

Data Management

S3: Experimental data (2)-Description

Monday 04/12/2023 - 14:00-16:00 (CET)

Cyril Pommier, Farzaneh Kazemipour-Ricci,
Isabelle Alic - INRAE

General objectives: Overview of data management for plant phenotyping - focus on FAIR data

Session 3

How to describe experiment metadata & data

Overview

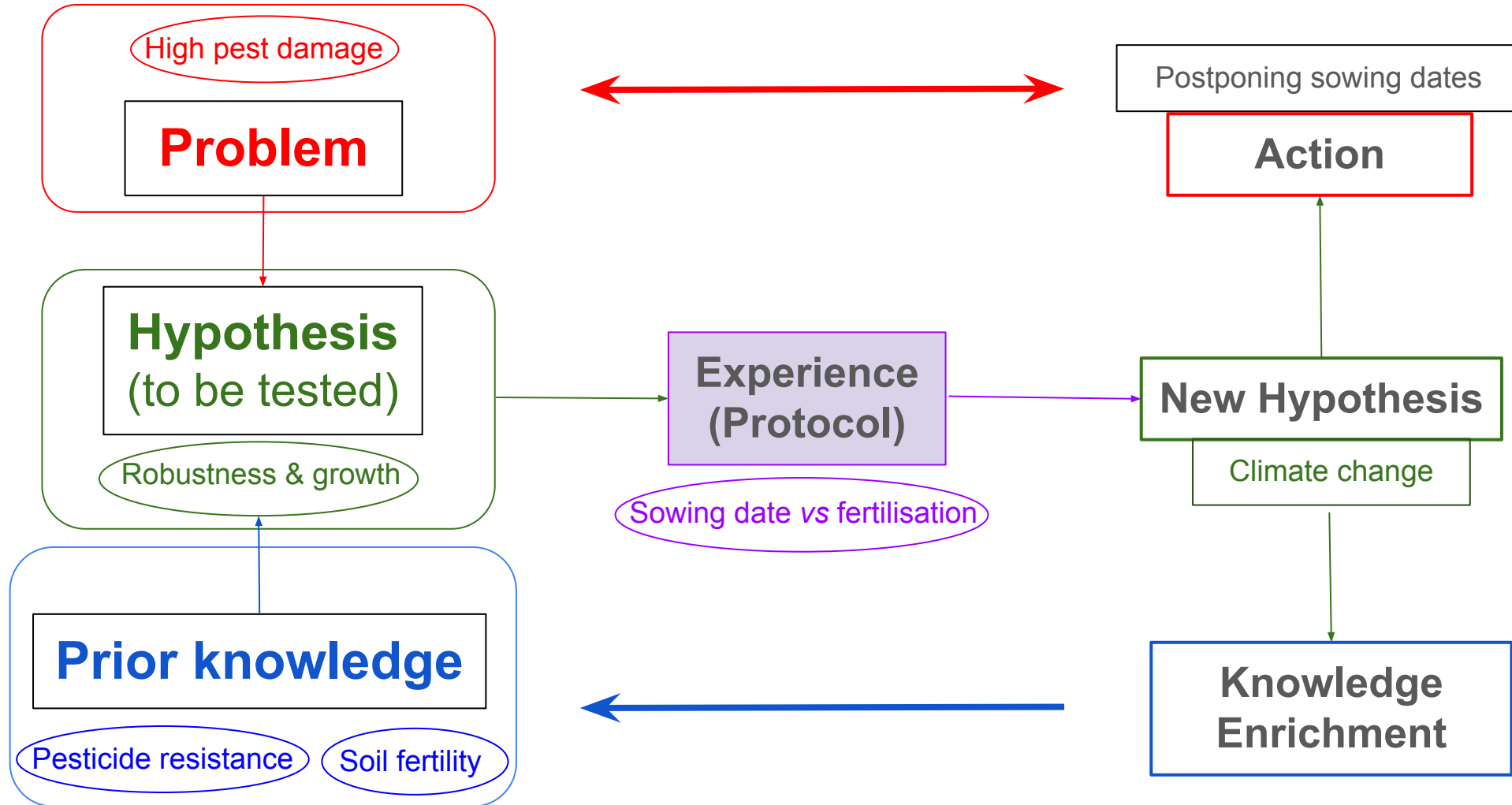
Experimental Data

- Introduction & Review of last session
- Minimum Information for Plant Phenotyping Experiments (MIAPPE)
- Managing experimental metadata in an IS - Illustration with PHIS

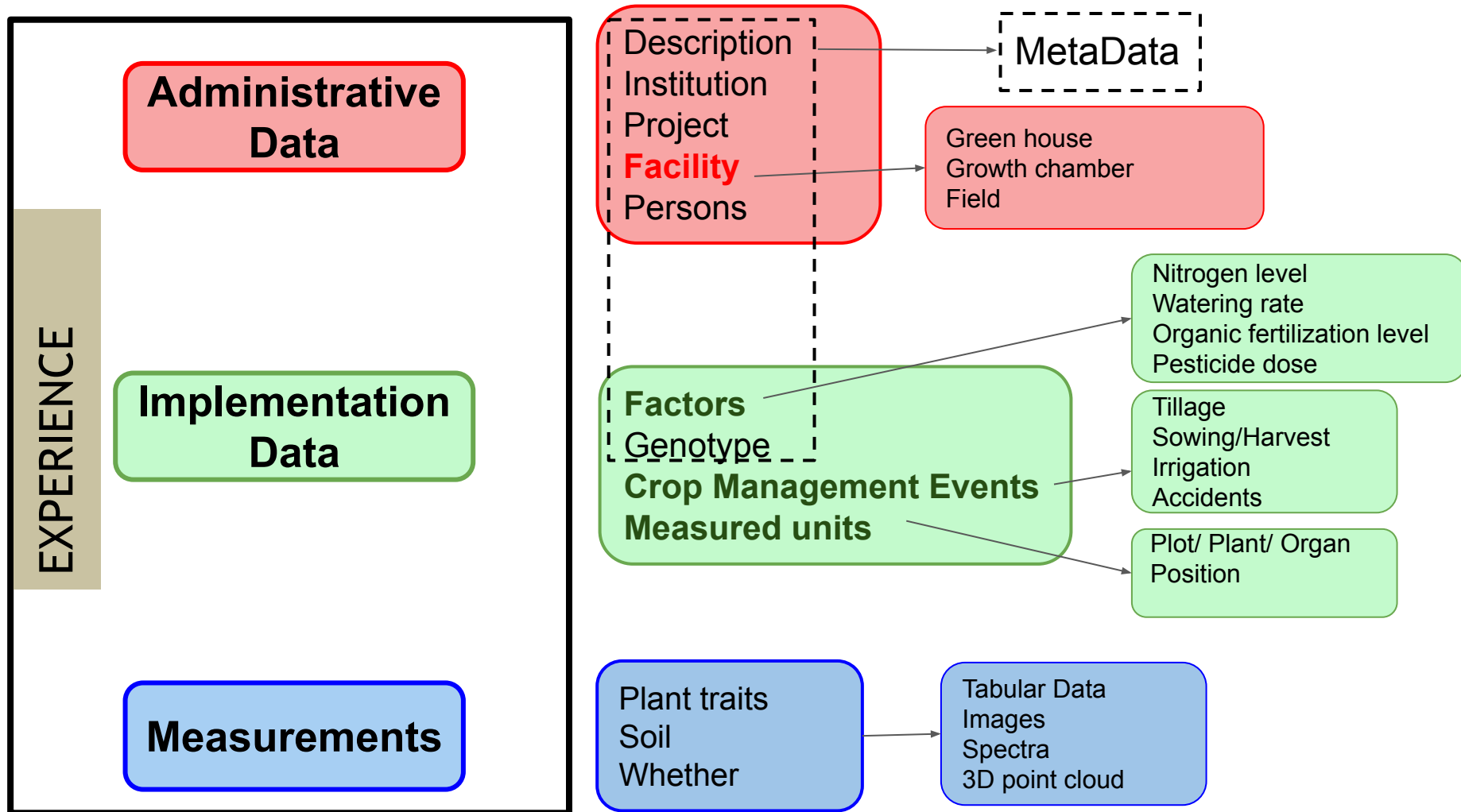
Review - Data storage, Protection & Rules

- Massive & complex phenomics data \Rightarrow No personal computer/storage
- Use Cloud Computing
 - Storage services
 - Processing services
 - EU solutions: EGI, EUDAT, FranceGrilles (coordination: EOSC)
- Data Management Plan & General Data Protection Regulation
 - FAIR Data \Leftrightarrow DMP
 - GDPR: individual rights but also more
 - RDM Kit: The Research Data Management toolkit for Life Sciences
https://rdmkit.elixir-europe.org/plant_sciences

Introduction: Plant Phenomics Experiment



Introduction: Plant Phenomics Experiment



Minimum Information for Plant Phenotyping Experiments (MIAPPE)

Data standards for plant research data in Phenomic and beyond

PLANT DATA STANDARDS: WHO

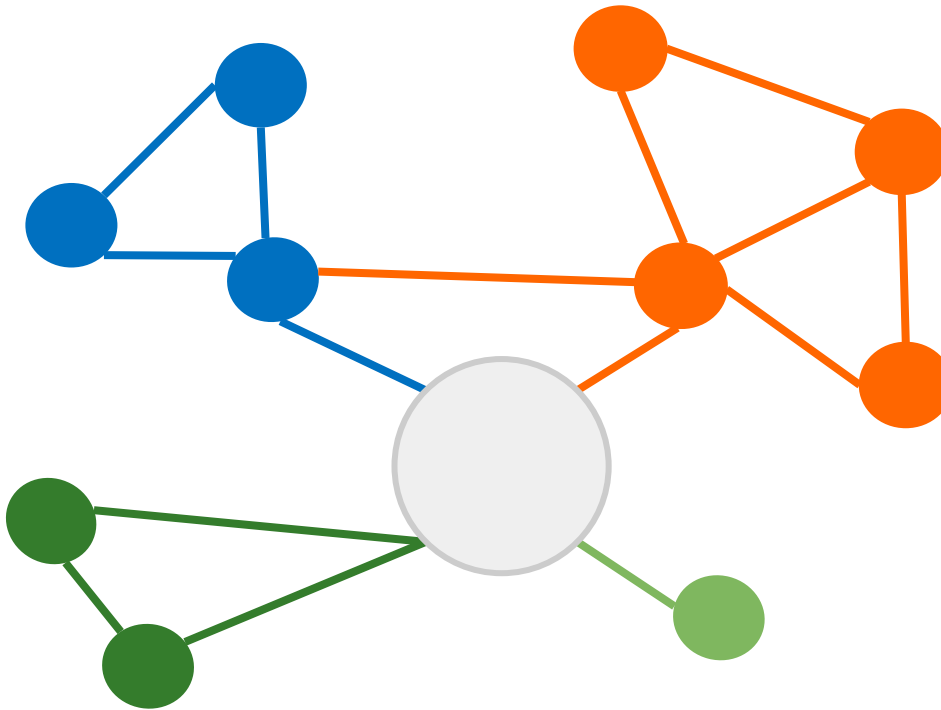


Interoperability in international network

National Networks



Global Networks



European Networks



International data standards



Plant data standards: contributors

Gathered to solve FAIR data problems for plants

ELIXIR



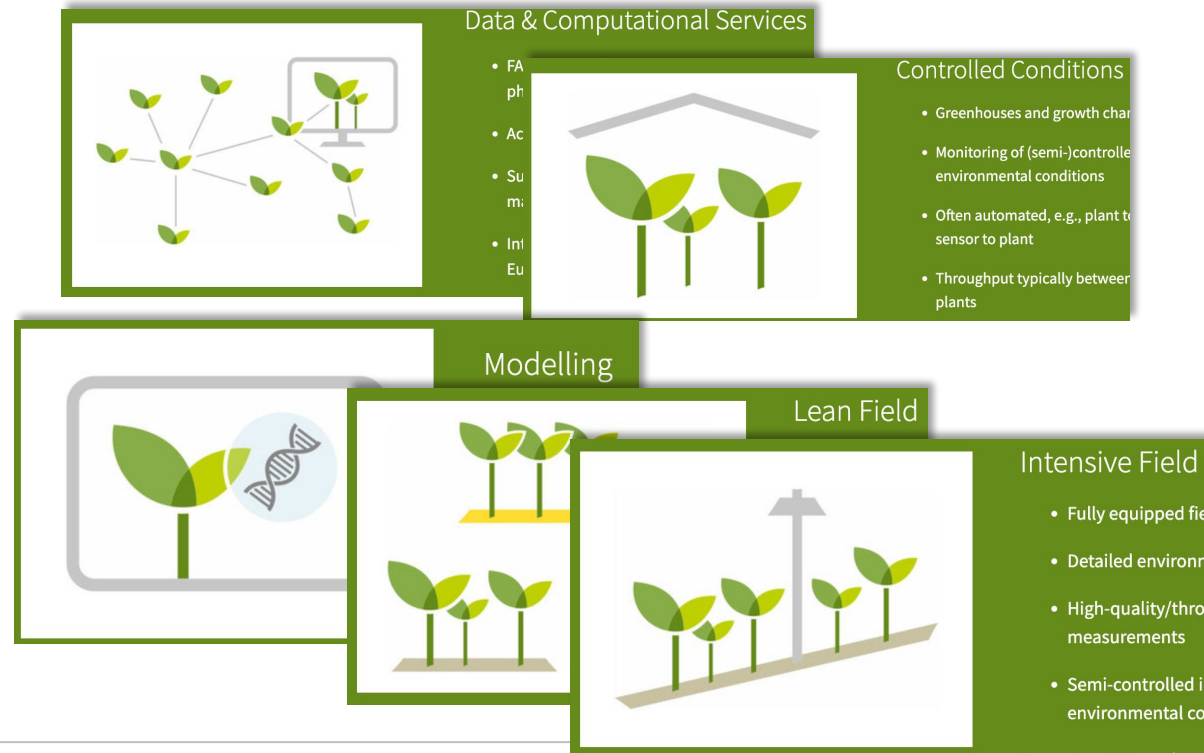
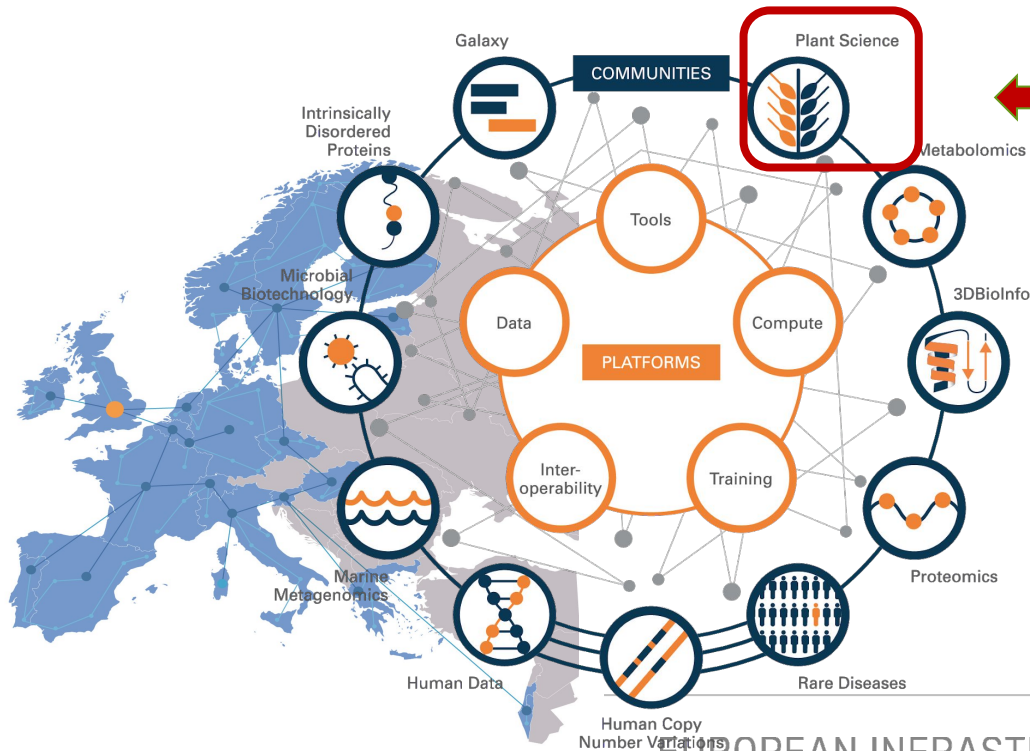
- European Infrastructure for life sciences
- Slovenia, France, Germany, Portugal, UK, ...
- <https://elixir-europe.org/communities/plant-sciences>
- FAIR data management, software, training



EMPHASIS



- European Infrastructure for Plant Phenotyping
- France, Germany, Belgium, UK, ...
- <https://www.plant-phenotyping.eu>



EUROPEAN INFRASTRUCTURE FOR PLANT PHENOTYPING

Plant data standards: contributors

Gathered to solve FAIR data problems for plants

CGIAR

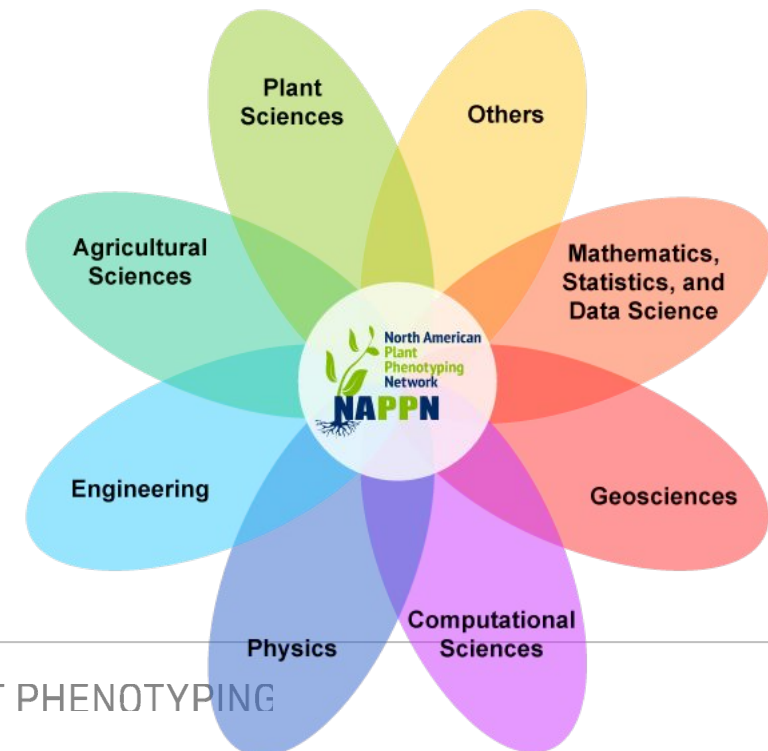
- Consortium of International Agricultural Research Centers
- CropOntology, MIAPPE
- <https://www.cgiar.org/>



NAPPN



- North American Plant Phenotyping Network
- Regional partner IPPN
- <https://nappn.plant-phenotyping.org/>

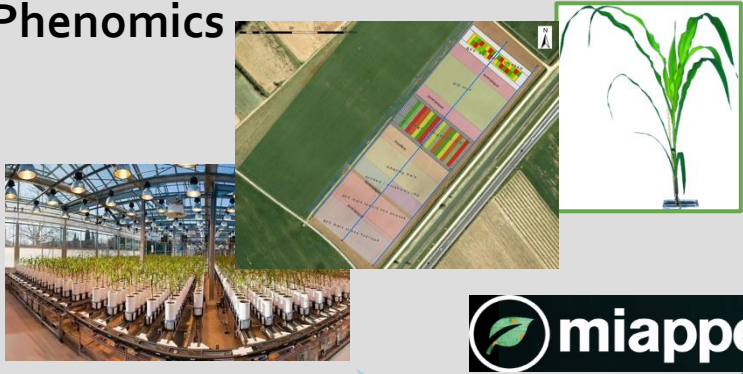


PLANT PHENOTYPING STANDARDS: WHY



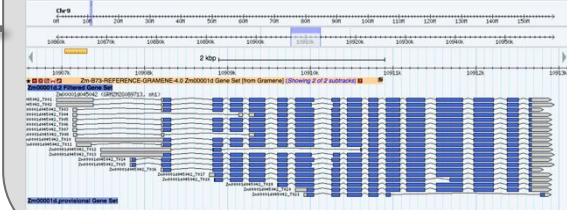
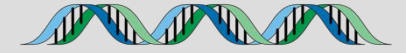
Integration * Omics / Environment / Phenomic

Phenomics



F_{indable} A_{ccessible} I_{nteroperable} R_{eusable}

Genetics
Genomics
Omics



Dispersed
Heterogenous
Getting Standardized

Plant Breeding
Genetic variations by Traits

Climate Change Studie
Genotype by Environment

Mostly centralized
Homogenous data
Heterogenous metadata

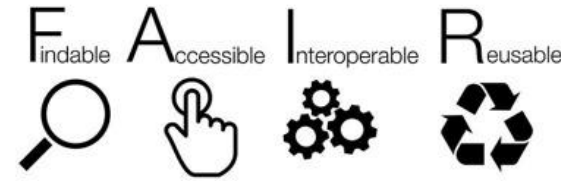
Dispersed
Heterogenous

F_{indable} A_{ccessible} I_{nteroperable} R_{eusable}

Environment



Why standardized data for phenotyping?



- Allow them to be re-used (including by yourself!)
 - Metadata/description experiment (who, why, where, how)



- Integration, automated processing
 - Linked data between datasets via identification of pivot objects

Phenotype 1 = measurement on a cultivar in an environment-GPS1-time1

Phenotype 2 = measurement on a cultivar in an environment-GPS2-time2

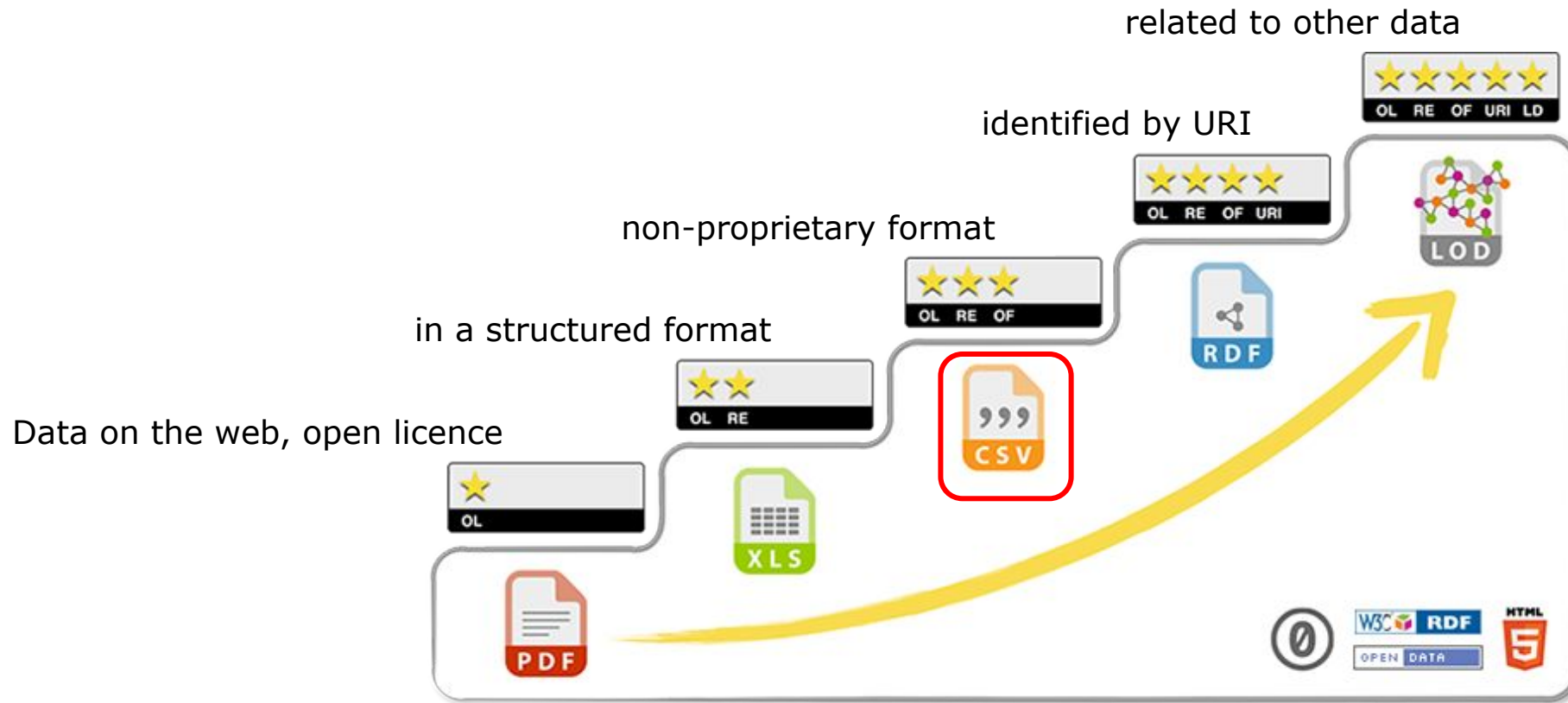
Genotype = observed marker's alleles on a cultivar

Climate 1 = climatic data at GPS1-time1



- Enable the discovery of data and knowledge:
 - Metadata, controlled vocabularies, ontologies

5 stars Open Data ⇒ Let's be pragmatic!



Progressing towards FAIR and Open Data requires a multidisciplinary cooperation :

- Biologists
- Bioinformaticians
- Specialists of ontologies/semantics

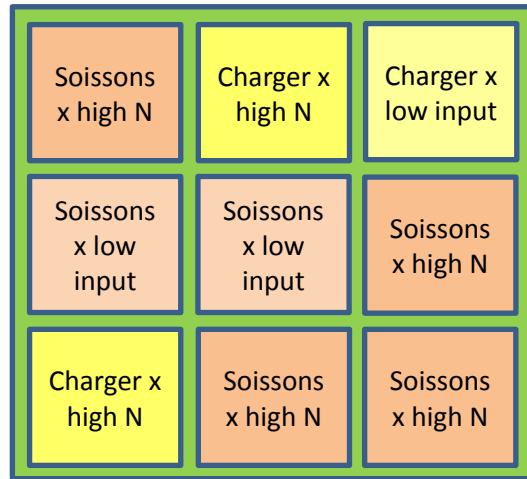
Phenotyping data life cycle

Data acquisition

Data computation

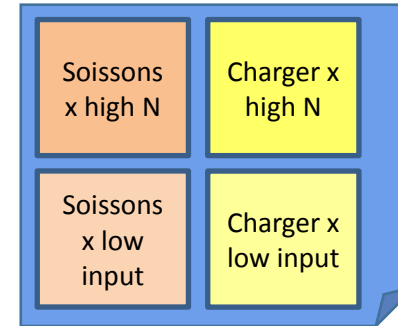
Knowledge

« Raw » data, pheno/env measures, variables



Derivation, Reduction

« computed » data, reduced, indicators



Publication

Variety charger is resistant to fusariose under intensive cultural practice

Genotype	Treatment	N input	Date	Rep	Fusariose
Soissons	low input	15,32253129	15/11/2011	1	5
Soissons	low input	15,31430556	16/11/2011	2	7

VARIABLES

Traceability

Raw measures

Data Cleaning

Platform IS (PhenoLims, PHIS, PIPPA ...)

Reproducibility & Provenance

Genotype	Treatment	Fusariose
Soissons	low input	6

INDICATORS

New computation for each scientific question

One raw dataset → many computed datasets

Raw data long term conservation

PLANT DATA STANDARDS: OVERVIEW



MIAPPE, Crop ontology, MCPD

Data Standards

Semantic



- Description of the data
- Controlled vocabularies: term name and definitions
- Ontologies: semantic links between terms
- Phenotyping/environment
- *Biologist* driven

Structure



- Formatting, Modelling and Organizing Data
 - Standards : CSV, VCF, GFF, MIAPPE , etc...
 - Whole Dataset organization
 - greenhouse, field, multiannual, multilocal
 - Includes Variables and plant material
- Biologist & Computer scientist* driven

Persistent Unique Identifiers

(sensor, gene, accessions, trait, URI, DOI,...)

Training

(guidelines, checklists, ...)

Technical

- Data integration and sharing
- Interoperability : tools and systems
 - GA4GH
 - Breeding API www.brapi.org
- *Computer scientist* driven



Plant Structure Standard



Minimal Information About Plant Phenotyping Experiment

www.miappe.org

- Several stakeholders
 - Elixir, Emphasis, Bioversity, North American PPN
- Open community:
 - Request for comments (next one within 3 months)
 - Github Feature requests
 - Mailing lists
 - Meetings, working groups
 - First version: 2015
- Annual plants (field crops, models, etc.)
- Perennial plants (vines, forest or fruit trees, etc.)

MIAPPE					
line #	MIAPPE Check list	Definition	Example	Format	Cardinality
DM-1	Investigation	Investigations are research programmes with defined aims. They can exist at various scales (for example, they could encompass a grant-funded programme of work, the various components comprising a peer-reviewed publication, or a single experiment).			1 per MIAPPE submission
DM-2	Investigation unique ID	Identifier comprising the unique name of the institution/database hosting the submission of the investigation data, and the accession number of the investigation in that institution.	EBI:12345678	Unique Identifier	0-1
	Investigation title	Human-readable string summarising the investigation.	Adaptation of Maize to Temperate Climates, Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with	Free text (short)	1
Environment					
ENV-1	Non-exhaustive list of Environment Parameters.				
line#	Environment parameters	Definition	Example	Format	
ENV-2					
ENV-3	Growth facility				
ENV-4	Air temperature	List of hourly air temperature throughout the experiment.	22 °C	Numeric	
ENV-5	Organ temperature	List of hourly organ temperatures throughout the experiment.	18 °C	Numeric	
Experimental Factors					
TR-1	Non-exhaustive list of Experimental Factors that can be applied.				
line #	Factor type	Definition	Example factor values	Format	
TR-2	Seasonal environment	A plant treatment (EO:0001001) involving an exposure to a given conditions of regional seasons.	Spring season; dry season	Plant Environment Ontology:'EO_0007038'	
TR-3	Air treatment regime	The treatment involving an exposure to wind/air with varying degree of temperature, which may depend on the study type or the regional environment.	28/25°C (Day/Night)	Plant Environment Ontology:'EO_0007161'	
TR-4	Soil temperature regime	A physical plant treatment (EO:0007316) involving an exposure to varying degree of temperature, which may depend on regional environment.	27/25°C (Day/Night)	Plant Environment Ontology:'EO_0007161'	
TR-5					

Plant Structure Standard: MIAPPE

Minimum Information for Biological and Biomedical Investigations

A collection of the historical MIBBI foundry reporting guidelines. The minimum information standard is a set of guidelines for reporting data derived by relevant methods in biosciences. If followed, it ensures that the data can be easily verified, analysed and clearly

- Biologist Friendly
 - Clear definitions and examples
 - Excel templates
 - Training
- Minimum and sufficient list of metadata
- Computer Scientist Friendly
 - Explicit data model: ISA-Tools and Breeding API (BrAPI) compatibility.
 - Validation framework and toolbox

Plant Structure Standard

- Imports established standards

- MultiCrop Passport Descriptors (MCPD): identification of plant material
- Crop Ontology: description of measurement and calculation methods

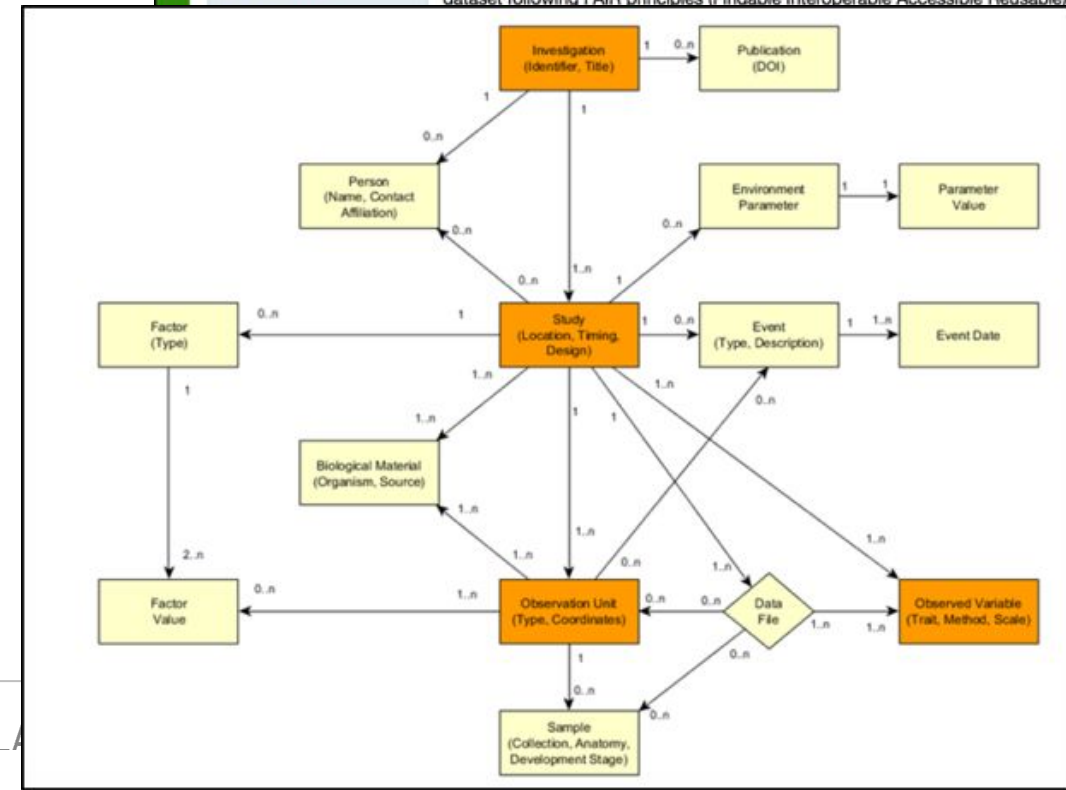
- Specification with several implementations

- Semantic: Plant Phenotyping Experiment Ontology
- Web Service: Breeding API
- File template Excel
- File Archive: ISA Tab
- Information systems: GnpIS, data.inrae.fr, BrAPI compliant databases(PHIS, PIPPA ...)

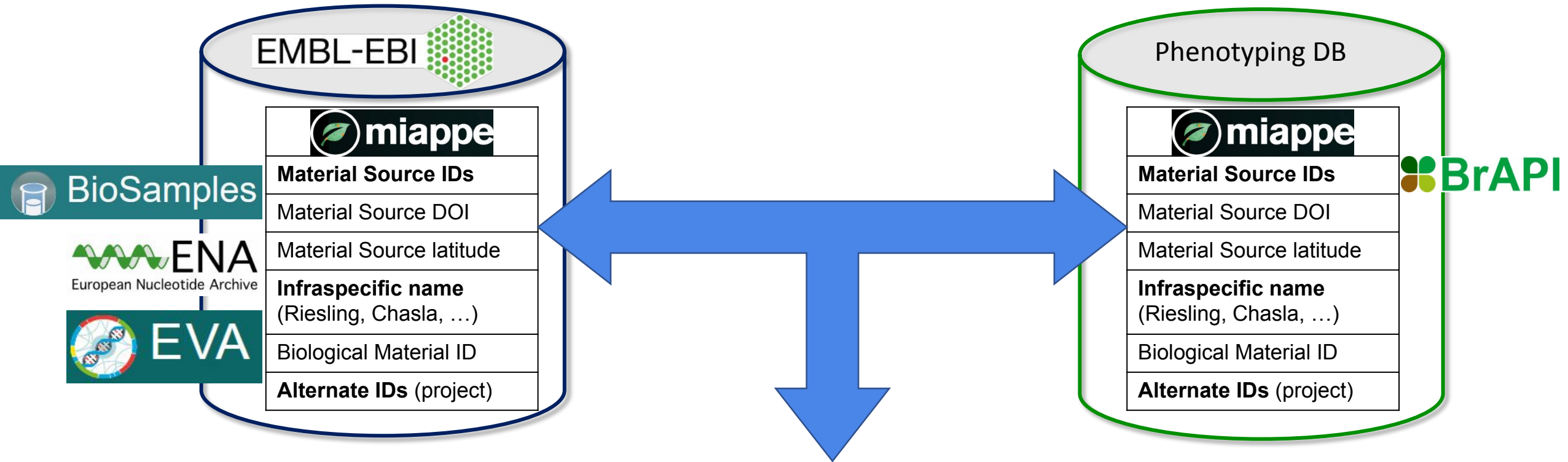
The screenshot shows the miappe website interface. At the top, there is a navigation bar with 'About' and 'Developers' links. Below that is the 'AgroPortal LIRMM' logo and a search bar with 'Browse', 'Search', 'Mappings', 'Recommender', 'Annotator', and 'Projects' options. The main heading is 'Plant Phenotype Experiment Ontology' with the RDA (Research Data Alliance) logo. Below the heading are links for 'Summary', 'Classes', 'Properties', 'Notes', 'Mappings', and 'Widgets'. The 'Details' section includes a table with the following information:

ACRONYM	PPEO
VISIBILITY	Public
DESCRIPTION	The Plant Phenotyping Experiment Ontology, PPEO, is an implementation of the Minimal Information About Plant Phenotyping Experiment. It lists and organises all the informations necessary to describe and reuse a phenotyping dataset following FAIR principles (Findable Interoperable Accessible Reusable).

Logos for Bioversity International, EMPHASIS, and elixir are also visible.



Data Integration between silos, From Phenotyping to Genotyping



Community data discovery portals

The screenshot shows the **FAIDARE FAIR Data-finder for Agronomic REsearch** portal. The URL is <https://urgi.versailles.inrae.fr/faidare/>. The page features a search bar for keywords, a "Reset all" button, and filters for "Germplasm" and "Trait".

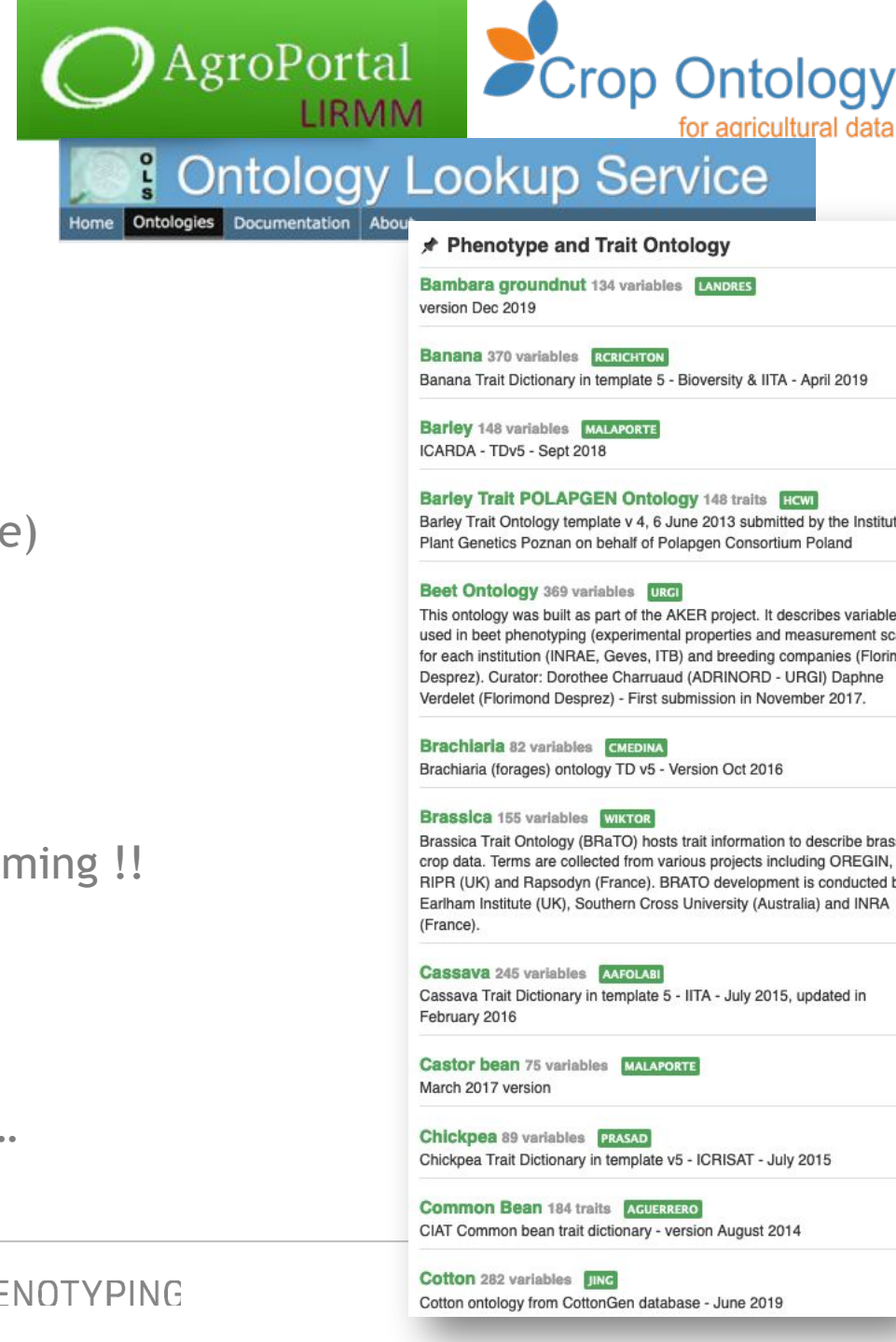
Sources

- URGI GnpIS (81,335)
- EBI European Nucleotide Archive (44,975)
- CIRAD TropGENE (722)
- VIB PIPPA (692)
- IBET BioData (67)
- IWGSC@GnpIS (18,814,632)
- Evoltree@GnpIS (5,354)
- OpenMinTeD@GnpIS (3,392)
- EBI Ensembl Plants

Crops
(common name, species, genus, subtaxa & synonyms)

Germplasm list
(panel, collection & population)

Phenotype Semantic Standard: Ontologies



The screenshot shows the AgroPortal LIRMM and Crop Ontology logos at the top. Below them is the 'Ontology Lookup Service' header with navigation links for Home, Ontologies, Documentation, and About. The main content area is titled 'Phenotype and Trait Ontology' and lists several ontologies with their respective variable counts and curators:

- Bambara groundnut** 134 variables **LANDRES**, version Dec 2019
- Banana** 370 variables **RCRICHTON**, Banana Trait Dictionary in template 5 - Bioversity & IITA - April 2019
- Barley** 148 variables **MALAPORTE**, ICARDA - TDv5 - Sept 2018
- Barley Trait POLAPGEN Ontology** 148 traits **HCWI**, Barley Trait Ontology template v 4, 6 June 2013 submitted by the Institut Plant Genetics Poznan on behalf of Polapgen Consortium Poland
- Beet Ontology** 369 variables **URGI**, This ontology was built as part of the AKER project. It describes variable used in beet phenotyping (experimental properties and measurement sc for each institution (INRAE, Geves, ITB) and breeding companies (Florim Desprez). Curator: Dorothee Charruaud (ADRINORD - URG) Daphne Verdelet (Florimond Desprez) - First submission in November 2017.
- Brachiaria** 82 variables **CMEDINA**, Brachiaria (forages) ontology TD v5 - Version Oct 2016
- Brassica** 155 variables **WIKTOR**, Brassica Trait Ontology (BRATO) hosts trait information to describe bras crop data. Terms are collected from various projects including OREGIN, RIPR (UK) and Rapsodyn (France). BRATO development is conducted b Earham Institute (UK), Southern Cross University (Australia) and INRA (France).
- Cassava** 245 variables **AAFOLABI**, Cassava Trait Dictionary in template 5 - IITA - July 2015, updated in February 2016
- Castor bean** 75 variables **MALAPORTE**, March 2017 version
- Chickpea** 89 variables **PRASAD**, Chickpea Trait Dictionary in template v5 - ICRISAT - July 2015
- Common Bean** 184 traits **AGUERRERO**, CIAT Common bean trait dictionary - version August 2014
- Cotton** 282 variables **JING**, Cotton ontology from CottonGen database - June 2019

www.cropontology.org

- Description of plant traits and characteristics
- Crop Ontology \Rightarrow Representation model, ie Framework
 - Trait + Method + Scale
 - Scale: unit, rating scale
 - Trait (e.g. leaf area): Entity (e.g. leaf) + Characteristic (surface)
 - (cf. *Quadruplet in Session 4*)

• www.cropontology.org \Rightarrow Repository of species ontologies

• \Rightarrow Can't find your variables (with the right method) ?

• \Rightarrow Contribute to existing ontologies

BUT !! Contribution to ontologies is potentially time-consuming !!

• \Rightarrow Or use Emphasis PHIS shared instance

<http://resources.plant-phenotyping.eu/emphasis/app/>

• \Rightarrow Or create ontologies by community

- Use existing template/tool : cropontology, MIAPPE, PHIS, ...
- Most pragmatic approach

MIAPPE: IMPLEMENTATIONS



Excel template

ISA Tab

Breeding API

Informations System

Formal Ontology

MIAPPE Excel template

- Available on MIAPPE github

https://github.com/MIAPPE/MIAPPE/tree/master/MIAPPE_Checklist-Data-Model-v1.1/MIAPPE_templates

- One sheet per MIAPPE section

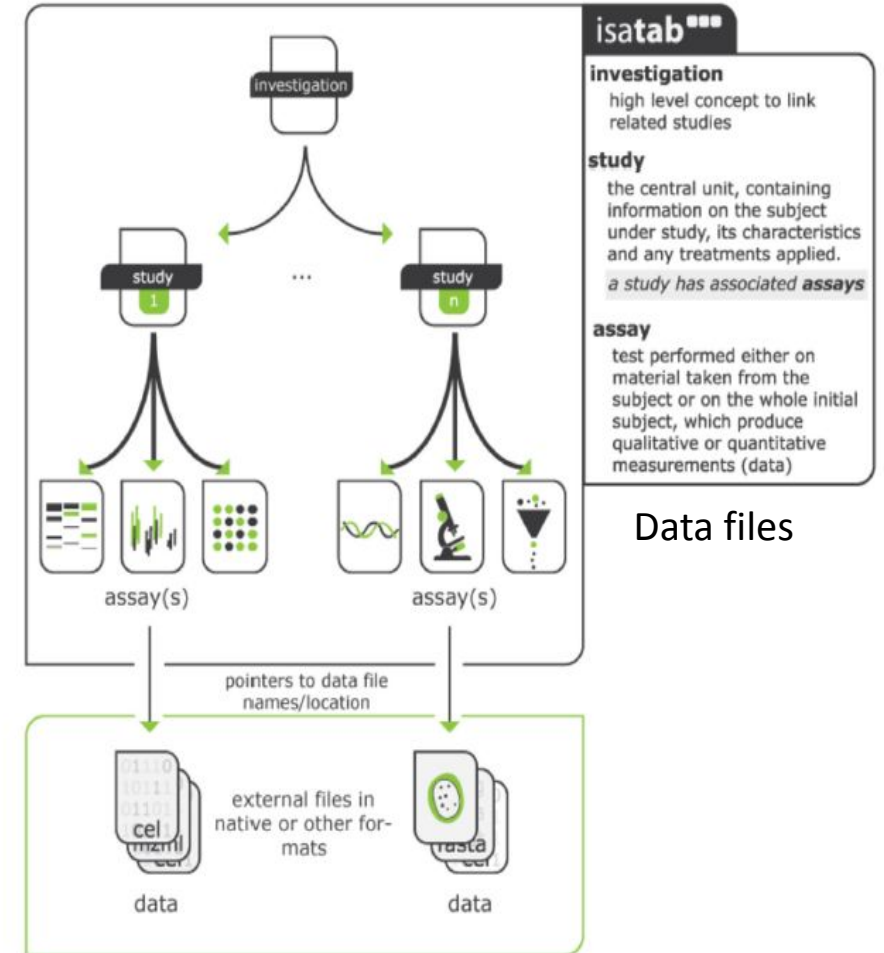
Environment	Investigation	Study	Person	Data file	Biological Material	Environment	Appendix I - Environment	Exp. Factor
General Definitions								

- Examples & Documentation

Field	Study unique ID*	Study title*	Study description	Start date of study*	End date of study	Contact institution*
Definition	Unique identifier comprising the name or identifier for the institution/database hosting the submission of the study data, and the identifier of the study in that institution. <i>If the study has no official unique ID, assign your own following the schema: study_1, study_2, study_..., study_n.</i>	Human-readable text summarising the study	Human-readable text describing the study	Date and, if relevant, time when the experiment started	Date and, if relevant, time when the experiment ended	Name and address of the institution responsible for the study.
Example	EBI:12345678; http://phenome-fppn.fr/maugio/2013/t2351	2002 evaluation of flowering time for a panel of 375 maize lines at the experimental station of Maugio (France).	2002 evaluation of male and female flowering time for a panel of 375 maize lines representing the worldwide genetic diversity at the experimental station of Maugio, France.	2002-04-04 2006-09-27T10:23:21+00:00	2002-11-27	UMR de Génétique Végétale, INRA – Université Paris-Sud – CNRS, Gif-sur-Yvette, France
Format	Unique identifier	Free text (short)	Free text	Date/Time (ISO 8601, optional time zone)	Date/Time (ISO 8601, optional time zone)	Free text (short)

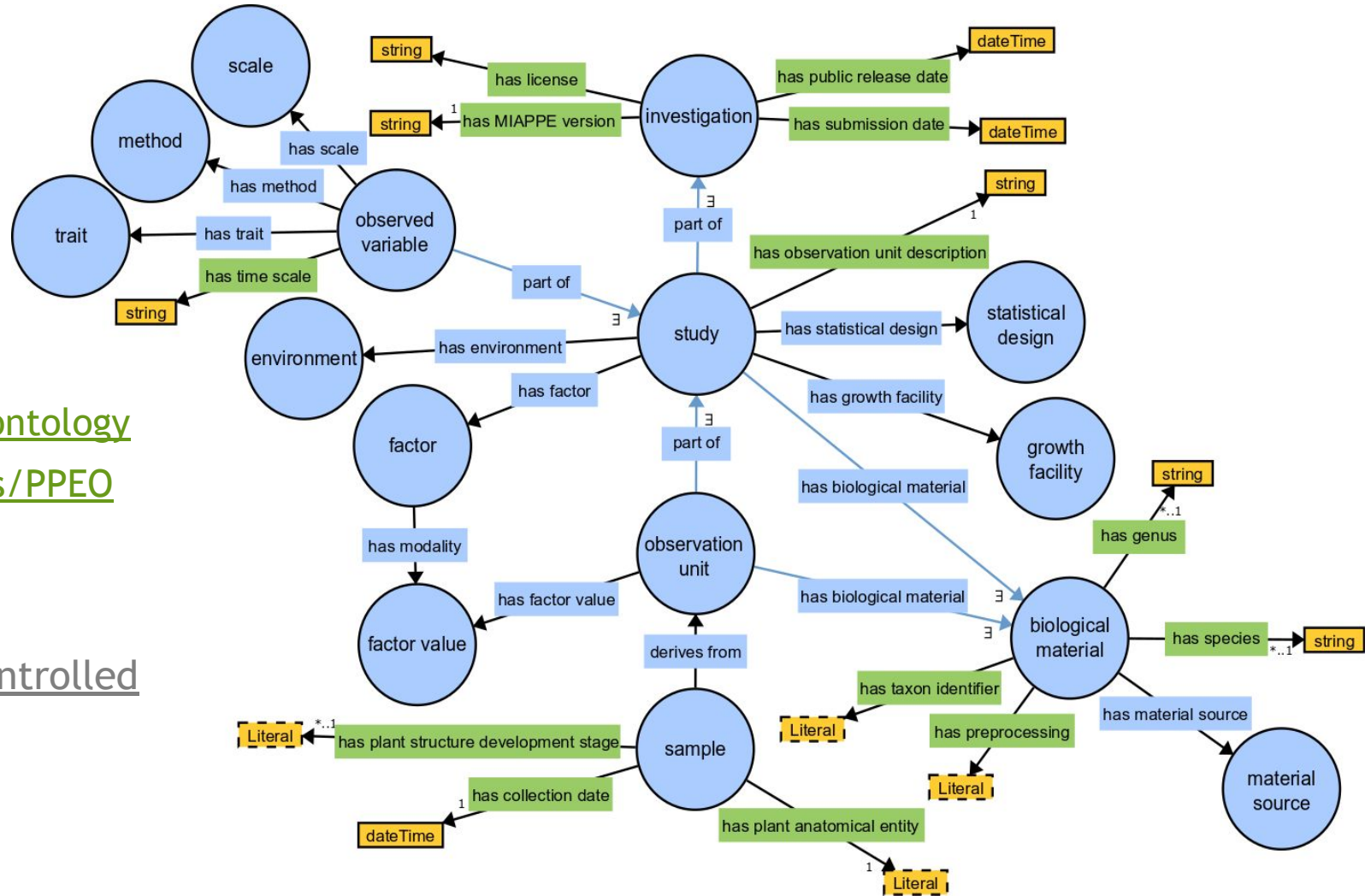
MIAPPE File Archive

- ISA Tab for Phenotyping
 - Investigation/Study/Assay
 - Zip Archive
 - MIAPPE Metadata
 - Raw data
 - CSV
 - Images or binary files
 - Reference to image archive (URI/URL)
 - Elaborated data
 - CSV
 - Provenance
- Training and improvements
 - File and metadata curation
 - Elixir
 - Célia Miguel & Anne Françoise Adam Blondon
 - BrAPI to IsaTab portable tool
 - Elixir implementation study for MIAPPE Validation



MIAPPE Semantic: PPEO

- MIAPPE OWL Implementation
 - <https://github.com/MIAPPE/MIAPPE-ontology>
 - <http://agroportal.lirmm.fr/ontologies/PPEO>
- Data model representation
- Formal concepts and constraints
- Ontology for Computer scientists, not a controlled vocabulary for biologists



Plant Technical Standard : Breeding API

- <http://brapi.org/>
- International collaboration
 - Standard Open Web Service API
 - Information Exchange, Main target: Breeding
 - Excellence in Breeding platform (CGIAR, Peter Selby)
- Major Elixir, Emphasis Contribution
 - Phenotyping specification
- Connect data repositories and tools:
 - Genotype visualization (Flapjack)
 - Studies graph preview and filtering
 - BrAPPS : Tools integrable in any BrAPI compliant System
 - <https://www.brapi.org/brapps.php>
 - R analysis environment
 - Field data capture
 - FAIR Data discovery → Elixir FAIDARE

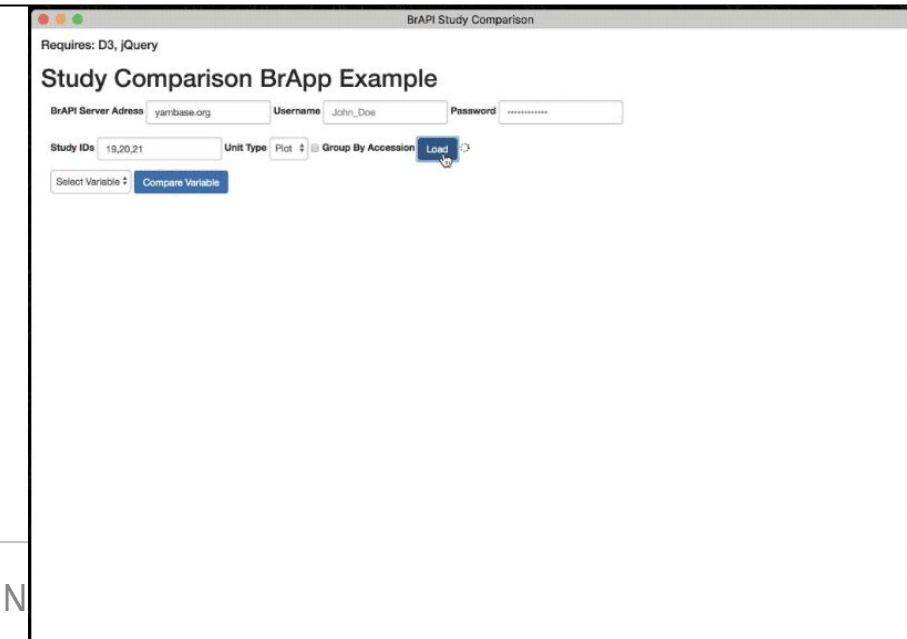


T3
CIRAD
GOBII
Wageningen
Cornell
iPlant

BILL & MELINDA
GATES foundation

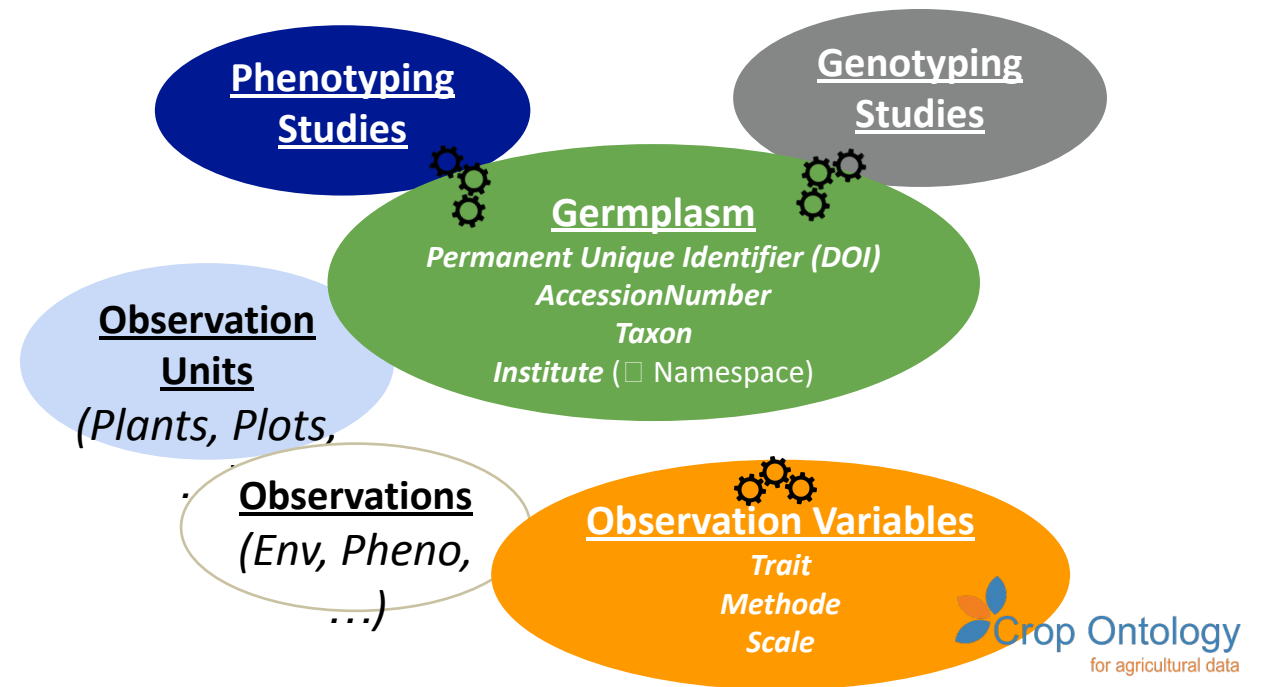


Selby *et al.* *Bioinformatics* (2019),
doi.org/10.1093/bioinformatics/btz190



- BrAPI has defined a **Standardized** set of data model structures to communicate the basic information of plant breeding
- BrAPI is a technical **Specification** which software developers can easily turn into code which communicates using the **Standard**

- MIAPPE compliant
- MIAPPE Investigation ⇒ BrAPI Trial
- MIAPPE Study ⇒ BrAPI Study
- Biological Material Germplasm MCPD dedicated BrAPI calls
- Observation variable BrAPI calls
- Datafile or Observation Unit



Standards Adoption



- French community for Plant Genetic Resources data exchange standard (Florilège, Olga, GnpIS, FAIDARE)
- ELIXIR plant data exchange standard (FAIDARE plant data portal <https://urgi.versailles.inrae.fr/aidare/>)
- Phenome-EMPHASIS.fr indexing standard
- EMPHASIS platform IS : PHIS, PIPPA, ...
- Adoption in the CGIAR
 - Breedbase
 - Solegenomics
 - ...
- FAIRDOM-SEEK

MIAPPE SPECIFICATIONS



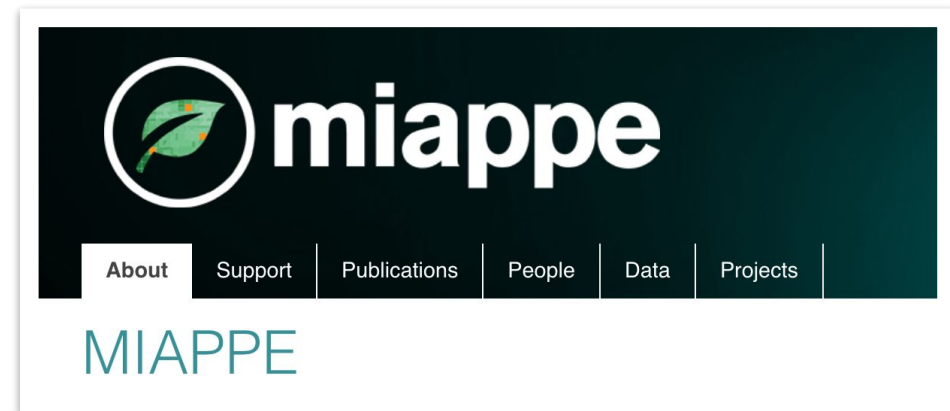
MIAPPE Specifications

- www.miappe.org

MIAPPE primers

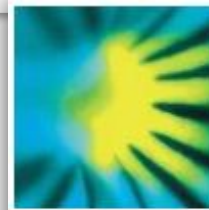
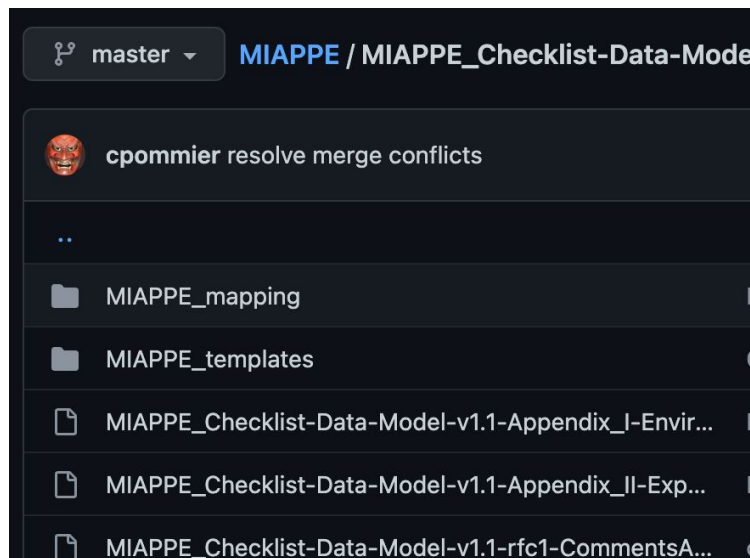
See the [support page](#) for full informations

- The latest specifications, [data model overview](#)
- The latest specifications [field list with description](#)



- Github

https://github.com/MIAPPE/MIAPPE/tree/master/MIAPPE_Checklist-Data-Model-v1.1



New Phytologist

Methods | [Open Access](#) | [CC](#) | [i](#)

Enabling reusability of plant phenomic datasets with MIAPPE 1.1

Evangelia A. Papoutsoglou [✉](#), Daniel Faria, Daniel Arend, Elizabeth Arnaud, Ioannis N. Athanasiadis, Inês Chaves, Frederik Coppens, Guillaume Cornut, Bruno V. Costa, Hanna Ćwiek-Kupczyńska, Bert Driesbeke, Richard Finkers, Kristina Gruden, Astrid Junker, Graham J. King, Paweł Krajewski, Matthias Lange, Marie-Angélique Laporte, Célia Michotey, Markus Oppermann, Richard Ostler, Hendrik Poorter, Ricardo Ramírez-Gonzalez, Živa Ramšak, Jochen C. Reif, Philippe Rocca-Serra, Susanna-Assunta Sansone, Uwe Scholz, François Tardieu, Cristobal Uauy, Björn Usadel, Richard G. F. Visser, Stephan Weise, Paul J. Kersey, Célia M. Miguel, Anne-Françoise Adam-Blondon, Cyril Pommier [✉](#) ... [See fewer authors](#) ^

First published: 14 March 2020 | <https://doi.org/10.1111/nph.16544> | Citations: 10

MIAPPE Specifications

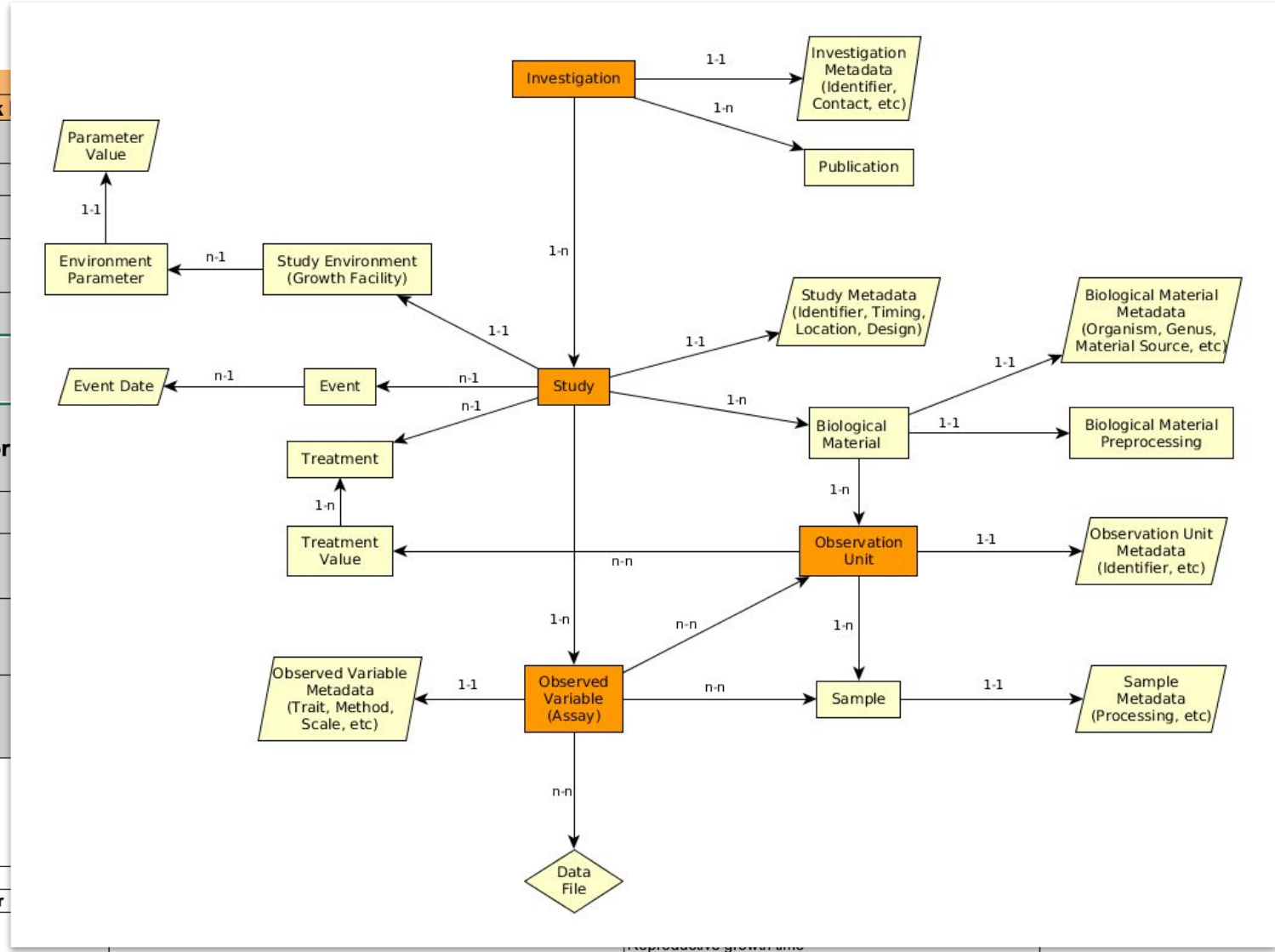
- Specifications table
- 11 Sections
- Metadata Fields
- Definitions
- Examples
- Cardinalities

line #	MIAPPE				Cardinality
	MIAPPE Check list	Definition	Example	Format	
DM-1	Investigation	Investigations are research programmes with defined aims. They can exist at various scales (for example, they could encompass a grant-funded programme of work, the various components comprising a peer-reviewed publication, or a single experiment).			1 per MIAPPE submission
DM-10	Study	A study (or experiment) comprises a series of assays (or measurements) of one or more types, undertaken to answer a particular biological question.			1+ per investigation
DM-30	Person	A human involved in the investigation or specifically any of its studies.			1+ per investigation / 0+
DM-36	Data File	A file or digital object holding observation data recorded during one or more assays of the study, typically in tabular form. Multiple data files may be provided per study, and each file can include observations for several observation units and several observed variables.			0+ per study
DM-40	Biological Material	The biological material being studied (e.g. plants grown from a certain bag or seed, or plants grown in a particular field). The original source of that material (e.g., the seeds or the original plant cloned) is called the material source, which, when held by a material repository, should have its stock identified.			1+ per study; 0+ per observation unit
DM-57	Environment	Environmental parameters that were kept constant throughout the study and did not change between observation units or assays. Environment characteristics that vary over time, i.e. environmental variables, should be recorded as Observed Variables (see below).			0-1 per study
DM-60	Experimental Factor	The object of a study is to ascertain the impact of one or more factors on the biological material. Thus, a factor is, by definition a condition that varies between observation units, which may be biotic (pest, disease interaction) or abiotic (treatment and cultural practice) in nature. Depending on the level of the data, an experimental factor can be either "what is the factor applied to the plant" (i.e. Unwatered), or the "environmental characterisation" (i.e. if no rain on unwatered plant : Drought ; if rain on unwatered plant: Irrigated)			0+ per study; 0+ per observation unit
DM-64	Event	An event is discrete occurrence at a particular time in the experiment (which can be natural, such as rain, or unnatural, such as planting, watering, etc). Events may be the realization of Factors or parts of Factors, or may be confounding to Factors. Can be applied at the whole study level or to only a subset of observation units.			0+ per study/observation unit
DM-69	Observation Unit	Observation units are objects that are subject to instances of observation and measurement. An observation unit comprises one or more plants, and/or their environment. There can be pure environment observation units with no plants. Synonym: Experimental unit.			1+ per study
DM-75	Sample	A sample is a portion of plant tissue harvested, non-harvested or extracted from an observation unit for the purpose of sub-plant observations and/or molecular studies. A sample must be used when there is a physical sample that needs to be stored and traced. Otherwise, observations made at the sub-plant level should be recorded as plant level observations using the observed variables to characterize the object of the observation (e.g. Berry sugar content, Fruit weight, Grain Protein content, Leaf 1 width, Leaf 2 width, Leaf 2 length).			0+ per observation unit
DM-82	Observed Variable	An observed variable describes how a measurement has been made. It typically takes the form of a measured characteristic of the observation unit (plant or environmental trait), associated to the method and unit of measurement. Multiple variables with the same combination of trait, method and scale can be used in association with different plant parts (leaf 1, leaf 2), when this distinction is necessary for observations referring to different parts of the same observation unit.			1+ per study
DM-83	Variable ID	Code used to identify the variable in the data file. We recommend using a variable definition from the Crop Ontology where possible. Otherwise, the Crop Ontology naming convention is recommended: <trait abbreviation>_<method abbreviation>_<scale abbreviation>. A variable ID must be unique within a given investigation.	Ant_Cmp_Cday	Unique identifier	1
DM-84	Variable name	Name of the variable.	Anthesis computed in growing degree days	Free text	0-1
DM-85	Variable accession number	Accession number of the variable in the Crop Ontology	CO_322:0000794	Crop Ontology term	0-1
DM-86	Trait	Name of the (plant or environmental) trait under observation	Anthesis time Reproductive growth time	Free text	1
DM-87	Trait accession number	Accession number of the trait in a suitable controlled vocabulary (Crop Ontology, Trait Ontology).	CO_322:0000030 TO:0000366	Term from Plant Trait Ontology, Crop Ontology, or XML Environment Ontology	0-1

MIAPPE Specifications

- Specifications table
- 11 Sections
- Metadata Fields
- Definitions
- Examples
- Cardinalities
- Links between sections/concepts

line #	MIAPPE Check
DM-1	Investigation
DM-10	Study
DM-30	Person
DM-36	Data File
DM-40	Biological Material
DM-57	Environment
DM-60	Experimental Factor
DM-64	Event
DM-69	Observation Unit
DM-75	Sample
DM-82	Observed Variable
DM-83	Variable ID
DM-84	Variable name
DM-85	Variable accession number
DM-86	Trait

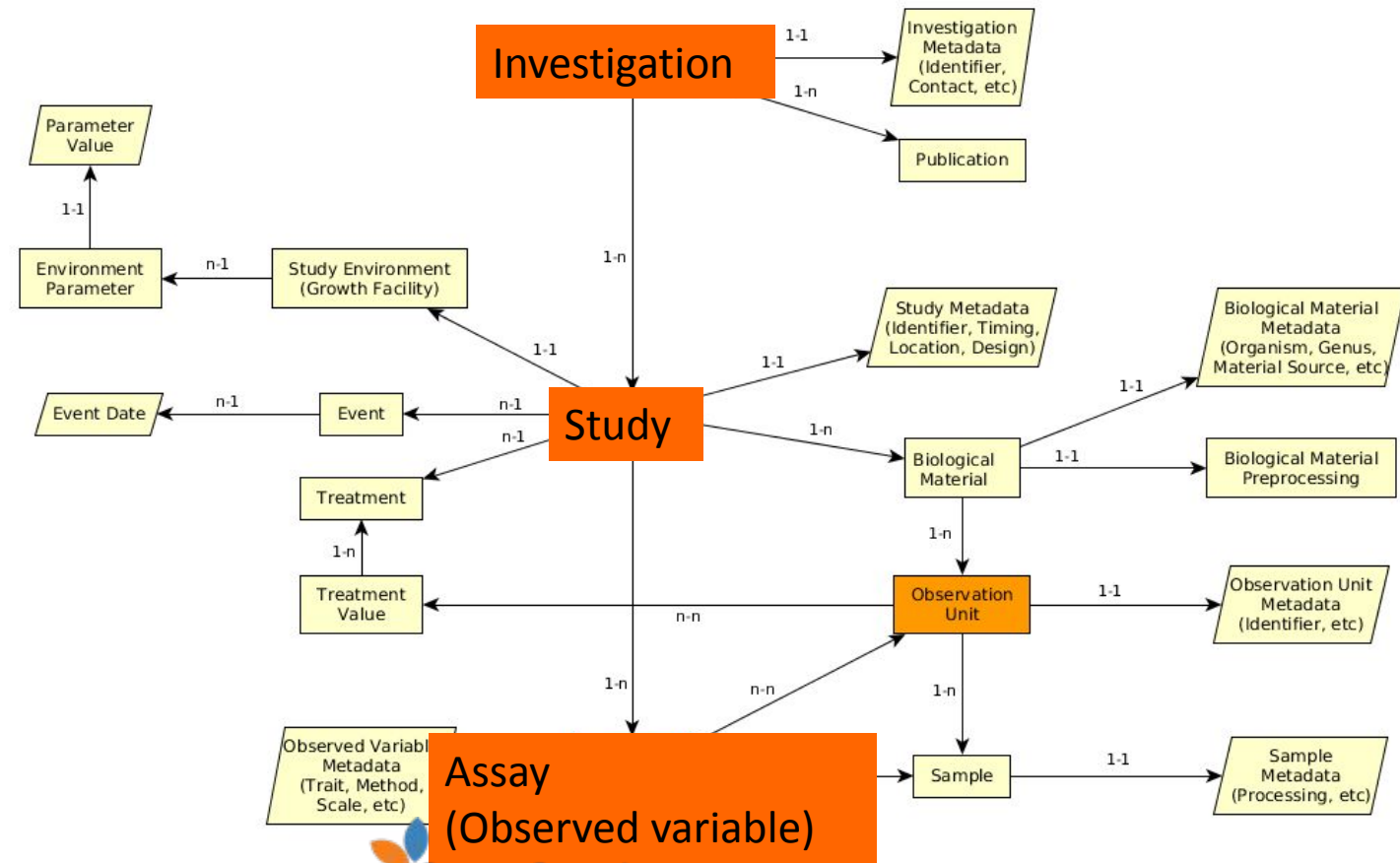
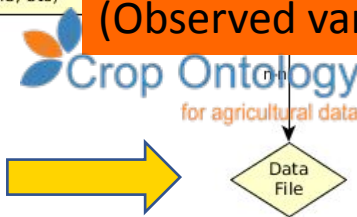
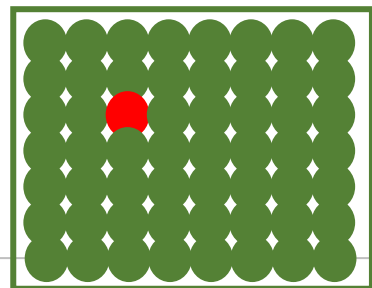


Cardinality
1 per MIAPPE submission
1+ per investigation
1+ per investigation / 0+
0+ per study
1+ per study; 0+ per observation unit
0-1 per study
0+ per study; 0+ per observation unit
0+ per study/observation unit
1+ per study
0+ per observation unit
1+ per study
1
0-1
0-1
1

DM-87	Trait accession number	Accession number of the trait in a suitable controlled vocabulary (Crop Ontology, Trait Ontology).	CO_322:0000030 TO:0000366	Term from Plant Trait Ontology, Crop Ontology, or XML Environment Ontology	0-1
-------	------------------------	--	------------------------------	--	-----

MIAPPE V1.1 overview (ISA)

- **Investigation:** the complete dataset
 - Fields, greenhouse, ...
 - One or more locations
 - One to several years
- **Study:** one experiment in one location for one or more years
- **Assay:**
 - Variable or Indice : Trait Pheno or Env
 - Level + Trait + Method + Scale/Unit
- **Level:**
 - Plant
 - Microplot
 - Block
 - Trial/Study
 - ...



Examples in MIAPPE

Two illustrations showing the experimental phenotyping documentation:

- **MAIZE:** [1] Millet *et al.* 2019 (<https://doi.org/10.15454/IASSTN>): A multi-site experiment in a network of European fields for assessing the maize yield response to environmental scenarios.
- **POPLAR:** [2] Monclus *et al.* 2012 (<http://dx.doi.org/10.1186/1471-2229-12-173>): Integrating genome annotation and QTL position to identify candidate genes for productivity, architecture and water-use efficiency in *Populus* spp

Millet et al 2019 [1] - Material & methods section on phenotyping experiments

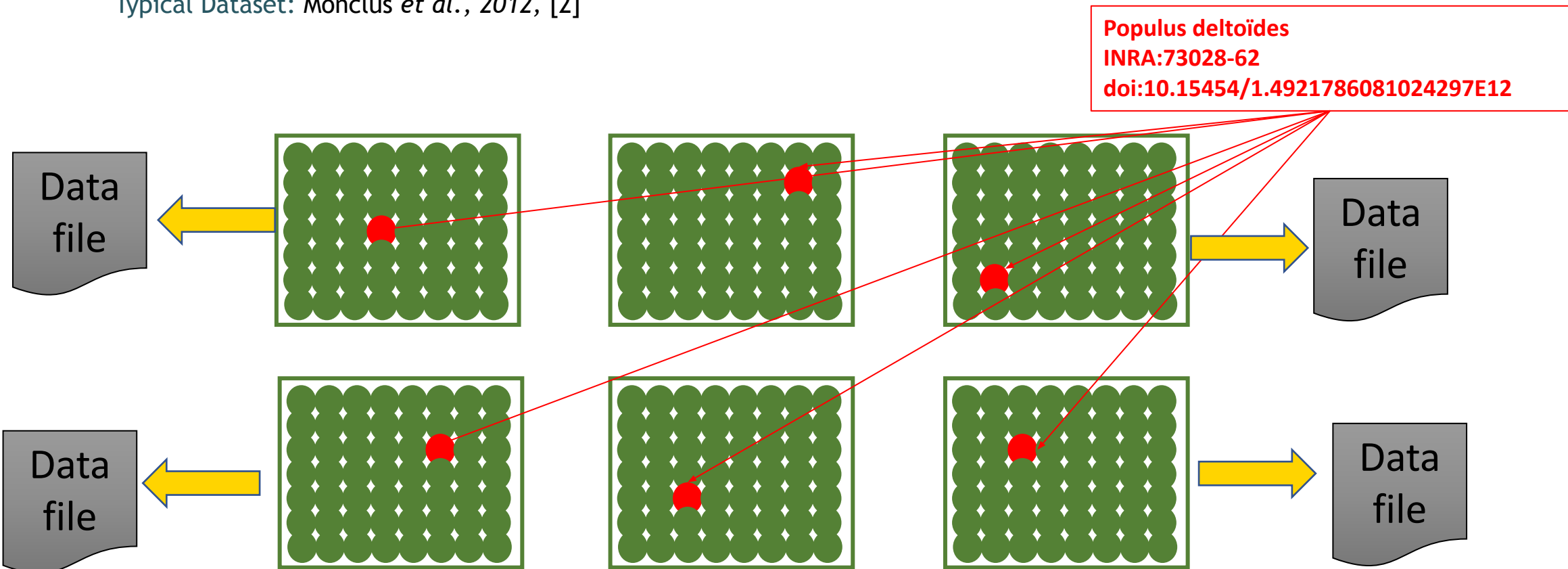
- A panel of **256 maize hybrids** was grown with **two water regimes** (irrigated or rainfed).
- **Location**: seven fields in 2012 and 2013, plus one site in Chile in 2013
- This resulted in 29 experiments defined as the combination of one year, one site and one **water regime**, with two and three **repetitions** for rainfed and irrigated treatments, respectively.
- A detailed **environmental characterisation** was carried out, with hourly records of micrometeorological data and soil water status, and associated with precise measurement of phenology.
- « **grain.yield** »: yield adjusted at 15% grain moisture, in ton per hectare (t ha⁻¹). « **grain.number** »: number of grain per square meter. « **grain.weight** »: individual grain weight (mg). « **anthesis** »: male flowering (pollen shed), in thermal time cumulated since emergence (d20°C). « **silking** »: female flowering (silking emergence), in thermal time cumulated since emergence (d20°C). « **plant.height** »: plant height, from ground level to the base of the flag leaf (highest) leaf (cm). « **tassel.height** »: plant height including tassel, from ground level to the highest point of the tassel (cm). « **ear.height** »: ear insertion height, from ground level to ligule of the highest ear leaf (cm).

Monclus et al 2012 [2] - Material & methods section on phenotyping experiments

- **3 Field trial** established in April 2003 **located** in France (Ardon, 47°49'41"N, 1°54'39"E, 110 m), Italy Cavallermaggiore ((44° 43' 0" N) , (7° 41' 0" E)), UK Headley ((51° 7' 0" N) , (-1° 10' 0" W))
- The **biological material** consisted of a cloned 336 F1 progeny from an interspecific cross between the female *Populus deltoides* (Bartr. Ex Marsh.) '73028-62' from Illinois and the male *P. trichocarpa* (Torr. and Gray) '101-74' from Washington State.
- The **trial was established** from 25 cm- homogenous hardwood cuttings planted at a plant density of 6670 trees per ha. The trial was and consisted in 6 randomized complete blocks where each F1 genotype and each parent was represented by one replicate.
- **Circumference and stem height** were measured at the end of the first (winter 2003–2004) and second (winter 2004–2005) as described in Dillen et al. Forest Ecol Manag. 2007, 252 (1–3): 12-23). Growth increment in height and circumference during the second growing season were calculated.
- **Leaf traits** were measured in 2003: one fully illuminated mature leaf was collected on each tree according to Monclus et al. <http://doi.org/10.1111/j.1469-8137.2005.01407.x>). Six calibrated discs of lamina were cut from this leaf, dried at 50 °C during 48 h and weighed, and specific leaf area (SLA, cm² g⁻¹) was computed. Leaf discs were ground to fine powder for analysis of leaf carbon isotope composition ($\delta^{13}\text{C}$), carbon (C_M) and nitrogen (N_M) contents. One-milligram subsamples of ground material were used for measuring the CO₂ produced by combustion and its ¹³CO₂/¹²CO₂ ratio by a continuous flux isotope ratio mass spectrometer. The discrimination between atmospheric CO₂ and plant material was calculated.

Experimental system in MIAPPE (crops & perennial plants)

Typical Dataset: Monclus *et al.*, 2012, [2]



6 randomized blocks

1 observation unit = one tree

No treatment

6 replicates defined by their position in each block: row and column

Experimental system in MIAPPE (crops & perennial plants)

- Free format (Near Infra Red Spectrum, tabular format, Images, Image Archives references,)
- Please note that a recommended format is currently being validated, see [github issue 71](#)
- Mainly tabular
- Metadata on each column heading
- Example : Poplar dataset

A	B	C	D	E	F	G	H	I
Accession Number	Trial Site	Campaign	Circum1: Tree circumference at 1 year	Date [Circum1]	Height1: Tree total height at 1 year	Date [Height1]	Shoots3: Number of resprouts at 3 years	Date [Shoots3]
661300270	Ardon	2004	45.645632645603683	12/01/2004	284.3	12/01/2004		
661300270	Ardon	2005					14.630625	12/05/2005
661300444	Ardon	2004	38.96112577281653	12/01/2004	228.8	12/01/2004		
661300444	Ardon	2005					8.5030559999999991	12/05/2005
661300312	Cavallermaggiore	2004	52.4	01/01/2004	249.9	01/01/2004		
661300312	Cavallermaggiore	2005					12.9816090000000001	01/05/2005
661300371	Cavallermaggiore	2004	45.74	01/01/2004	230.2	01/01/2004		
661300371	Cavallermaggiore	2005					10.3041	01/05/2005
661300487	Cavallermaggiore	2004	72.52	01/01/2004	309.8	01/01/2004		
661300487	Cavallermaggiore	2005					10.6798239999999998	01/05/2005
661300585	Cavallermaggiore	2004	71.739999999999995	01/01/2004	305.7	01/01/2004		
661300585	Cavallermaggiore	2005					10.9561000000000001	01/05/2005
661300468	Headley	2004	45.27	01/01/2004	247	01/01/2004		
661300468	Headley	2005					15.8881960000000002	01/05/2005
661300469	Headley	2004	70.9300000000000007	01/01/2004	313	01/01/2004		
661300469	Headley	2005					13.2714489999999999	01/05/2005
661300533	Headley	2004	57.67	01/01/2004	258.8	01/01/2004		

MIAPPE Mandatory sections - Investigation

Investigations = Dataset

Or Research programmes with defined objectives

Examples:

- Unique experiment
- Multilocal network
- Funded project

=> An investigation contains one or more studies.

Metadata: Information on the publication of papers or data sets

=> Title, description, keywords, auteurs, DOI, ...

Examples:

MAIZE [1]: Full set of experiments, all locations and yearq

POPLAR [2]: All multi local measurements over three years

Investigation
Investigation unique ID
Investigation title
Investigation description
Submission date
Public release date
License
MIAPPE version
Associated publication

MIAPPE Mandatory sections - Study

"Study" (or Experiment)

- Includes assays (ie analyses or measurements)
- One study \Rightarrow one location

Metadata: description of the whole experiment

=> Timing, location, statistical design, cultural practices, etc.

Examples:

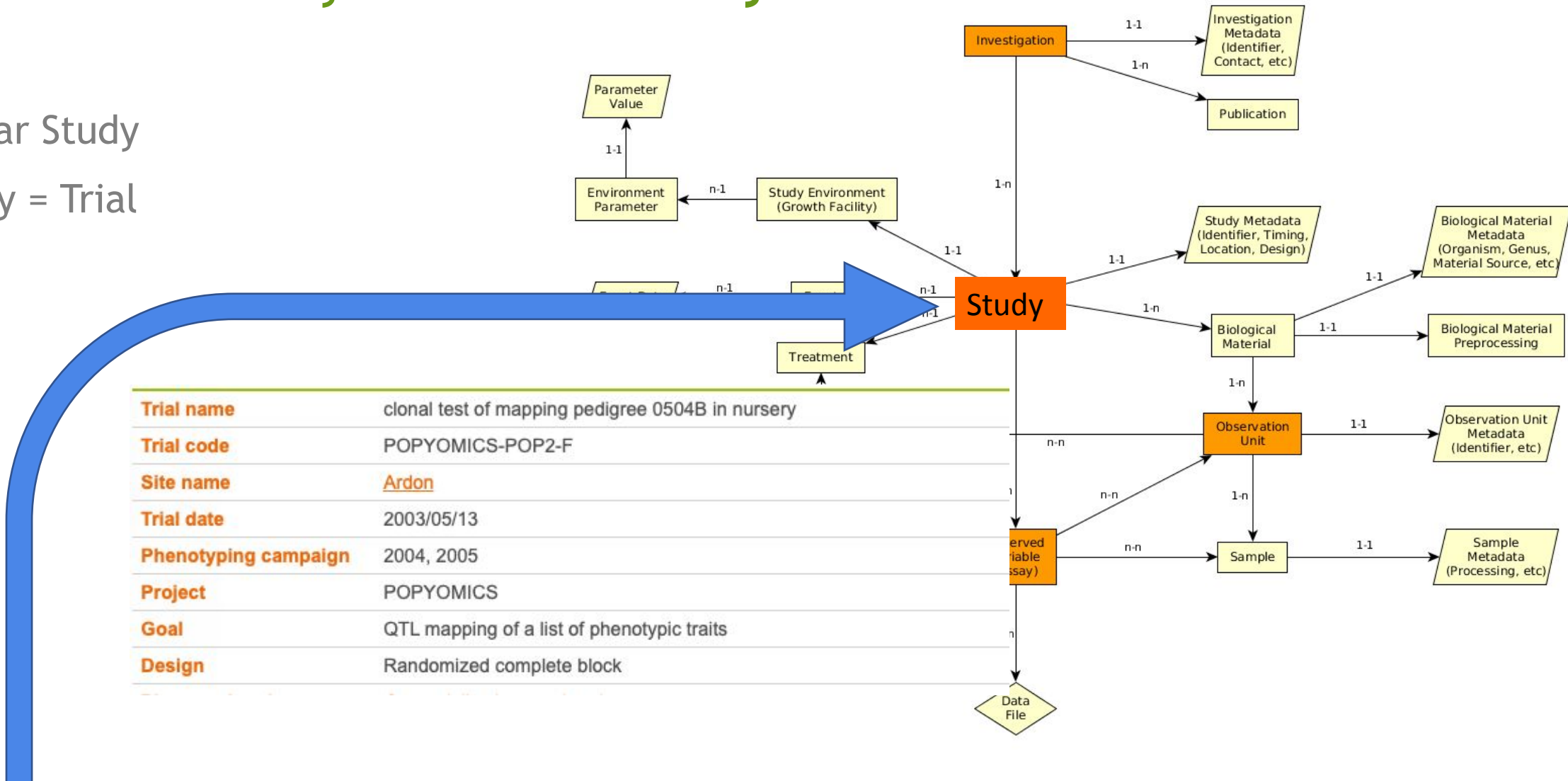
MAIZE [1] : 29 studies => 2 years x 7 locations x 2 treatments (Gaillac 2012 rain, Gaillac 2013 watered, ...)

POPLAR [2] : 6 studies => 3 locations over 2 years (Ardon_2003-2004, Ardon_2004-2005)

Study
Study unique ID
Study title
Study description
Start date of study
End date of study
Contact institution
Geographic location (country)
Experimental site name
Geographic location (latitude)
Geographic location (longitude)
Geographic location (altitude)
Description of the experimental design
Type of experimental design
Observation unit level hierarchy
Observation unit description
Description of growth facility
Type of growth facility
Cultural practices
Map of experimental design

MIAPPE Mandatory sections - Study Metadata

- Poplar Study
- Study = Trial



Trial name	clonal test of mapping pedigree 0504B in nursery
Trial code	POPYOMICS-POP2-F
Site name	<u>Ardon</u>
Trial date	2003/05/13
Phenotyping campaign	2004, 2005
Project	POPYOMICS
Goal	QTL mapping of a list of phenotypic traits
Design	Randomized complete block

A	B	C	D	E	F	G	H	I
Accession Number	Trial Site	Campaign	Circum1: Tree circumference at 1 year	Date [Circum1]	Height1: Tree total height at 1 year	Date [Height1]	Shoots3: Number of resprouts at 3 years	Date [Shoots3]
661300270	Ardon	2004	45.645632645603683	12/01/2004	284.3	12/01/2004		
661300270	Ardon	2005					14.630625	12/05/2005
661300444	Ardon	2004	38.96112577281653	12/01/2004	228.8	12/01/2004		
661300444	Ardon	2005					8.5030559999999991	12/05/2005

Biological Material
Biological material ID
Organism
Genus
Species
Infraspecific name
Biological material latitude
Biological material longitude
Biological material altitude
Biological material coordinates uncertainty
Biological material preprocessing
Material source ID (Holding institute/stock centre, accession)
Material source DOI
Material source latitude
Material source longitude
Material source altitude
Material source coordinates uncertainty
Material source description

MIAPPE Mandatory sections - Biological material

List of biological material, accessions, varieties, ...

- Identifier *and* source (storage centre, gene bank, etc.).
- Crucial for integrating pheno-genomic or genetic data

Metadata

- Minimum fields of the standard Multicultural Passport Descriptor (MCPD)
- PLUS GPS location of forest tree / provenance of in situ material

Material source: accession, cultivar/variety, region of provenance, laboratory cross, ...



Biological material: seed lot, cuttings, plant in the study, ...



Plant Samples: part of the plant taken for study, detached leaves, ...

MCPD identification system:

- Genebank/Lab + Species + accession number (mandatory)
- DOI

- Lab + internal accession number (mandatory)
- URI

- Lab + internal accession number (mandatory)
- BioSample ID

MIAPPE important section - Observation Unit & Samples

Observation units :

- Physical objects (e.g. microplot) of the study on which measurements and observations are made.
- Virtual objects (genotype) on which variables or indices are calculated.
- Typically one or more plants
- Object of the phenological or environmental measurements

Specific Metadata

=> identifiers, location, replication, treatments, ...

Sample

- Piece of tissue extracted from an observation unit
- Sub-plant observations or molecular study.

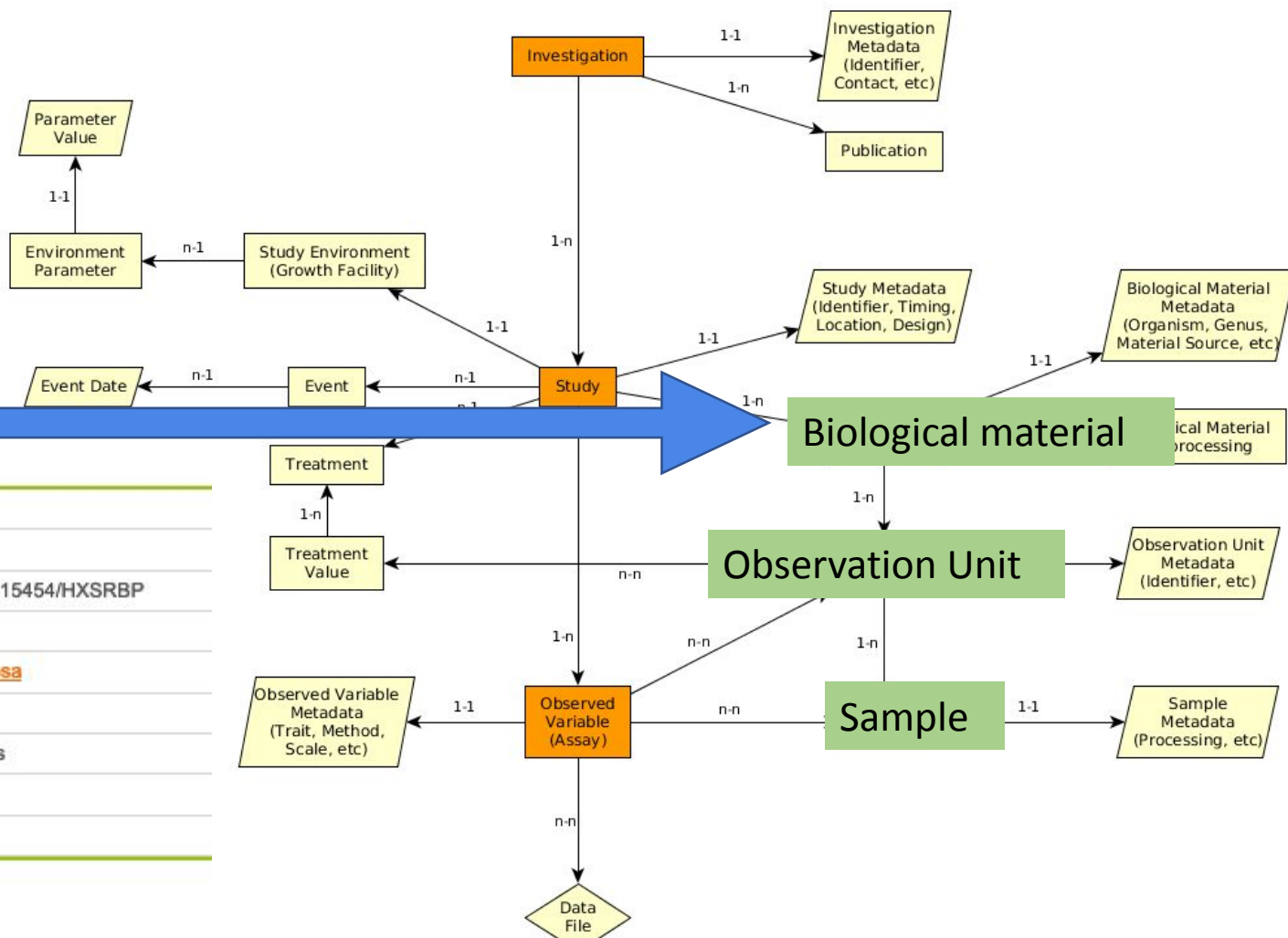
Metadata: aligned with BioSample checklists

=> identifiers, information about processing, ...

Observation Unit	
Observation unit ID	
Observation unit type	
External ID	Sample
Spatial distribution	Sample ID
Observation Unit fact	Plant structure development stage
	Plant anatomical entity
	Sample description
	Collection date
	External ID

MIAPPE Mandatory sections - Biological material

- Plant Material
 - Identification
 - Description
- Multi Crop Passport Descriptor



IDENTIFICATION	
Accession number	661300333
Accession name	661300333
Permanent Unique Identifier	https://doi.org/10.15454/HXSRBP
Synonyms	0054B115
species	<i>Populus x generosa</i>
Pedigree	-
Biological status	Interspecific cross
Genetic nature	Clone
Comment	-

HOLDING	
Holding stock center	Forest BRC (I)
Presence status	

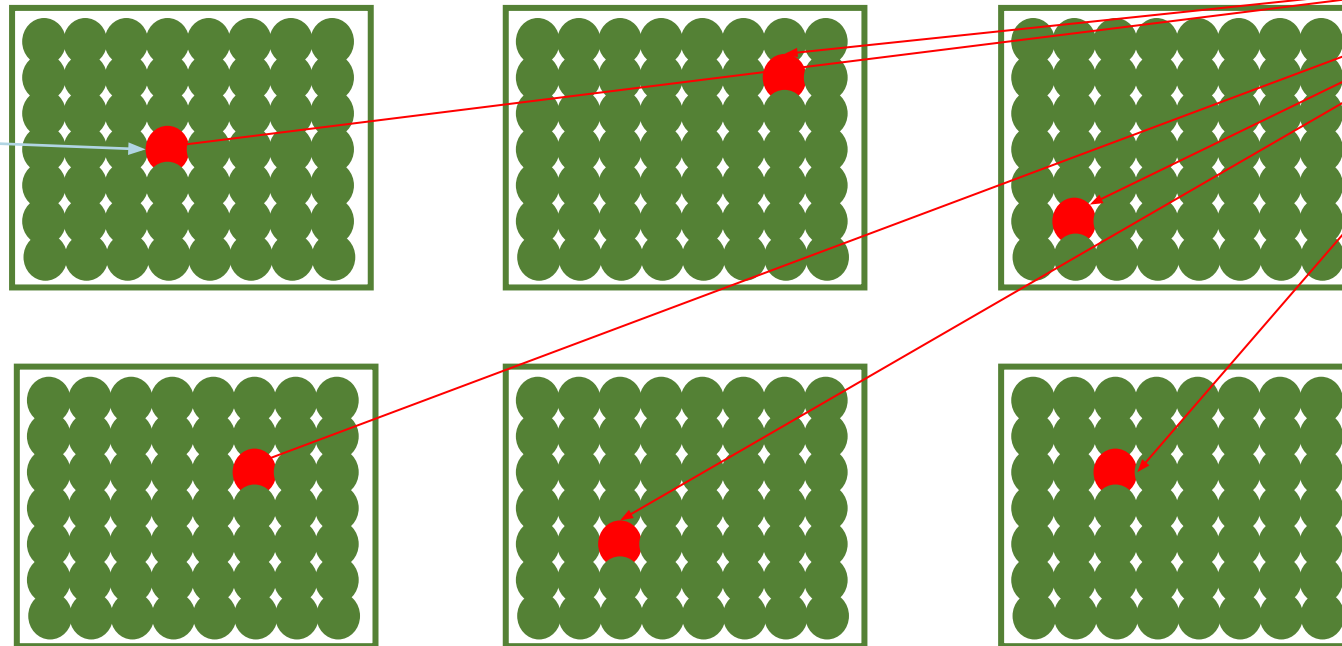
A	B	C	D	E	F	G	H	I
Accession Number	Trial Site	Campaign	Collection Date	Height [m]	Tree total height at 1 year	Date [Height1]	Shoots3: Number of resprouts at 3 years	Date [Shoots3]
661300270	Ardon	2004	4			12/01/2004		
661300270	Ardon	2005					14.630625	12/05/2005
661300444	Ardon	2004	38.96112577281653		12/01/2004	228.8		
661300444	Ardon	2005					8.5030559999999991	12/05/2005

MIAPPE main sections - Biological Material & Observation Unit

Example: POPLAR [2]

[observation unit]
Populus deltoïdes
INRA:73028-62
Block1-Row4-Col4

[Biological material]
Populus deltoïdes
INRA:73028-62
doi:10.15454/1.4921786081024297E12

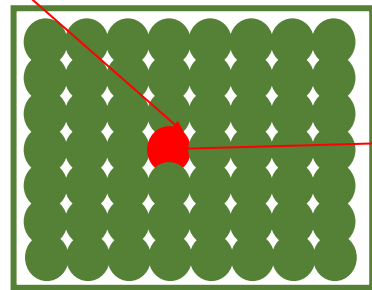


- 6 randomized blocks
- 1 observation unit = one tree
- No treatment
- 6 replicates defined by their position in each block: row and column

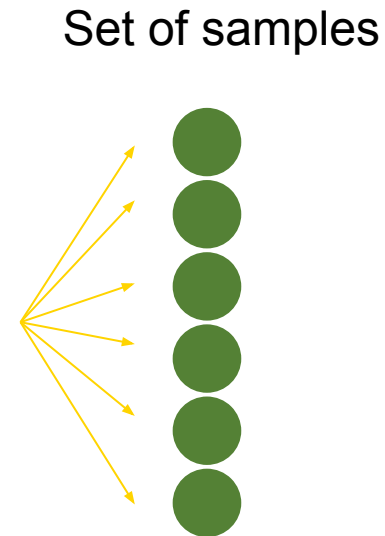
MIAPPE main sections - Observation Unit & Samples

Example: Monclus *et al.*, 2012, [2]

Populus deltoïdes
INRA:73028-62
doi:10.15454/1.4921786081024297E12
[observation unit] : Block1-Row4-Col4



One leaf



6 leaf disks

For each tree of each block
= for each observation unit

Different types of processing of the leaf disks depending on the measurement: can be captured by different sets of samples (e.g. if subsampling generates repetitions) or in the method of the observed variable.

MIAPPE Mandatory sections - Observed Variable (Assay)

Observed Variable (assay)

- Trait
- Measurement / computation method
- Unité, Scale
- Trait: Phenotype or Environment

Metadata

- Minimal: Trait, Method, Scale
- Complete: Crop Ontology Trait Dictionary
 - See following slides

Observed Variable
Variable ID
Variable name
Variable accession number
Trait
Trait accession number
Method
Method accession number
Method description
Reference associated to the method
Scale
Scale accession number
Time scale

Crop Ontology Trait Dictionary Standard

Variable = trait + method + scale

Study	Genotype	CIR	HT	SUR
Domaine de Valcros	6579	600	17	1
Domaine de Valcros	6580	482	14	1

The phenotypic observation

Variable	Variable ID
	Variable name
	Variable synonyms
	Context of use
	Growth stage
	Variable status
	Variable Xref
	Institution
	Scientist
	Date
	Language
Crop	

Crop Ontology Trait Dictionary Standard

Variable = **trait** + method + scale

Study	Genotype	CIR	HT	SUR
Domaine de Valcros	6579	600	17	1
Domaine de Valcros	6580	482	14	1

Circumference of the tree

Total height of the tree

Survival state of the tree

What is the studied character?

Trait	Trait ID
	Trait
	Trait class
	Trait description
	Trait synonyms
	Main trait abbreviation
	Alternative trait abbreviations
	Entity
	Attribute
	Trait status
Trait Xref	

Crop Ontology Trait Dictionary Standard

Variable = trait + **method** + scale

Study	Genotype	CIR	HT	SUR
Domaine de Valcros	6579	600	17	1
Domaine de Valcros	6580	482	14	1

Measured at breast height with a graduated ribbon

Measured from soil to crown with a pole or a clinometer

Visual assessment with a reference scoring scale

How is it observed?

Method	Method ID
	Method
	Method class
	Method description
	Formula
	Method reference

Crop Ontology Trait Dictionary Standard

Variable = trait + method + **scale**

Study	Genotype	CIR	HT	SUR
Domaine de Valcros	6579	600	17	1
Domaine de Valcros	6580	482	14	1

cm

mm

0 = Alive
1 = Dead
2 = Doubtful

How is it expressed (unit or scale)?

Scale	Scale ID
	Scale name
	Scale class
	Decimal places
	Lower limit
	Upper limit
	Scale Xref
	Category 1
	Category 2
	...
Category n	

MIAPPE main sections - Observed Variable (Assay)

Example MAIZE [1] :

In each study measurements are made at three levels:

- Plant level: 20 variables
- Genotype/Study level: 19 variable both phenotype and environment
- Plot level: 9 variables

=> For instance

- Female flowering days to silking D20deg
- Plant height (cm)

FFLW_D20deg: Female flowering days to silking D20deg VARIABLE

Synonyms	Female flowering days to silking D20deg FFLW
Growth stage	Flowering
Crop	Maize
Silking time TRAIT	
Identifier	CO_322:0000031
Name	Silking time
Description	Silking time
Synonyms	Female flowering time
Main abbreviation	Silk
Alternative abbreviations	S FFlw
Entity	Flower
Attribute	Silking time
Class	Phenological

Thermal time between emergence and silking – Computation METHOD

Identifier	MIPO:0000027
Name	Thermal time between emergence and silking – Computation
Description	Calculated as equivalent days at 20 °C unit between emergence and 50% anthesis.
Reference	B. Parent, O. Turc, Y. Gibon, M. Stitt and F. Tardieu (2010) Modelling temperature-compensated physiological rates, based on the co-ordination of responses to temperature of developmental processes. Journal of Experimental Botany
Class	Computation

PTHT: Plant height (cm) VARIABLE

Ontology name	Maize Traits
Identifier	MIPO:0000006
Name	PTHT
Synonyms	Plant height (cm)
Xref	CO_322:0000994
Crop	Maize

Plant height TRAIT

Identifier	CO_322:0000994
Name	Plant height
Description	Plant height from the base to the top part (in reproductive stages to the top of the tassel).
Main abbreviation	PH
Entity	Plant
Attribute	height
Class	Agronomical

PH - Measurement METHOD

Identifier	CO_322:0000995
Name	PH - Measurement
Description	Recommended to take multiple plants and measure the height from the base of a plant to the top of the tassel, enter the data individually in the FieldBook and calculate the average.
Reference	DTMA drought phenotyping protocol. 2009. CIMMYT. Magorokosho et al. 2010. Characterization of maize germplasm grown in eastern and southern Africa: Results of the 2009 regional trials coordinated by CIMMYT. Zimbabwe

D20deg: days equivalent time at 20 °C SCALE

Identifier	MIPO:0000030
Name	D20deg: days equivalent time at 20 °C
Data type	Numerical

MIAPPE Mandatory sections - Observed variable

Each column: Trait + Method + Scale

Traits, methods and scales English

Search terms...

- Terminal bud diameter TRAIT
- Tree circumference TRAIT
- Tree diameter TRAIT
- Tree flexuosity TRAIT
- Tree height TRAIT
- HT: Tree total height VARIABLE
- HTm: Tree total height in m VARIABLE
- Tree ring width TRAIT
- Tree shape TRAIT
- Tree status TRAIT
- Type c p e TRAIT
- Volume TRAIT
- Whorl defects TRAIT
- Wood axial parenchyma high TRAIT
- Wood axial parenchyma thickness TRAIT
- Wood density TRAIT

HT: Tree total height VARIABLE

Ontology name: Woody Plant Ontology
 Identifier: CO_357:0000048
 Name: HT

Tree height TRAIT

Identifier: CO_357:1000037
 Name: Tree height
 Description: Total height of the tree, from the ground to the tallest part of the crown

Main abbreviation: HT
 Alternative abbreviations: Height, HP, HPL, h, H, TH, Height.F, Height.LUK, Height.I

Entity: tree
 Attribute: height
 Status: Standard for INRA & GenTree project
 Class: Morphological

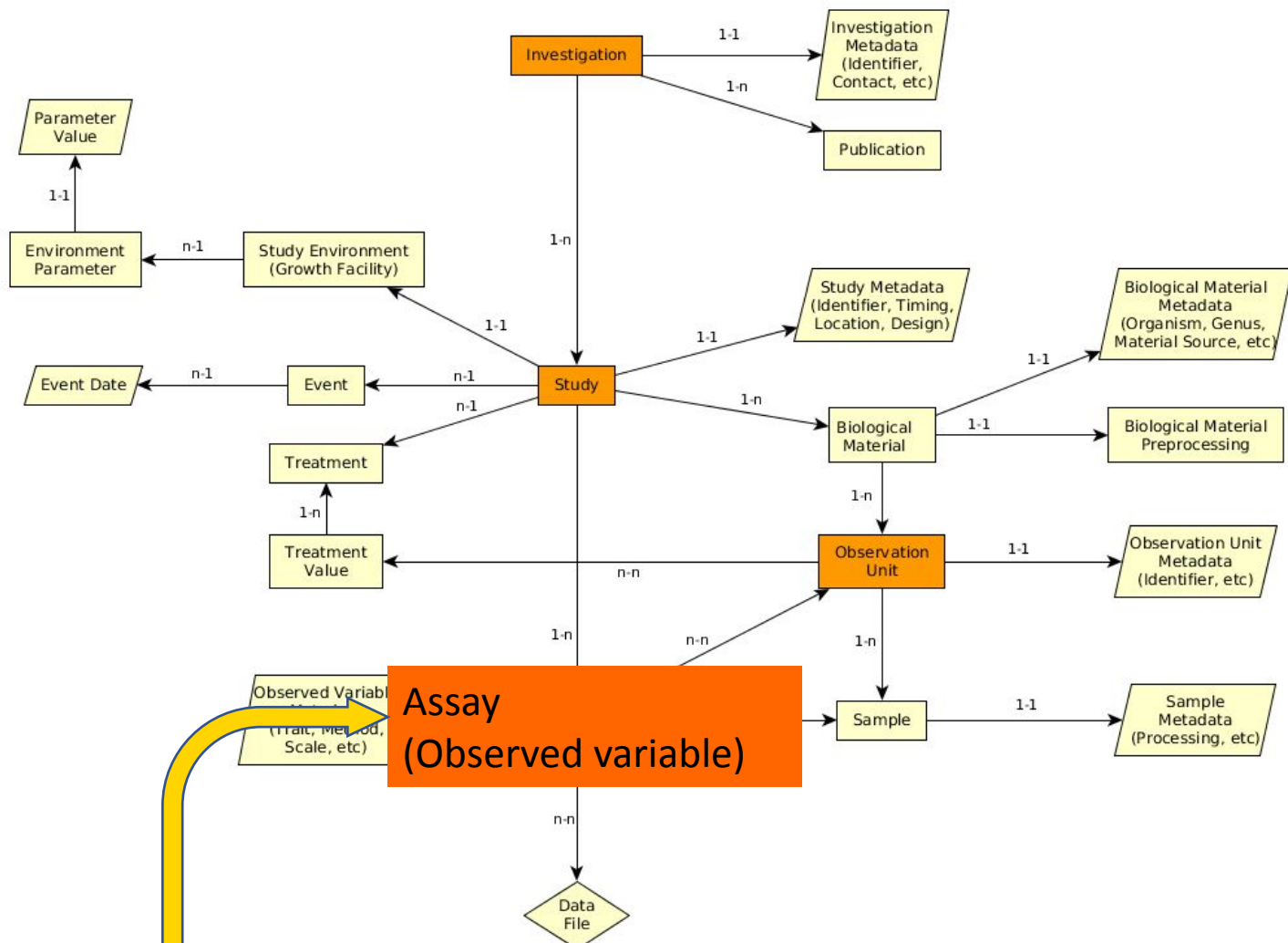
Tree height protocol METHOD

Identifier: CO_357:2000027
 Name: Tree height protocol
 Description: Measured from soil to basis of the apical meristem or bud (depending of time in season) with a pole or a clinometer
 Reference: GenTree_protocols_0.99.pdf page 16, https://en.wikipedia.org/wiki/Tree_measurement#Height

Class: Measurement

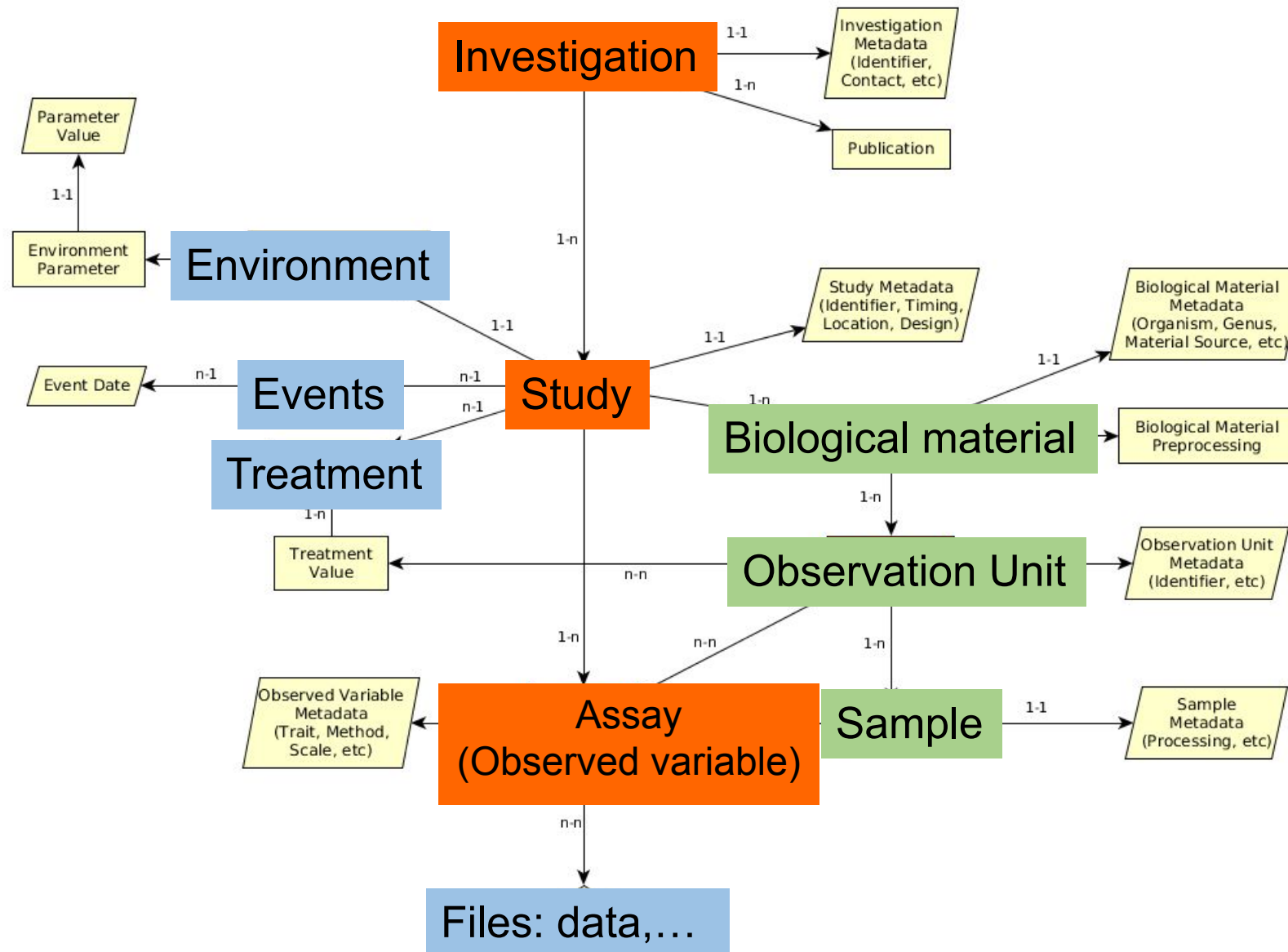
cm SCALE

Identifier: CO_357:3000107
 Name: cm



A	B	C	D					
Accession Number	Trial Site	Campaign	Circum1: Tree circumference at 1 year	Date [Circum1]	Height1: Tree total height at 1 year	Date [Height1]	Shoots3: Number of resprouts at 3 years	Date [Shoots3]
661300270	Ardon	2004	45.645632645603683	12/01/2004	284.3	12/01/2004		
661300270	Ardon	2005					14.630625	12/05/2005
661300444	Ardon	2004	38.96112577281653	12/01/2004	228.8	12/01/2004		
661300444	Ardon	2005					8.5030559999999991	12/05/2005

Other key sections



MIAPPE important section - Experimental Factor

Experimental factor:

- Treatment or biotic or abiotic factor whose effect on the study is being evaluated.
- Those are not cultural practices.
- A priori or a posteriori categorisation
- Takes on different values in the study

Experimental Factor
Experimental Factor type
Experimental Factor description
Experimental Factor values

Metadata include a name, a description and a value of the treatment

Example: MAIZE [1]

- Two experimental factors/treatments: Rainfed, Watered
- Block organisation

GENOTYPE ID		TREATMENT	Trial Name	Trial Site	LEVEL	LEVEL	LEVEL
Accession Number	Accession Name	water_regime			BLOCK	PLOT	REPLICATE
FR19_H	FR19_H	rainfed	KWS Karlsruhe 2011	Karlsruhe	5	124	1
FR19_H	FR19_H	rainfed	KWS Karlsruhe 2011	Karlsruhe	10	250	3
FR19_H	FR19_H	rainfed	KWS Karlsruhe 2013	Karlsruhe	11	401503	2
FR19_H	FR19_H	watered	KWS Karlsruhe 2012	Karlsruhe	23	400353	1
FR19_H	FR19_H	watered	KWS Karlsruhe 2011	Karlsruhe	2	6	1

MIAPPE important section - Event

Event

- Discrete, dated or time-stamped occurrence
- Natural (e.g. rain, pathogen attack)
- Cultural practice (e.g. sowing, irrigation)
- Applied to the whole study or by observation unit
- **This is not the factor**, but additional information.
- Event can be used to achieve a factor.

Event
Event type
Event accession number
Event description
Event date

Metadata : name, description and time/date

Example: POPLAR [2]

- the field establishment date, 2003
- the orchard was subjected to 15mm of rain on March 15, 2012 (fiction)



Study	Event		
	Name	Description	Date
Monclus <i>et al.</i> , 2012	Rain	15mm of rain on the orchard	2012-03-15

MIAPPE important section - Environment

Environment:

- Environmental parameters, experimental conditions
- Constant (*i.e.* a single value) over the whole study
- "Descriptors" of the study, key value pairs
- Environmental variables are treated as observed variables.

Environment
Environment parameter
Environment parameter value

Metadata: name, description and value

Example:

Environment parameter	Environment parameter Value
Mean day air temperature	22°C
Rooting medium composition	Ca (XEO:00058): 5 mg/L; ...

MIAPPE Mandatory sections - Data file

- Any format (Data matrix, NIRS, Images,)
- Mostly tabular
- Metadata on each column header

Data File
Data file link
Data file description
Data file version

A	B	C	D	E	F	G	H	I
Accession Number	Trial Site	Campaign	Circum1: Tree circumference at 1 year	Date [Circum1]	Height1: Tree total height at 1 year	Date [Height1]	Shoots3: Number of resprouts at 3 years	Date [Shoots3]
661300270	Ardon	2004	45.645632645603683	12/01/2004	284.3	12/01/2004		
661300270	Ardon	2005					14.630625	12/05/2005
661300444	Ardon	2004	38.96112577281653	12/01/2004	228.8	12/01/2004		
661300444	Ardon	2005					8.5030559999999991	12/05/2005
661300312	Cavallermaggiore	2004	52.4	01/01/2004	249.9	01/01/2004		
661300312	Cavallermaggiore	2005					12.9816090000000001	01/05/2005
661300371	Cavallermaggiore	2004	45.74	01/01/2004	230.2	01/01/2004		
661300371	Cavallermaggiore	2005					10.3041	01/05/2005
661300487	Cavallermaggiore	2004	72.52	01/01/2004	309.8	01/01/2004		
661300487	Cavallermaggiore	2005					10.6798239999999998	01/05/2005
661300585	Cavallermaggiore	2004	71.739999999999995	01/01/2004	305.7	01/01/2004		
661300585	Cavallermaggiore	2005					10.9561000000000001	01/05/2005
661300468	Headley	2004	45.27	01/01/2004	247	01/01/2004		
661300468	Headley	2005					15.8881960000000002	01/05/2005
661300469	Headley	2004	70.9300000000000007	01/01/2004	313	01/01/2004		
661300469	Headley	2005					13.2714489999999999	01/05/2005
661300533	Headley	2004	57.67	01/01/2004	258.8	01/01/2004		



Acknowledgments

Elixir Plant community & platforms

Beier S., Gruden C., Pommier C., Michotey C, Coppens Scholz U., Lange M., Contreras B., Adam Blondon AF, Faria D, Chavez I, Miguel C, Droedsbek B, Finkers R, Papoutsoglou E, Olster R, Ramsak Z, ...



H2020 AGENT



N. Stein (IPK, coord), P. Kersey (RBGK), M. Alaux (INRAE), S. Weise (IPK), C. Pommier (INRAE), M. Lange (IPK), R. Finkers (WUR), J. Destin (INRAE)

Crop Ontology

Arnaud E, Laporte MA, ...



MIAPPE community

ELIXIR Plant Community, Krajewsky P, Cwiek H, Tardieu F, Usadel B, Arend D, Arnaud E, Junker A, King G, Laporte MA, Michotey C, Poorter H, Reif J, Rocca-Serra P, Sansone SA, Kersey P, And many more!



Breeding API

Selby P, Mueller L, Robbins K, Backlund JE, ... , And many more!

Emphasis

Tardieu F, Usadel B, Arend D, Junker A, Poorter H, Neveu P, Pierushka R, Shur U... And many more!



Plant Phenomics Experiment Data

Standards

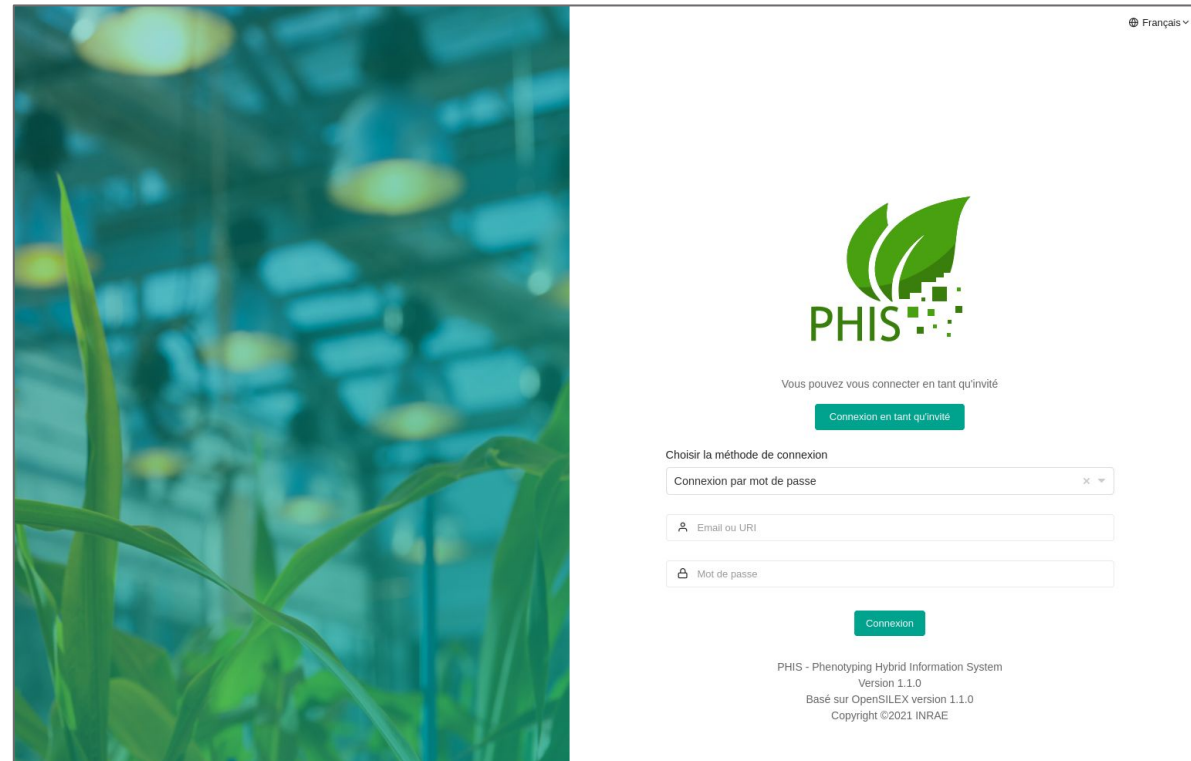


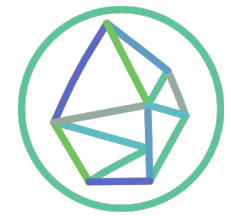
Tools



Managing experimental metadata in an Information System

Illustration with PHIS





PHIS: Ontology driven Information System for Plant Phenomics



Identification



- Every item can be identified : plant, experiment, sensor, event, etc.
- Persistent, non ambiguous, resolvable

Semantic



- Naming conventions
- Controlled vocabulary
- Links
- Annotation & Data enrichment



PHIS: Ontology driven Information System for Plant Phenomics



Identification



Semantic



URI of plant:

<<http://phenome.fr/arch/2017/c17000118>>

URI of pot:

<<http://phenome.fr/arch/2013/pc13001542>>

URI of cart:

<<http://phenome.fr/arch/2013/ct1300123>>

URI of cabin:

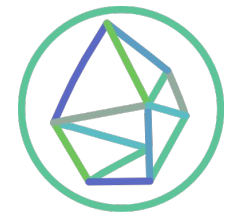
<<http://phenome.fr/arch/2018/ac180015>>

URI of camera:

<<http://phenome.fr/arch/2018/ac180019>>



URI of image: <<http://phenome.fr/arch/2017/ic17002295855>>



PHIS: Ontology driven Information System for Plant Phenomics



Identification



Semantic



URI of plant:

<<http://phenome.fr/arch/2017/c17000118>>

URI of pot:

<<http://phenome.fr/arch/2013/pc13001542>>

URI of cart:

<<http://phenome.fr/arch/2013/ct1300123>>

URI of cabin:

<<http://phenome.fr/arch/2018/ac180015>>

URI of camera:




<<http://phenome.fr/arch/2018/ac180019>>

URI of image: <<http://phenome.fr/arch/2017/ic17002295855>>



Plant_Perimeter_ImageProcessing_millimetre
Variable

Details Annotations Device associated Data Visualization Documents

Description   

Structure

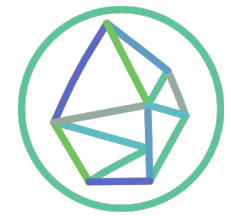
Entity	Plant
Observation level	
Characteristic	Perimeter
Method	ImageProcessing
Unit/Scale	millimetre

URI
http://phis.egi-demo.eu/id/variable/plant_perimeter_imag...

Name
Plant_Perimeter_ImageProcessing_millimetre

Alternative name
Plant_Perimeter

Description



PHIS: Ontology driven Information System for Plant Phenomics



Identification



Semantic



- URI of plant:**
<<http://phenome.fr/arch/2017/c17000118>>
- URI of pot:**
<<http://phenome.fr/arch/2013/pc13001542>>
- URI of cart:**
<<http://phenome.fr/arch/2013/ct1300123>>
- URI of cabin:**
<<http://phenome.fr/arch/2018/ac180015>>
- URI of camera:**
<<http://phenome.fr/arch/2018/ac180019>>



URI of image: <<http://phenome.fr/arch/2017/ic17002295855>>

Plant Perimeter ImageProcessing millimetre

Variables
Manage and configure variables, entities and observations

Interoperability References

Add references to http://phis.egi-demo.eu/id/variable/plant_perimeter_imageprocessing_millimetre

Reference ontologies

- AGROPORTAL
- AGROVOC
- BioPortal
- Crop Ontology
- Plant Ontology
- Planteome
- Units of measurement ontology (UO)
- Units of Measure (OM)
- QUDT Ontologies (QUDT)
- XML/XSD Datatype Schemas

Relations

Close match

Reference URI

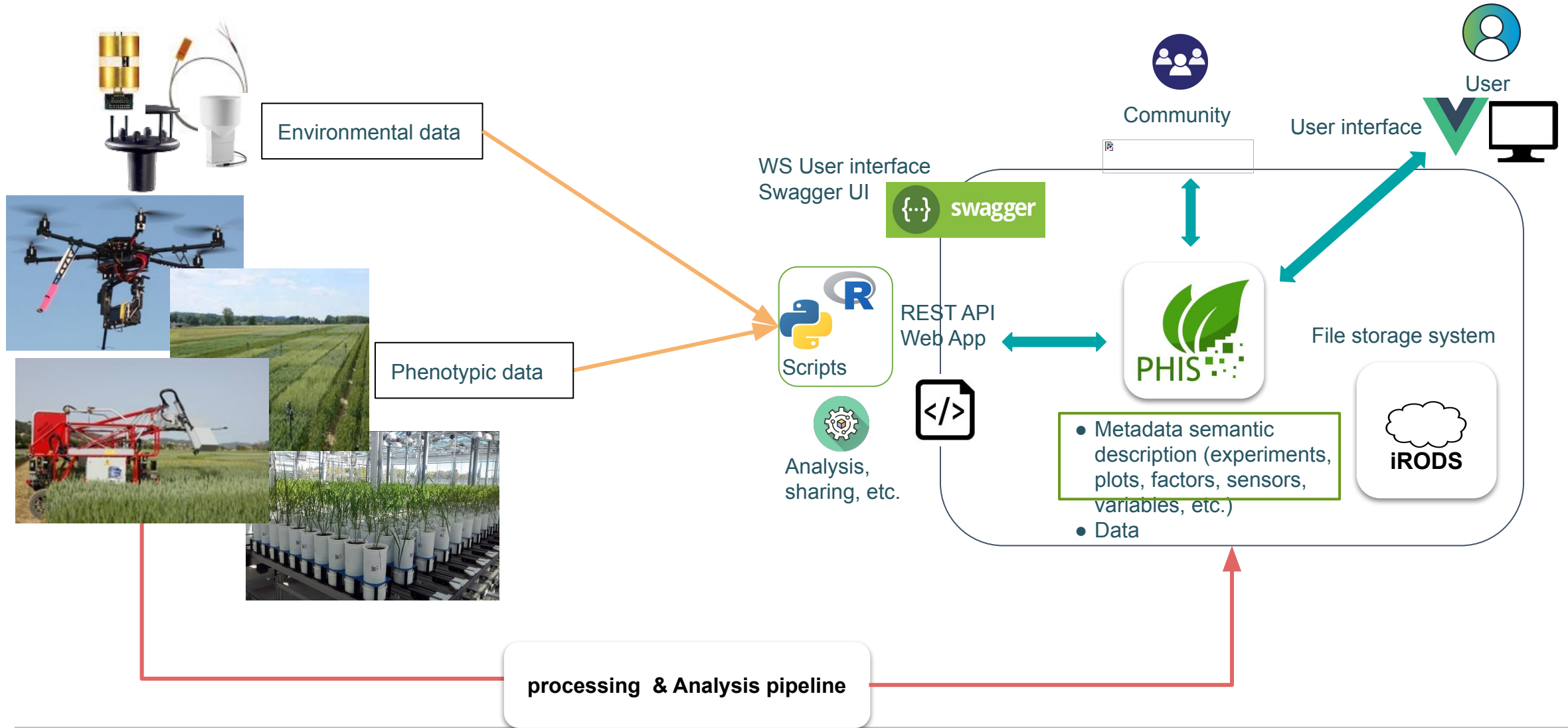
http://purl.obolibrary.org/obo/PATO_0001711

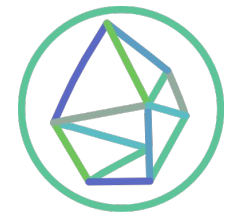
Add references

No reference available

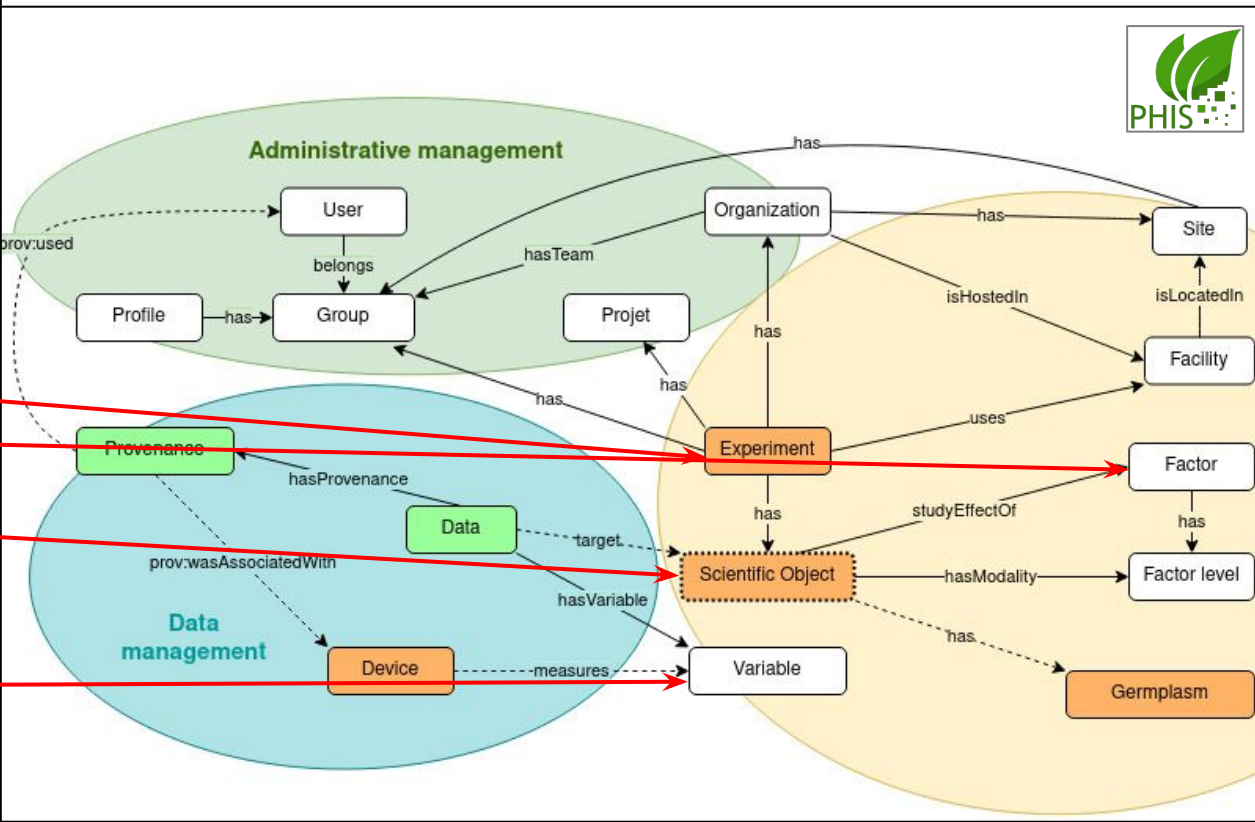
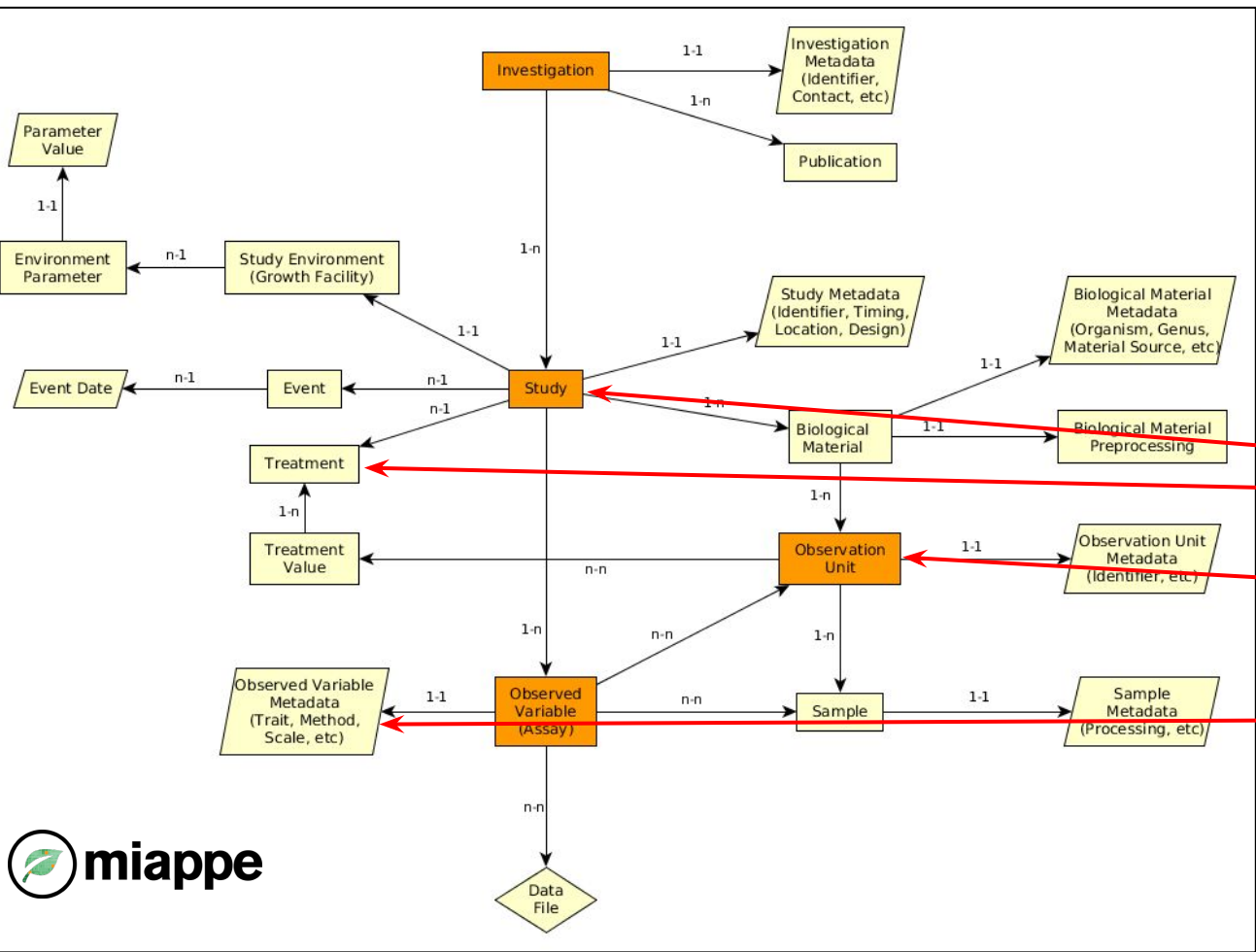


PHIS: Ontology driven Information System for Plant Phenomics





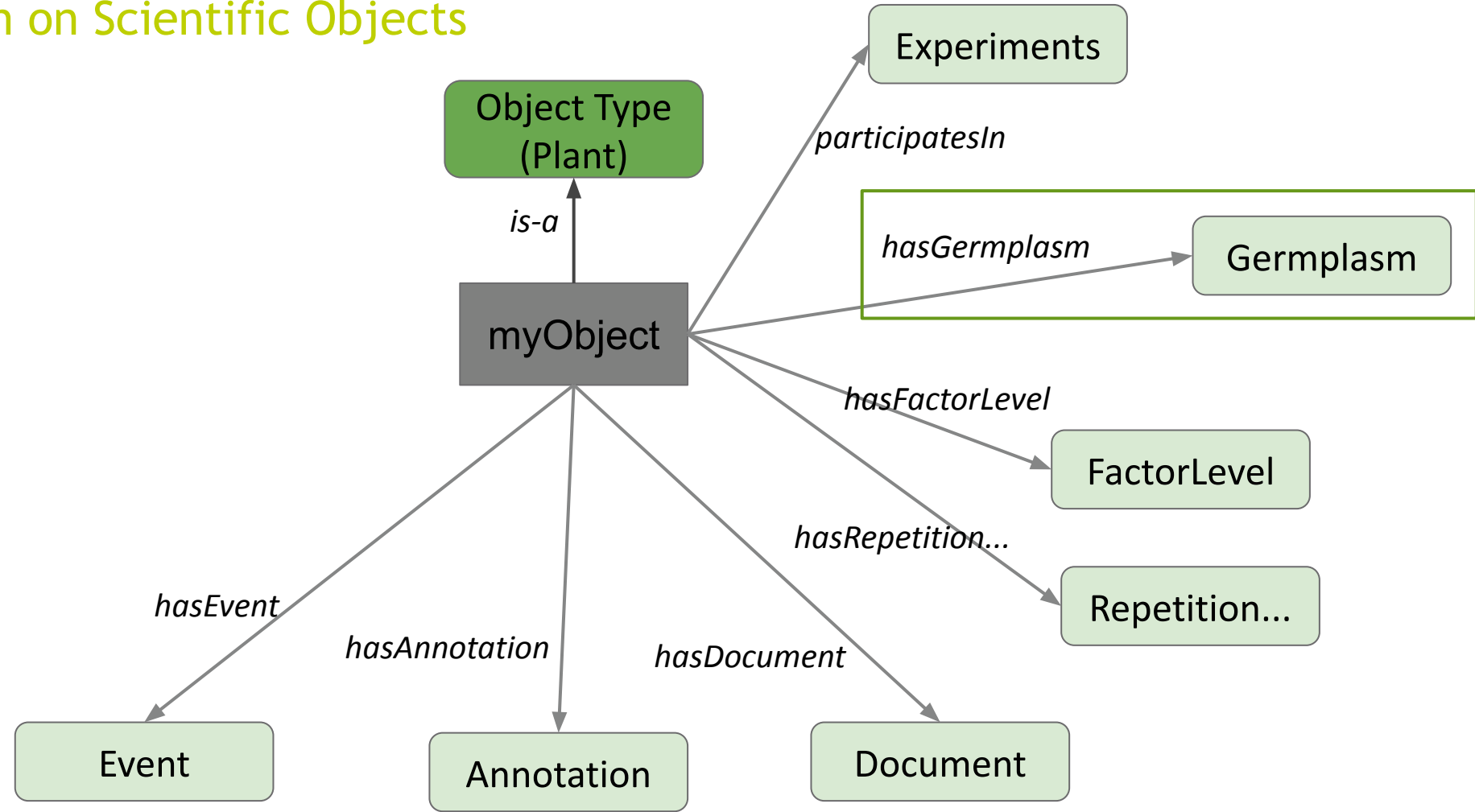
PHIS & Standards (MIAPPE)





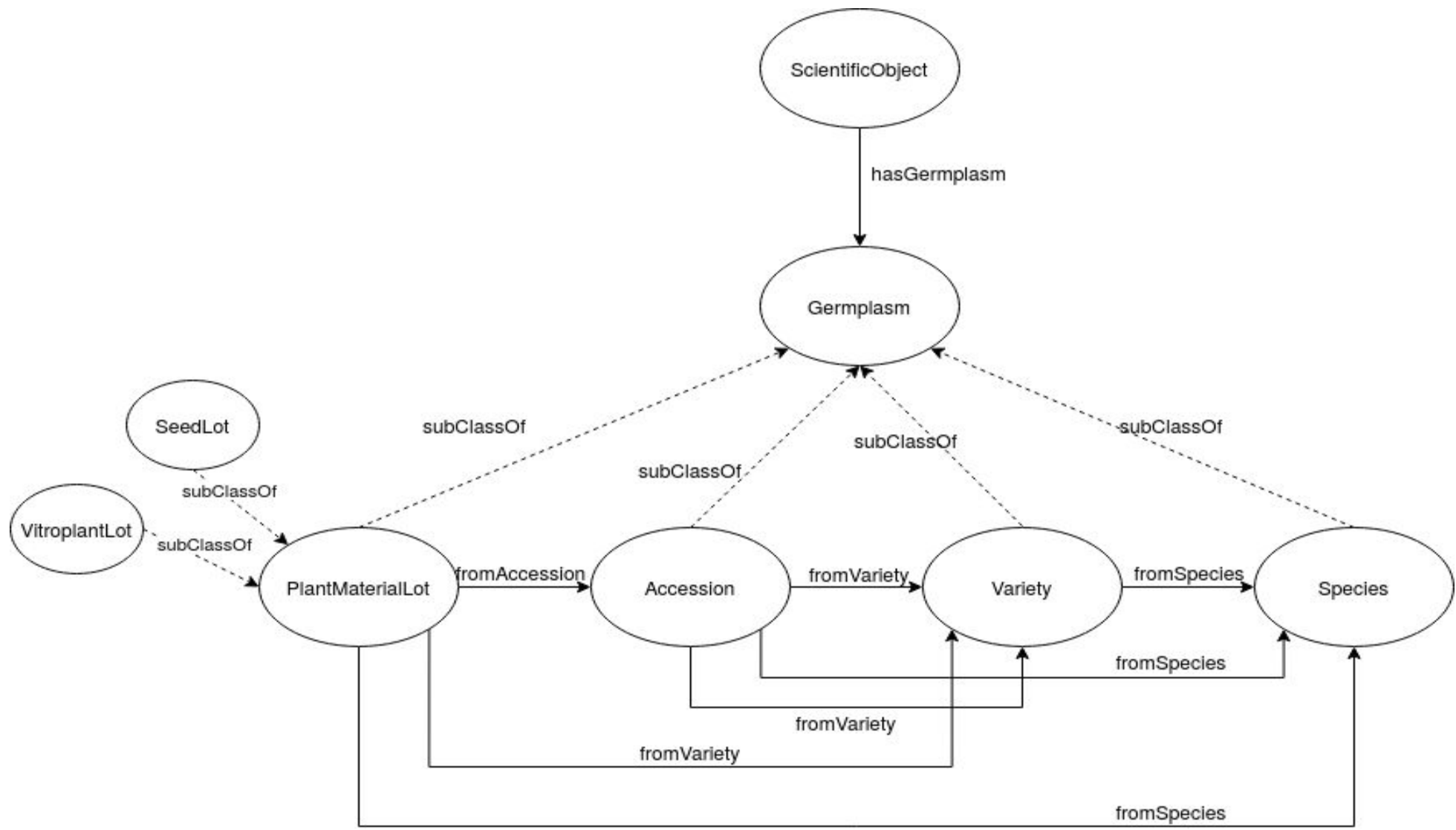
PHIS - Ontology for Experimental Scientific Objects (OESO)

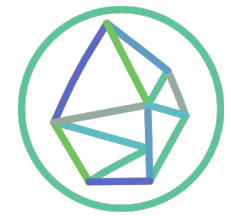
Zoom on Scientific Objects





PHIS - Germplasm (Genetic resources)





Germplasm information (MCPD-PHIS vocabulary)



MCPD	Accession dans PHIS
0. PUID	URI
1. Institute code	Institute
2. Accession number	AccessionNumber
5. Genus	<i>Unknown</i>
6. Species	Species URI
7. Species Authority	<i>Unknown</i>
8. Subtaxon	Variety URI
10. Common crop name	species name
11. Accession name	Name



PHIS Web Interfaces

Scientific information management

- Variables
- Scientific Objects
- Germplasm
- Documents

Huachano-R3

Germplasm

Details Annotations Documents

Description

URI <http://phis.egi-demo.eu/id/germplasm/variety.huachano-r3>

Type Variety

Name Huachano-R3

Description Transgenic sweetpotato expressing a viral protein (SPCSV RNase III)

Species sweetpotato http://aims.fao.org/aos/agrovoc/c_3937

SweetPotatoViruses

Experiment

Details Factors Scientific objects Data Visualization Map Annotations Documents

+ Add scientific object + CSV Import

Selected Scientific Objects 0 Actions + Export all

Select all			
<input type="checkbox"/>	SP004054 (plant)		
<input type="checkbox"/>	SP004055 (plant)		
<input type="checkbox"/>	SP004056 (plant)		
<input type="checkbox"/>	SP004057 (plant)		
<input type="checkbox"/>	SP004058 (plant)		
<input type="checkbox"/>	SP004059 (plant)		
<input type="checkbox"/>	SP004060 (plant)		
<input type="checkbox"/>	SP004061 (plant)		

Details Events Positions Annotations Documents

URI phis-egi-demo.id/scientific-object/so-sp004061

Name SP004061

Type Plant

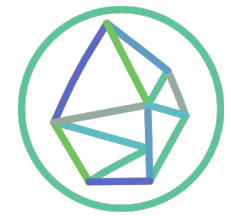
Creation date 07/13/2018

Factor level

- SPFMV (Viral Disease)

Germplasm

- [Huachano-R3](#)



PHIS Web Interfaces: Scientific Objects

OpenSILEX MAU17-PG Experiment

Scientific Organization

- Organizations
- Projects
- Experiments
- Facilities
- Devices
- Persons

Scientific Information

- Scientific Objects
- Variables
- Germplasm

Data

Vocabulary

Administration

Tools

component.menu.publication.li

Web API

Description Factors Scientific objects 2K+ Data 729 Visualization Map Annotations Documents

General informations

Name MAU17-PG

State Finished

Period 2017-05-19 - 2017-09-22 (4 months, 4 days)

URI <http://www.phenome-fppn.fr/diaphen/DIA2017-05-19>

Objective Genetic gain, maize, water deficit, field, diaphen

Description

Context



Projects [French plant phenomic network \(FPPN\)](#)

Organizations [DIAPHEN](#)

Facilities [mac3](#)

Species [Zea mays](#)


Factors [Irrigation](#)

Groups M3P , Guest 

Contacts

Scientific supervisors [Llorenç CABRERA-BOSQUET](#)

Technical supervisors [Romain Chapuis](#)

Declared by [admin admin](#) 



PHIS Web Interfaces: Scientific Objects

OpenSILEX MAU17-PG Experiment

Scientific Organization

Organizations
Projects
Experiments
Facilities
Devices
Persons

Scientific Information

Scientific Objects
Variables
Germplasm

Data
Vocabulary
Administration
Tools

Description Factors 1 Scientific objects 2K+ Data 729 Visualization Map Annotations Documents

+ Add scientific object + CSV Import

Selected Scientific Objects 0 Actions + Export all

Select all

- > 1/DZ_PG_67/ZM4394/WW/1/DIA2017-05-19 (plot)
- > 10/DZ_PG_48/ZM4421/WW/1/DIA2017-05-19 (plot)
- > 100/DZ_PG_16/ZM4359/WW/2/DIA2017-05-19 (plot)
- > 101/DZ_PG_11/ZM4386/WW/2/DIA2017-05-19 (plot)
- > 102/DZ_PG_13/ZM4379/WW/2/DIA2017-05-19 (plot)
- > 103/DZ_PG_48/ZM4421/WW/2/DIA2017-05-19 (plot)
- > 104/DZ_PG_14/ZM4375/WW/2/DIA2017-05-19 (plot)
- > 105/DZ_PG_02/ZM4356/WW/2/DIA2017-05-19 (plot)
- > 106/DZ_PG_24/ZM4342/WW/2/DIA2017-05-19 (plot)

100/DZ_PG_16/ZM4359/WW/2/DIA2017-05-19

Description Events 1 Positions 1 Annotations Documents

URI <http://www.phenome-fppn.fr/diaphen/2017/o17000100>

Name 100/DZ_PG_16/ZM4359/WW/2/DIA2017-05-19

Type Plot

Geometry WKT GeoJSON

Creation date 05/19/2017

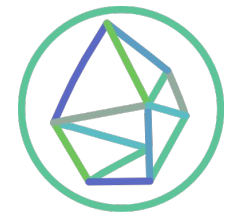
Germplasm

- ZM4359

Replication 2

Factor level

- WW (Irrigation)



PHIS Web Interfaces: Germplasm (Genetic resources)

PHIS Germplasm
Manage Genetic Resources Information

Scientific Organization
Organizations
Projects
Experiments
Facilities
Devices
Persons

Scientific Information
Scientific Objects
Variables
Germplasm
Data
Vocabulary
Administration
Tools
Open science

Type
Select a germplasm type

Species
Bread wheat

Production year
Enter a year

Institute code
Enter an institute code

Name
Enter germplasm name

Experiment
Select one experiment

Germplasm Group
Select one germplasm group

URI
Enter a part of an uri

Advanced Search
Reset Search

Selected Germplasm 0 Display Actions + Export all

Showing 60 to 80 of 563 entries

<input type="checkbox"/>	Name	Type	Species	Actions
<input type="checkbox"/>	ADJIKLI	Variety	Bread wheat	
<input type="checkbox"/>	AGASSIZ	Variety	Bread wheat	
<input type="checkbox"/>	AGRI_NAC_KAUZ_3_1D13_1_MLT	Variety	Bread wheat	
<input type="checkbox"/>	AGROR	Variety	Bread wheat	
<input type="checkbox"/>	AGROUNIA	Variety	Bread wheat	
<input type="checkbox"/>	AJAX_1994	Variety	Bread wheat	
<input type="checkbox"/>	AKASABI_SHIRAZU	Variety	Bread wheat	
<input type="checkbox"/>	AKULA_BONITO_F10S_3	Variety	Bread wheat	
<input type="checkbox"/>	ALBATROS_1950	Variety	Bread wheat	
<input type="checkbox"/>	ALBIDUM	Variety	Bread wheat	
<input type="checkbox"/>	ALCY_CAMBIER	Variety	Bread wheat	
<input type="checkbox"/>	ALKA	Variety	Bread wheat	



PHIS Web Interfaces: Germplasm (Genetic resources)

AGROR
Germplasm

Description Annotations Documents

General informations

URI <http://phenome.inrae.fr/diaphen/id/germplasm/variety.agror>

Type Variety

Name AGROR

Species Bread wheat http://aims.fao.org/aos/agrovoc/c_7951

Related Experiments

Search on experiment name

Showing 0 to 1 of 1 entries

Name	Species	Start date	End date	State
2024_PlastIGE	Bread wheat	2023-11-13	2024-12-31	

« < 1 > »

Related groups

Search by group name

No entries found



PHIS Web Interfaces: Germplasm (Genetic resources)

PHIS Declare Germplasm
Add species, varieties, accessions ...

Scientific Organization >
Scientific Information >
Scientific Objects
Variables
Germplasm
Data >
Vocabulary >
Administration >
About >
Tools >
Web API

Type ?
Variety

Download template Load CSV Reset table Add Row Add column

Check Insert

	URI	Name*	Subtaxa	Variety Code	Species URI*	Institute Code
1	http://Triticum_aestivum/AARDEN/SAATEN_UNION	AARDEN			http://aims.fao.org/aos/agrovoc/c_7951	SAATEN-UNION
2	http://Triticum_aestivum/ACCESS/CPB_TWYFORD	ACCESS			http://aims.fao.org/aos/agrovoc/c_7951	CPB TWYFORD
3	http://Triticum_aestivum/ACTIVUS/SAATZUCHT_DONAU_Gmbh	ACTIVUS			http://aims.fao.org/aos/agrovoc/c_7951	SAATZUCHT DONAU Gmbh
4	http://Triticum_aestivum/ADAGIO/R2N	ADAGIO			http://aims.fao.org/aos/agrovoc/c_7951	R 2 N
5	http://Triticum_aestivum/ADAPTO/SYNGENTA	ADAPTO			http://aims.fao.org/aos/agrovoc/c_7951	SYNGENTA
6	http://Triticum_aestivum/ADES/SYNGENTA_SEEDS_SAS_(F)	ADES			http://aims.fao.org/aos/agrovoc/c_7951	SYNGENTA SEEDS SAS (F)
7	http://Triticum_aestivum/ADESSO/SAATZUCHT_DONAU_Gmbh	ADESSO			http://aims.fao.org/aos/agrovoc/c_7951	SAATZUCHT DONAU Gmbh
8	http://Triticum_aestivum/AFRICA/GAE_RECHERCHE_SA	AFRICA			http://aims.fao.org/aos/agrovoc/c_7951	GAE RECHERCHE SA
9	http://Triticum_aestivum/AFRODITE/ISEA_SPA	AFRODITE			http://aims.fao.org/aos/agrovoc/c_7951	ISEA SPA
10	http://Triticum_aestivum/AGADIR/GAE_RECHERCHE_SA	AGADIR			http://aims.fao.org/aos/agrovoc/c_7951	GAE RECHERCHE SA
11	http://Triticum_aestivum/AGUILA/SECOBRA_RECHERCHES_SA	AGUILA			http://aims.fao.org/aos/agrovoc/c_7951	SECOBRA RECHERCHES SA
12	http://Triticum_aestivum/AIRBUS/LIMAGRAIN_EUROPE	AIRBUS			http://aims.fao.org/aos/agrovoc/c_7951	LIMAGRAIN EUROPE
13	http://Triticum_aestivum/AKAMAR/NICKERSON_INTERNATIONAL_RESEAR	AKAMAR			http://aims.fao.org/aos/agrovoc/c_7951	NICKERSON INTERNATIONAL RESEAR
14	http://Triticum_aestivum/AKIM/NICKERSON_INTERNATIONAL_RESEAR	AKIM			http://aims.fao.org/aos/agrovoc/c_7951	NICKERSON INTERNATIONAL RESEAR
15	http://Triticum_aestivum/ALATUS/	ALATUS			http://aims.fao.org/aos/agrovoc/c_7951	
16	http://Triticum_aestivum/ALBAGRAN/MOMONT_HENNETTE_ET_FILS	ALBAGRAN			http://aims.fao.org/aos/agrovoc/c_7951	MOMONT HENNETTE ET FILS
17	http://Triticum_aestivum/ALBATOR/UNISIGMA	ALBATOR			http://aims.fao.org/aos/agrovoc/c_7951	UNISIGMA

Conclusion

How to describe your plant phenomic experiment?



MCPD



- Standards & Tools : make your life easier !
- The Plant science community is active, with well-identified and shared standards, as well as tools to implement these standards
- https://rdmkit.elixir-europe.org/plant_sciences
- MIAPPE can be implemented using Excel files, but also using MIAPPE compliant information systems.
 - PIPPA: <https://pippa.psb.ugent.be/>
 - FAIRDOMSeek: <https://seek4science.org/>
 - PHIS: <http://www.phis.inrae.fr/>

Thank you for your attention!

INRAE

PHENOME
EMPHASIS | FRANCE

<https://www.phenome-emphasis.fr/>

Mistea
Mathématiques, Informatique et Statistique
pour l'Environnement et l'Agronomie

URGI

LEPSE
PLANT ADAPTATION
TO CLIMATE CHANGE



OpenSILEX Team - <http://opensilex.org/>



<http://www.phis.inrae.fr/>

Special thanks to: Silvana Moscatelli, François Tardieu, ...

Thank you very much
for your attention!

 emphasis@fz-juelich.de

 emphasis.plant-phenotyping.eu

 EMPHASIS_EU

 EMPHASIS.EU

 EMPHASIS on Plant Phenomics



EMPHASIS is an ESFRI-listed project.



EMPHASIS-PREP is funded by the European Union (Grant Agreement: 739514).

EUROPEAN INFRASTRUCTURE
FOR PLANT PHENOTYPING