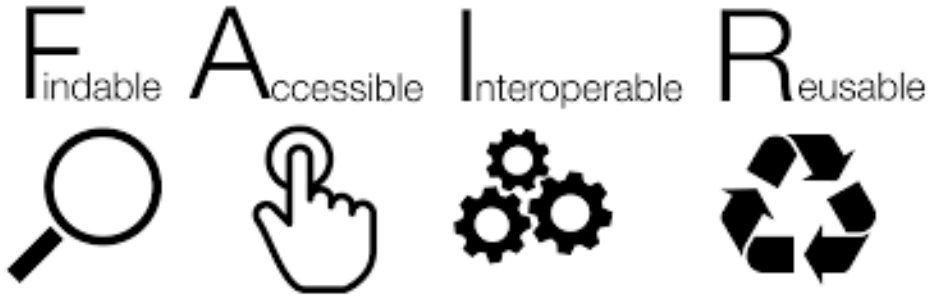
PHIS is being installed in other installations (5 PHENOME nodes, University of Nottingham and WUR)

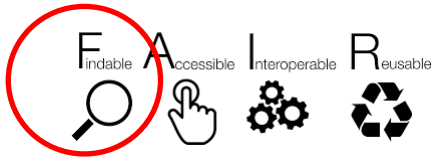
TAKE HOME MESSAGES

Why PHIS?

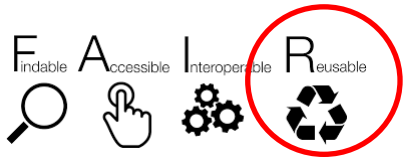
- Allows integration, management and visualisation of multi-source and multi-scale data
- Analysis and meta-analyses
- Integration of external data
- Flexible design => adaptability to other installations

TAKE HOME MESSAGES

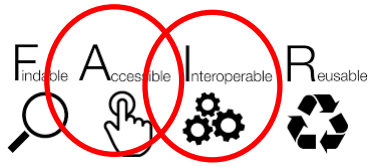




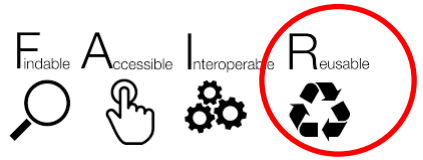
- International identification (**URI**)



- Semantics (**Ontologies**, standardized vocabularies)



- Portal interoperability (**Web services**) with external resources



- **Tracability** (eg. Events) and reproducibility

TAKE HOME MESSAGES

Skills needed:

- System administrator (software installation)
- Data scientist (data management, development of data pipeline and data analysis tools)
- Scientific knowledge (ontologies development)

TAKE HOME MESSAGES

Skills needed:

- System administrator (software installation)
- Data scientist (data management, development of data pipeline and data analysis tools)
- Scientific knowledge (ontologies development)

Acknowledgements

MISTEA

Pascal Neveu
Anne Tireau
Morgane Vidal
Arnaud Charleroy
Guilhem Heindrich
Nadine Hilgert
Isabelle Sanchez
Pierre-Etienne Alary

LEPSE

François Tardieu
Llorenç Cabrera-Bosquet
Vincent Nègre
Jonathan Mineau-Cesari
Nicolas Brichet

PHENOME partners

Patrick Moreau
Cyril Pommier
Romain Chapuis
…