



**HAL**  
open science

## Standards for distributed Plant Phenotyping data integration: challenges and solutions.

Cyril Pommier

► **To cite this version:**

Cyril Pommier. Standards for distributed Plant Phenotyping data integration: challenges and solutions.. COST FA1306 "The quest for tolerant varieties - Phenotyping at plant and cellular level" - WG1 Meeting "Phenotyping: from the lab to the field", Mar 2018, Leuven, Belgium. pp.67 slides. hal-01779751

**HAL Id: hal-01779751**

**<https://hal.science/hal-01779751>**

Submitted on 5 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

# Standards for Distributed Plant Phenotyping data integration



## Challenges and Solutions

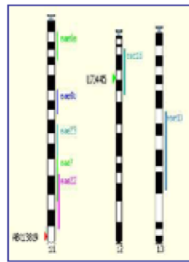
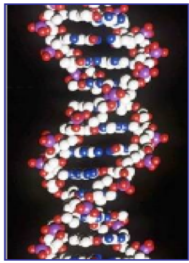


# Outline

- FAIR principles
- Data model
- Standards
- Use cases
  - ◆ Information systems : traceability and integration
  - ◆ Standard File format
  - ◆ Web services for analysis and interoperability
  - ◆ Semantic Web & RDF
- Perspectives



# Phenotyping $P = G \times E \times C$

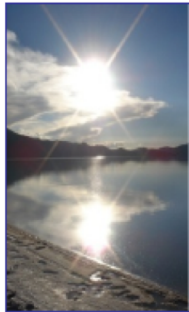


**Plant Genotype**

**Interactions**



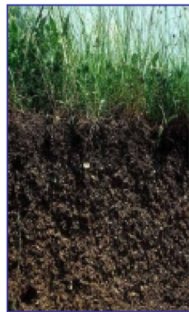
**Environment**



Climate



Cultural Practice



Soil



Stress

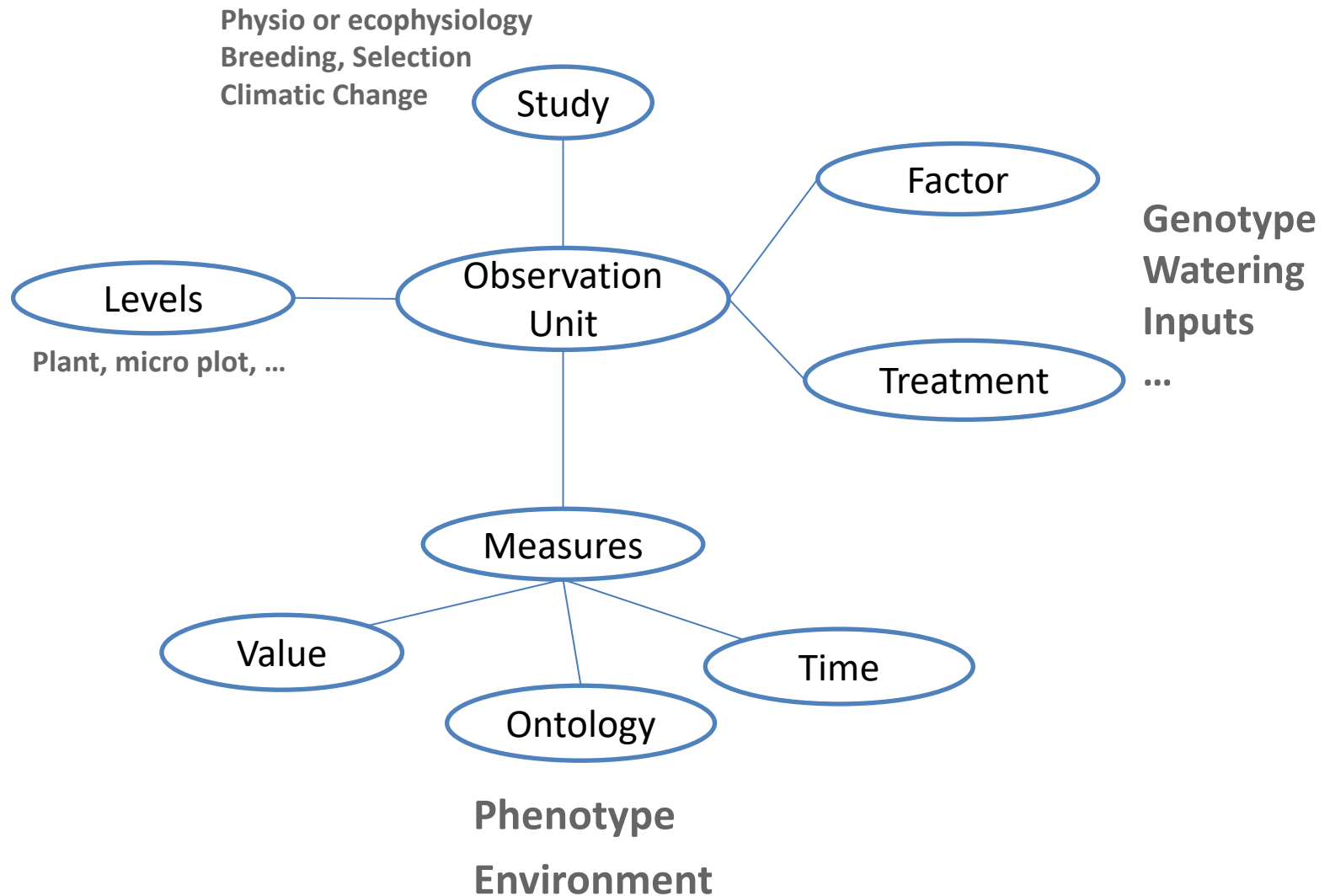
**Phenotype**



Phenology, Structure, Biochemical content

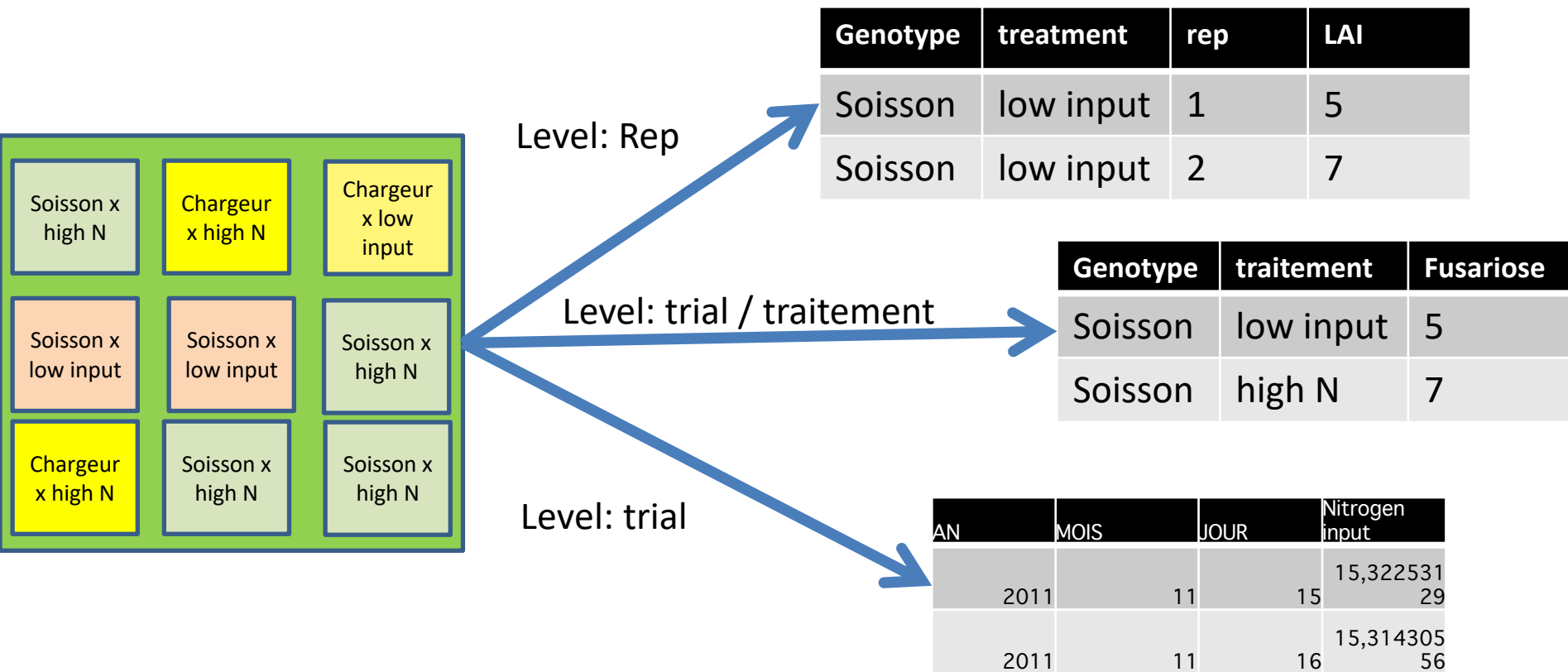
- Trait/Variable : a definition of an observable characteristic
  - ◆ Yield
- Phenotype : A trait/Variable with a value
  - ◆ Yield = 15 Kg by hect

# Study/Trial organisation



# Data life cycle

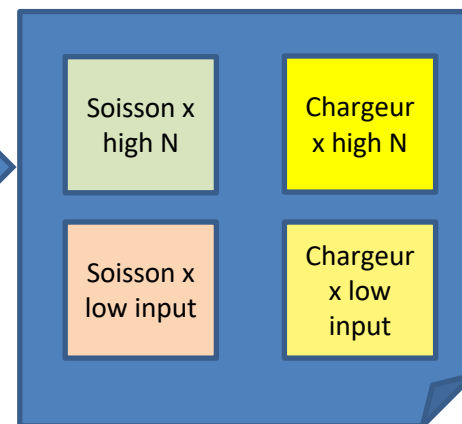
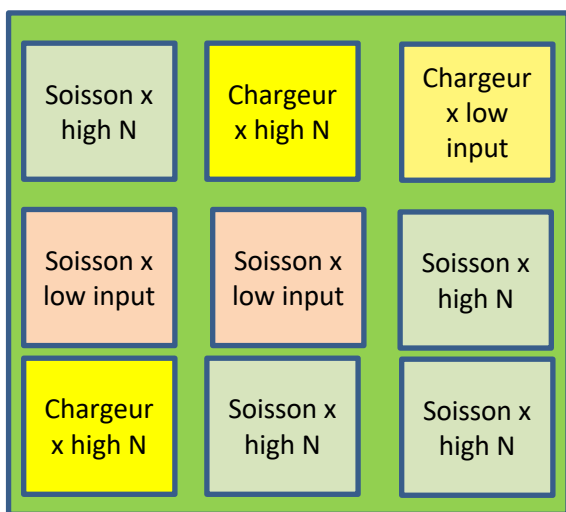
## Raw data



# Data Elaboration

Raw data

Elaborated data



| Genotype | treatment | rep | LAI |
|----------|-----------|-----|-----|
| Soisson  | fi        | 1   | 5   |
| Soisson  | fi        | 2   | 7   |

| Genotype | treatment | LAI |
|----------|-----------|-----|
| Soisson  | low input | 6   |

# Data elaboration

- NIRS

| Individual  | Genotype  | Repetition | Date  | 4000                  | 4002                  | 4004                  | 4006                  |
|-------------|-----------|------------|-------|-----------------------|-----------------------|-----------------------|-----------------------|
| 662200657-1 | 662200657 | 1          | Nov11 | 0.790383794657<br>181 | 0.790834341907<br>288 | 0.791261293580<br>598 | 0.791659233526<br>804 |
| 662200657-2 | 662200657 | 2          | Nov11 | 0.817254078896<br>638 | 0.817865344054<br>208 | 0.818450087726<br>549 | 0.819007314250<br>037 |
| 662200659-1 | 662200659 | 1          | Nov11 | 0.822434379133<br>726 | 0.822871701993<br>426 | 0.823268552780<br>517 | 0.823649256266<br>632 |

- After Calibration and Analysis

| Individual  | Genotype  | Repetition | Gluc_sol    | Gluc_solub | Gluc_hydrol | NIRS_id | Date  |
|-------------|-----------|------------|-------------|------------|-------------|---------|-------|
| 662200659-1 | 662200659 | 1          | 0.064074733 | 0.38557928 | 0.321504547 | 2274    | Nov11 |



# Data elaboration

- LIDAR
  - ◆ Very Big files
  - ◆ Plant Architecture
  - ◆ Raw extraction :
    - Leaf area, leaf angle
  - ◆ Elaboration → % interception
- Metabolomic
  - ◆ MSI Compliant Full description
  - ◆ Elaboration →
    - Leaf Protein concentration
    - Root Cellulose concentration

# Data elaboration

- Transformation and computation operation
- Future new analysis to elaborate new variables.
  - Raw data long term conservation
- Need
  - ◆ Reproducibility
  - ◆ Traceability
  - ◆ Provenance

→ FAIR

# FAIR Data Principles

Ids  
Metadata  
Index

Open Protocols  
Persistent Metadata

Semantics  
Linked Data  
Vocabularies

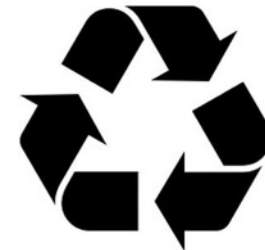
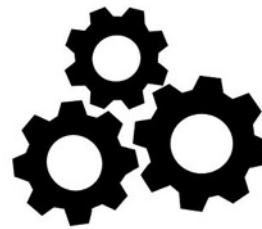
License  
Well described  
Provenance  
Standards

F  
Findable

A  
Accessible

I  
Interoperable

R  
Reusable



- Avoid Data Silos
- Interoperable Data and Tools Integration
  - ◆ Linked data
- Repositories Complementarity
  - ◆ Volume and Variety Burden Sharing

## FAIR: DATA STANDARDS



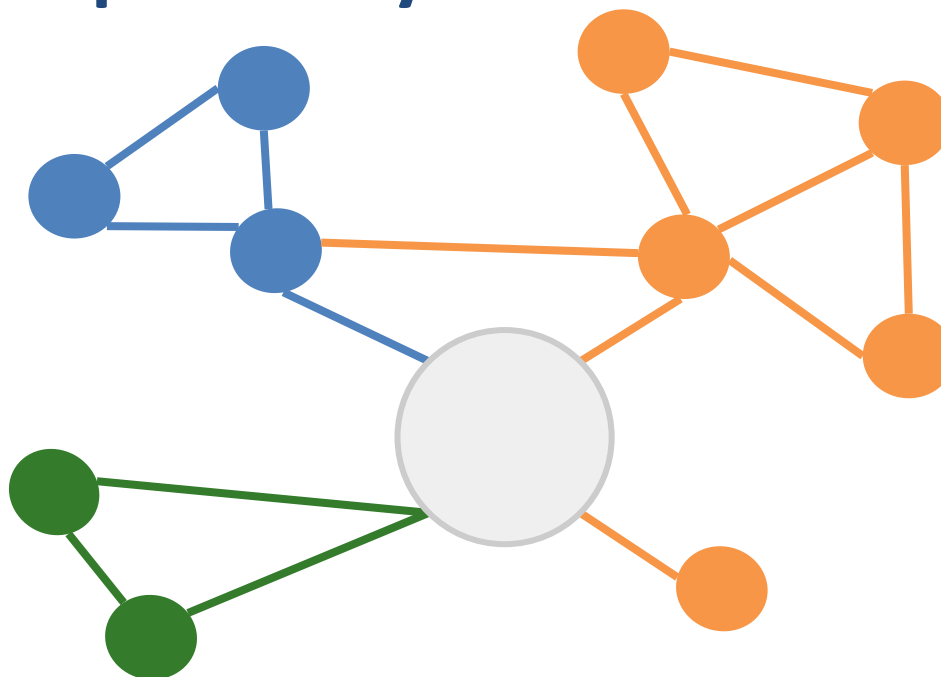
Metadata  
Semantic

# Interoperability in International network

## National Networks



## European Networks



## Global Networks



## International data standards



Controlled vocabularies  
Trait dictionaries



# WheatIS standards

- <http://wheatis.org/DataStandards.php>
- Phenotypes guidelines
- MIAPPE
- Croponology
- Breeding API

## Wheat Data Interoperability Guidelines

Home Guidelines ▾ Ontologies & Vocabularies Use cases ▾ Getting involved About ▾

Home / Phenotypes

### Phenotypes

Characteristics of an organism resulting from interactions between genes and the environment. Plant phenotyping can be used in several fields, such as breeding and agronomy. Phenotypic data (traits and associated value) can be used in several fields of research in multi-location environments, germplasm bank and crop improvement. Furthermore, phenotypes are in interaction, like for instance in crop yield.

Getting Involved



## 1. Data formats

We recommend following minimum format principles with data variables (trait along with method, units and scales or environmental conditions). This principle has been formalized in the [MIAPPE](#) under improvement.

The [Breeding API](#), described below, is a MIAPPE compliant web service API specification.

There is also an emerging initiative to produce a semantic web version of MIAPPE through [to RDF transformation \(new!\)](#).

See the [germplasm recommendations](#) for data format regarding germplasm information.

- Reuse for other crops: Rice, ...

## Wheat Data Interoperability Guidelines

[Home](#) [Guidelines](#) [Ontologies & Vocabularies](#) [Use cases](#) [Getting involved](#) [About](#)

[Home](#) / [Getting involved](#)

### Getting involved

## See who endorses the guidelines

To stay relevant and useful for the wheat data community, it is important that the recommendations contained here are maintained. This will be especially true as new technologies are developed requiring standards to be agreed. The maintenance of the guidelines will depend on an ongoing process of monitoring practices across the wheat community, and maintaining communication with those who are producing and using wheat data. There are many ways you can help.

**Become an adopter:** Please let us know why and how these guidelines are helpful to you and your institution. See the [Current list of adopters](#)

**Give your feedback:** You can contact us either by email or by leaving a comment on a specific page. This is an easy way to propose improvements or corrections, update the state of the art and share best practices.

#### Getting Involved



Whe

# INTEROPERABILITY: SEMANTIC

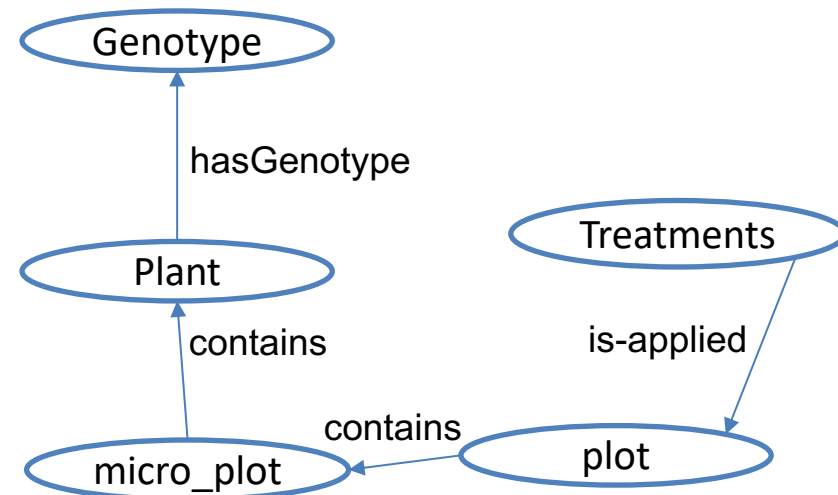
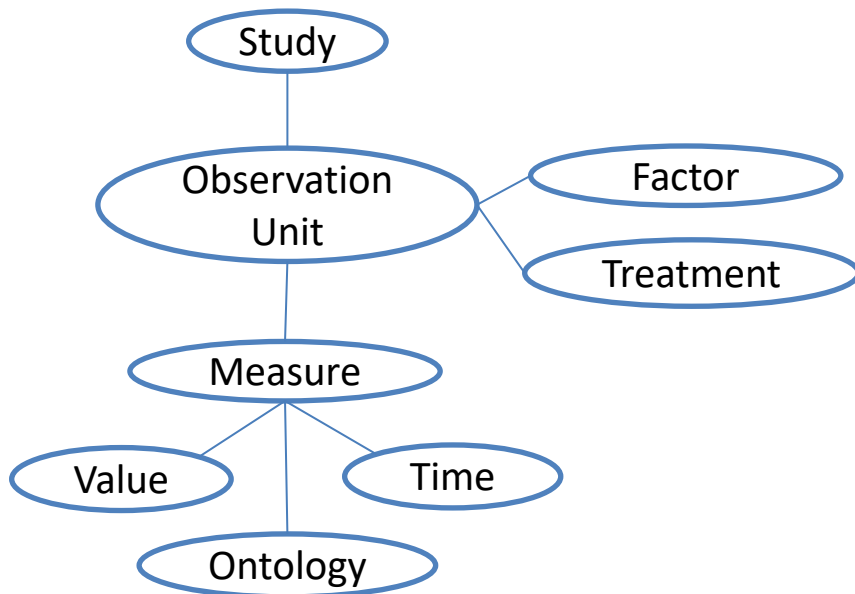


FAIR

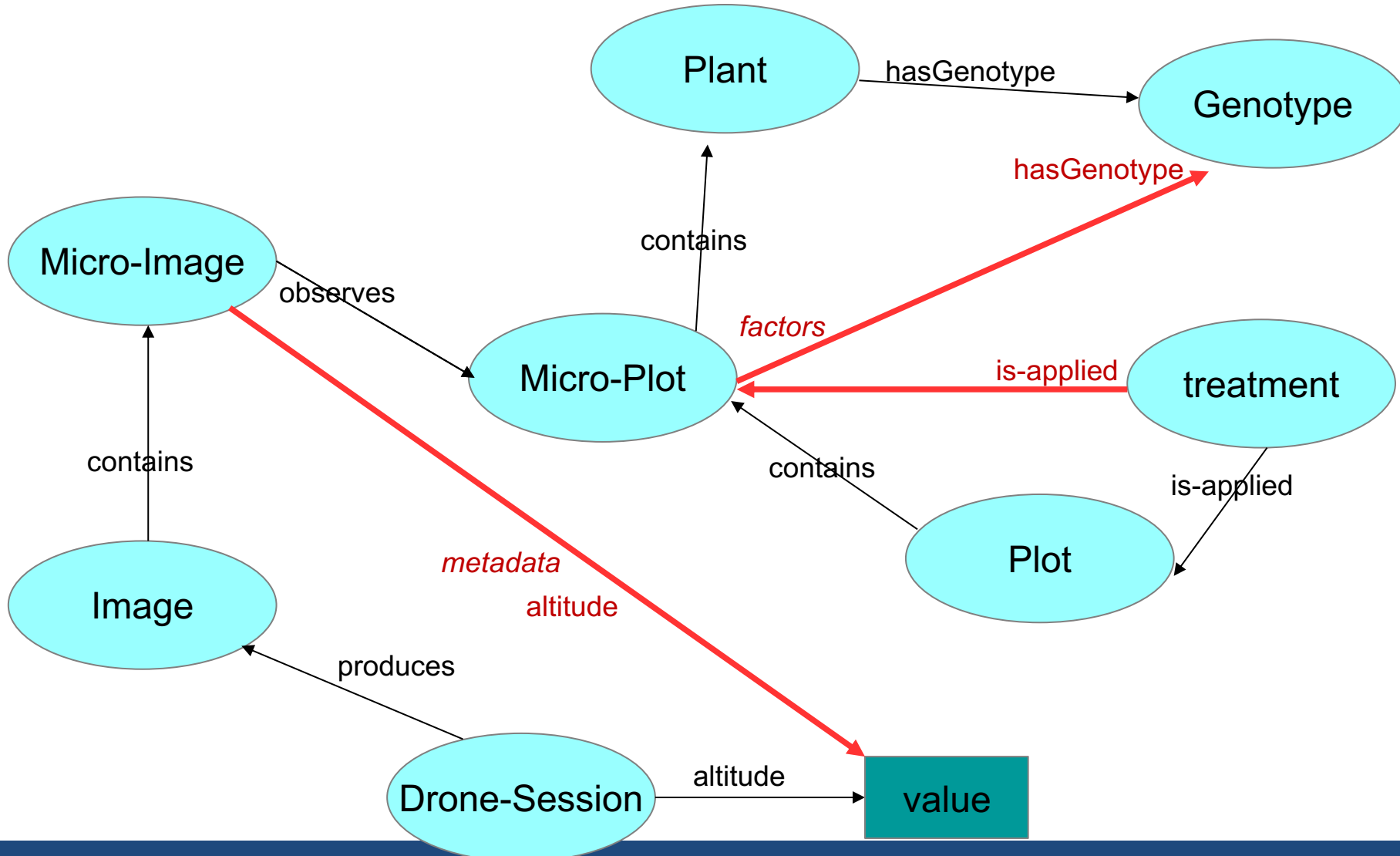


# Semantic

- Define concepts: plant, micro\_plot, phenotype, experiment, event
- Define Relations: micro\_plot contains plant
- Common Vocabularies and ontologies
  - ◆ Agrovoc, Plant Ontology, Crop Ontology, Environment Ontology, ...



# Semantic Inference



# Semantic Inference

- Enables Provenance and Reuse
- Uncouple
  - ◆ Data capture
  - ◆ Data analysis
  - ◆ Data reuse
- Inference allows to format the data for a target use

Elizabeth Arnaud et al.

## Phenotype data in breeding

| Germplasm ID | PH  | PH2 | GCOL | GY |
|--------------|-----|-----|------|----|
| 24530        | 80  | 1   | 2    | 35 |
| 85432        | 120 | 3   | 4    | 48 |
| 78452        | 95  | 2   | 4    | 43 |
| 56093        | 100 | 2   | 1    | 50 |

The height of the plant

The color of the grain

The grain yield

Annotation must explain:

1/ What is the observation about ?

=

**TRAIT or Environment**

Elizabeth Arnaud et al.

## Phenotype data in breeding

Annotation must explain:

| Germplasm ID | PH  | PH2 | GCOL | GY |
|--------------|-----|-----|------|----|
| 24530        | 80  | 1   | 2    | 35 |
| 85432        | 120 | 3   | 4    | 48 |
| 78452        | 95  | 2   | 4    | 43 |
| 56093        | 100 | 2   | 1    | 50 |

2/ How is the trait observed ?

=

**METHOD**

With a measuring tape

Estimated visually

Calculated  
Panicle weight x panicle density

Elizabeth Arnaud et al.

## Phenotype data in breeding

Annotation must explain:

| Germplasm ID | PH  | PH2 | GCOL | GY |
|--------------|-----|-----|------|----|
| 24530        | 80  | 1   | 2    | 35 |
| 85432        | 120 | 3   | 3    | 48 |
| 78452        | 95  | 2   | 3    | 43 |
| 56093        | 100 | 2   | 1    | 50 |

3/ How is the trait observation expressed?

=

**SCALE**

cm

1 = short (<90cm)  
2 = medium (>90cm and <110cm)  
3 = tall (>110cm)

1 = white  
2 = cream  
3 = yellow

g/m<sup>2</sup>

Elizabeth Arnaud et al.

# Phenotype data in breeding

| Germplasm ID | PH  | PH2 | GCOL | GY |
|--------------|-----|-----|------|----|
| 24530        | 80  | 1   | 2    | 35 |
| 85432        | 120 | 3   | 3    | 48 |
| 78452        | 95  | 2   | 3    | 43 |
| 56093        | 100 | 2   | 1    | 50 |

1 VARIABLE  
=  
1 TRAIT/ENV  
+ 1 METHOD  
+ 1 SCALE

**Woody Plant Ontology** ONTOLOGY

Ontology name: Woody Plant Ontology  
 Authors: Célia Michotey  
 Version: v1  
 Licence: CC BY 4.0  
 Links: [CropOntology TVD5 format](#)  
[Adonis format](#)

Community  
Species  
Ontologies

### Traits, methods and scales

Search terms...

Woody Plant Ontology ONTOLOGY

- └ Agronomical TRAIT CLASS
- └ Fertility TRAIT CLASS
- └ Morphological TRAIT CLASS
- └ Other TRAIT CLASS
- └ Phenological TRAIT CLASS
- └ Physiological TRAIT CLASS
- └ Quality TRAIT CLASS
- └ Stress TRAIT CLASS
- └ Stress: Abiotic stress TRAIT CLASS
- └ Stress: Biotic stress TRAIT CLASS
  - └ Bacterial canker girdling index TRAIT
  - └ Bacterial canker lesion length TRAIT
  - └ Game damages TRAIT
  - └ Nests presence TRAIT
  - └ NID: Nests presence VARIABLE

### NID: Nests presence VARIABLE

|                                |  |
|--------------------------------|--|
| Identifier                     | WOODY:0000067                                    |
| Name                           | NID  |
| Synonyms                       | Nests presence<br>NID[Adonis]                    |
| Context of use                 | Trial evaluation                                 |
| Status                         | Standard for INRA                                |
| Institution                    | INRA   |
| Scientist                      | Célia Michotey                                   |
| Date                           | 13/03/2017                                       |
| Crop                           | Woody Plant                                      |
| <b>Nests presence TRAIT</b>    |  |
| Identifier                     | WOODY:1000056                                    |
| Name                           | Nests presence                                   |
| Description                    | Presence or absence of nests                     |
| Main abbreviation              | NID  |
| Status                         | Standard for INRA                                |
| Class                          | Stress: Biotic stress                            |
| <b>Visual scoring METHOD</b>   |  |
| Identifier                     | WOODY:2000003                                    |
| Name                           | Visual scoring                                   |
| Description                    | Visual assessment with a reference scoring scale |
| Class                          | Estimation                                       |
| <b>Presence/ absence SCALE</b> |  |
| Identifier                     | WOODY:3000003                                    |
| Name                           | Presence/Absence                                 |
| Data type                      | Nominal  |
| Decimal places                 | 0  |
| Min                            | 0  |

# INTEROPERABILITY: STANDARD





# Minimum Information About Plant Phenotyping Experiment

- Standard
  - ◆ Data exchange and traceability
  - ◆ Repositories
- Input and output for analysis pipelines developed for phenotyping and genetic
  - ◆ Transplant, Elixir : European eInfrastructure for life science
  - ◆ EPPN and Emphasis European eInfrastructure (ESFRI) for plant Phenotyping
  - ◆ CGIARs
- Implementations
  - ◆ File Archive
  - ◆ Web Service
  - ◆ RDF

## Measures for interoperability of phenotypic data: minimum information requirements and formatting

[Hanna Ćwiek-Kupczyńska](#), [Thomas Altmann](#), [Daniel Arend](#), [Elizabeth Arnaud](#), [Dijun Chen](#), [Guillaume Cornut](#), [Fabio Fiorani](#), [Wojciech Frohberg](#), [Astrid Junker](#), [Christian Klukas](#), [Matthias Lange](#), [Cezary Mazurek](#), [Anahita Nafissi](#), [Pascal Neveu](#), [Jan van Oeveren](#), [Cyril Pommier](#), [Hendrik Poorter](#), [Philippe Rocca-Serra](#), [Susanna-Assunta Sansone](#), [Uwe Scholz](#), [Marco van Schriek](#), [Ümit Seren](#), [Björn Usadel](#), [Stephan Weise](#), [Paul Kersey](#) and [Paweł Krajewski](#) ✉

*Plant Methods* 2016 12:44 | DOI: 10.1186/s13007-016-0144-4 | © The Author(s) 2016

Received: 15 April 2016 | Accepted: 18 October 2016 | Published: 9 November 2016



## Towards recommendations for metadata and data handling in plant phenotyping

[Paweł Krajewski](#)<sup>1,\*</sup>, [Dijun Chen](#)<sup>2</sup>, [Hanna Ćwiek](#)<sup>1</sup>, [Aalt D.J. van Dijk](#)<sup>3</sup>, [Fabio Fiorani](#)<sup>4</sup>, [Paul Kersey](#)<sup>5</sup>, [Christian Klukas](#)<sup>2</sup>, [Matthias Lange](#)<sup>2</sup>, [Augustyn Markiewicz](#)<sup>6</sup>, [Jan Peter Nap](#)<sup>3</sup>, [Jan van Oeveren](#)<sup>7</sup>, [Cyril Pommier](#)<sup>8</sup>, [Uwe Scholz](#)<sup>2</sup>, [Marco van Schriek](#)<sup>7</sup>, [Björn Usadel](#)<sup>4,9</sup> and [Stephan Weise](#)<sup>2</sup>

[« Previous](#) | [Next Article »](#)  
[Table of Contents](#)

### This Article

*J. Exp. Bot.* (2015) 66 (18): 5417-5427.  
doi: 10.1093/jxb/erv271  
First published online: June 4, 2015  
This article appears in: [Special](#)

# MIAPPE Construction

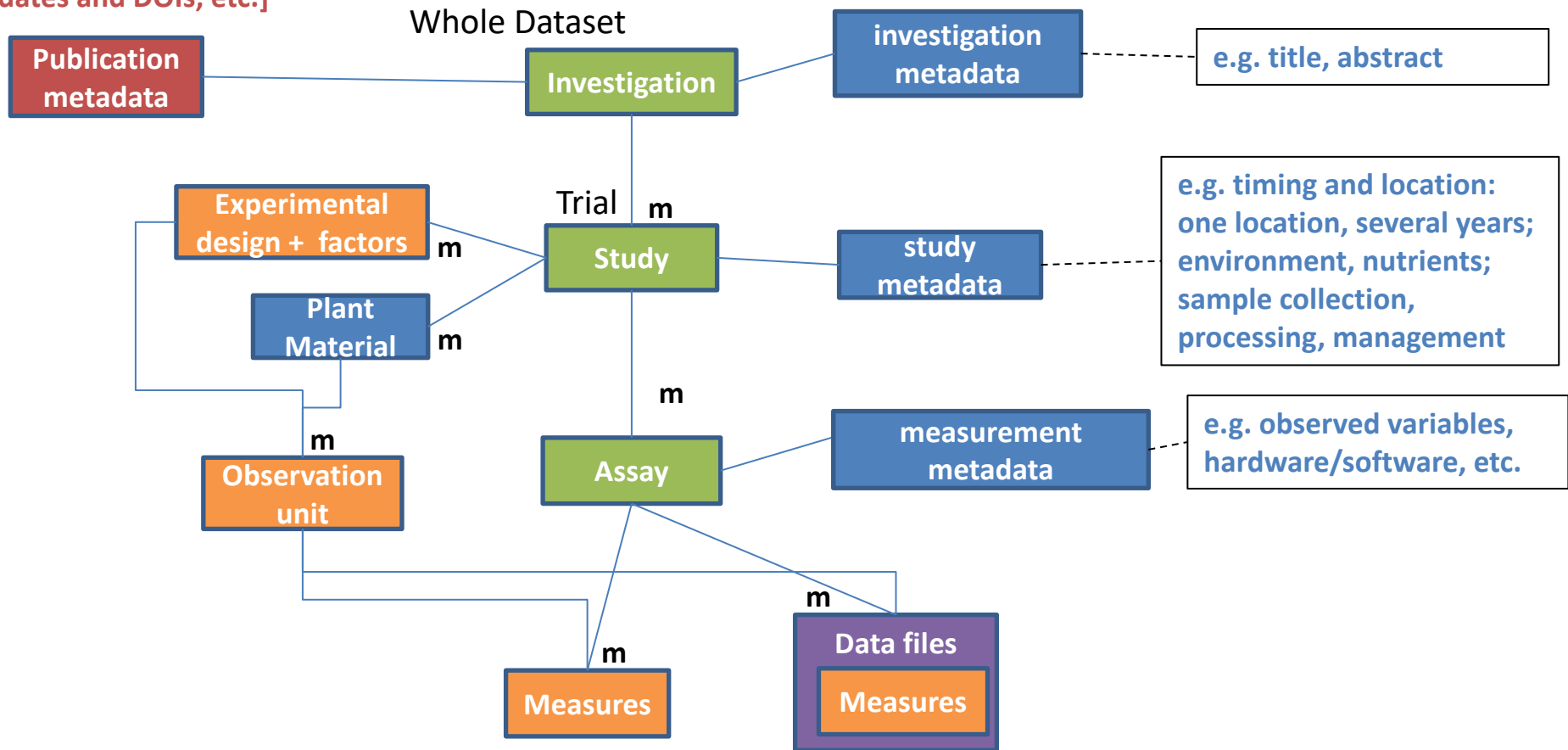
- Contribution
  - ◆ European Infrastructures : Elixir (Bioinformatics), Emphasis/EPPN (Phenotyping)
  - ◆ National Institutes: France, Germany, Poland, UK, Portugal, Slovenia, Nederland, Belgium, Italy
- Steering committee
  - ◆ Elixir: European Bioinformatic infrastructure
  - ◆ Emphasis: European plant phenotyping infrastructure
  - ◆ Bioversity International - CGIAR
  - ◆ Elizabeth Arnaud, Paul Kersey, Pawel Krajewsky, Matthias Lange, Cyril Pommier, Bjorn Usadel
- Current Versions
  - ◆ Version 1 : Transplant + EPPN + Phenome
  - ◆ Version 1.1 : Elixir, validation in progress
    - Documentation and model: Biologist friendly
    - Adaptations to new implementations
    - Proposal open for consultation contribution welcomed
    - See <http://www.miappe.org> for providing feedback.



# MIAPPE v1.1 Overview

Ownership, contact details, [database accessions, publication dates and DOIs, etc.]

Primary metadata



# MIAPPE v1.1 Overview

## Plant Material

- **Biosource : Plant Material identification**
- Multi Crop Passport Descriptor (MCPD) compliant
  - ◆ <https://www.biodiversityinternational.org/e-library/publications/detail/faobiodiversity-multi-crop-passport-descriptors-v21-mcpd-v21/>
  - ◆ FAO & Biodiversity International
  - ◆ Genebanks
- ID Field are reused in MIAPPE
- Key : Identification
  - ◆ 0. Persistent unique identifier : PUID = URI/DOI
  - ◆ 1. Institute code
  - ◆ 2. Accession number
  - ◆ 5. Genus
  - ◆ 6. Species
  - ◆ 7. Subspecies : ‘subsp.’ (for subspecies); ‘convar.’ (for convariety); ‘var.’ (for variety); ‘f.’ (for form); ‘Group’ (for ‘cultivar group’)
  - ◆ 28. Remarks
- Suitable for non Genbank collections : breeder varietal list, laboratory collection, ...

- Observation Variables
  - ◆ Phenotype & environment
  - ◆ Trait : What is measured
  - ◆ Method : How is it measured
  - ◆ Scale : How is it observed
  - ◆ Variable : trait + method + scale

**Traits, methods and scales**

DOWNLOAD SHOW OBSOLETE TERMS English

- Wheat traits
  - Abiotic stress traits
    - Acid soil tolerance
      - AcidTol Estimation (method\_of)
    - Aluminum tolerance
    - Boron tolerance
    - Cold tolerance
    - Drought tolerance
    - Frost damage
    - Hail damage
    - Heat tolerance
    - Herbicide damage
    - Leaf rolling incidence

**Term information**

AcidTol Estimation

Identifier CO\_321:0000299

Method class Estimation

Method description Observe the growth low pH (acid) soil fi

Method name AcidTol Estimation

Method reference Reynolds et al., 200

**Traits, methods and scales**

Search terms...

Woody Plant Ontology ONTOLOGY

**Woody Plant Ontology ONTOLOGY**

Ontology name Woody Plant Ontology

Authors Célia Michotey

Version v1

Licence CC BY 4.0

Links [CropOntology.TVD5 format](#)  
[Adonis format](#)

Physiological TRAIT CLASS

Quality TRAIT CLASS

Stress TRAIT CLASS

Stress: Abiotic stress TRAIT CLASS

Stress: Biotic stress TRAIT CLASS

Bacterial canker girdling index TRAIT

Bacterial canker lesion length TRAIT

Game damages TRAIT

**NID: Nests presence VARIABLE**

Identifier WOODY:0000067

Name NID

Synonyms Nests presence  
NID[Adonis]

Context of use Trial evaluation

Status Standard for INRA

Institution INRA

Scientist Célia Michotey

Date 13/03/2017

Crop Woody Plant

**Nests presence TRAIT**

Identifier WOODY:1000056

Name Nests presence

Description Presence or absence

Main abbreviation NID

Status Standard for INRA

Class Stress: Biotic stress

**Visual scoring METHOD**

Identifier WOODY:2000003

Name Visual scoring

Description Visual assessment scoring scale

Class Estimation

**Presence/ absence SCALE**

Identifier WOODY:3000003

# MIAPPE v1.1 overview:

## Metabolomic

- Existing standards
  - ◆ Metabolights, Minimum information MSI
  - ◆ Full traceability
  - ◆ Sample treatment : temperature, tissue processing, molecule extraction protocol, measurement protocol....
- MIAPPE
  - ◆ Elaborated data
  - ◆ Elaborated, computed variables: Leaf Sugar content by microplot
  - ◆ Link to Metabo archive
    - Metabolight
    - Metabohub
    - ...

## USE CASES



Information Systems for publication and traceability

File archive

Web services

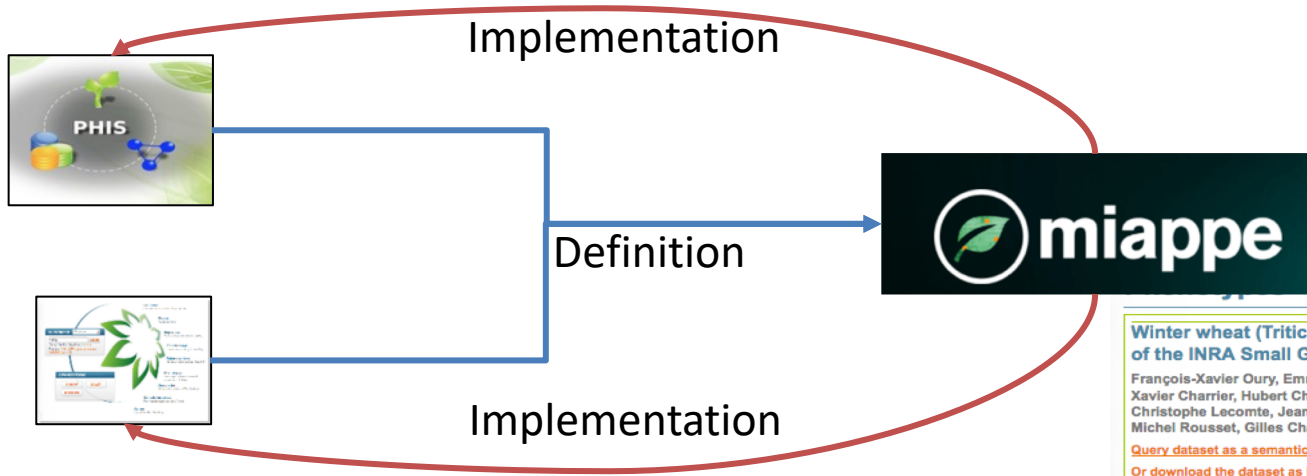
Semantic web and RDF

## USE CASE: DATA REPOSITORIES





# MIAPPE: Databases



- Production databases
  - ◆ PHIS
- Cleaned data publication repositories
  - ◆ GnpIS
    - <http://dx.doi.org/10.15454/1.4489666216568333E12>
  - ◆ eDale
  - ◆ PlantPhenoDB
- Implementation
  - ◆ Internal
  - ◆ Web services API

Winter wheat (*Triticum aestivum* L) phenotypic data from the multiannual, multilocal trial of the INRA Small Grain Cereals Network.

François-Xavier Oury, Emmanuel Heumez, Bernard Rolland, Jérôme Auzanneau, Pierre Bérard, Maryse Brancourt, Xavier Charrier, Hubert Chiron, Camille Depatureaux, Laurent Falchetto, Olivier Gardet, Stéphane Gilles, Alex Girard, Christophe Lecomte, Jean-Yves Morlais, Pierre Pluchard, Didier Tropée, Maxime Trottet, Patrice Walczak, Gérard Walczak, Michel Rousset, Gilles Charmet

[Query dataset as a semantic graph.](#)  
[Or download the dataset as RDF archive.](#)

Abstract  
 Published 2015 by INRA

[Back to Form](#)

Search parameter(s):  
 Geolocation

DATA SETS: 4  
 Network Data Set : [INRA Wheat Network BRC accession \(A series\)](#)  
 Network Data Set : [INRA Small Grain Cereals Network](#)  
 DOI: <http://dx.doi.org/10.15454/1.4489666216568333E12>  
 Network Data Set : [INRA Wheat Network not BRC accession \(B and C series\)](#)

Origin site Collecting site Evaluation site

Phenotyping campaign(s)  
 2000 x 2001 x 2002 x 2003 x 2004 x 2005 x 2006 x 2007 x  
 2009 x 2010 x 2011 x 2012 x 2013 x 2014 x 2015 x  
[remove all](#) [add all](#)

Trial list Phenotypic data

ZURGI GnpIS 813 trials

| Name                              | Campaign | Site               | Direct |
|-----------------------------------|----------|--------------------|--------|
| BTH_Champagne-céréales_2004_SetA1 | 2004     | Champagne-céréales |        |
| BTH_Champagne-céréales_2005_SetA1 | 2005     | Champagne-céréales |        |

# MIAPPE databases & FAIR

- Findable: Indexed metadata
- Interoperable : Vocabularies/Ontologies, Plant material ID

## Phenotypes

Genotype Variable Trial Reset Form Results

Traits, methods and scales

Search terms...

- Wheat Inra Phenotype Ontology **ONTOLOGY**
  - Abiotic stress **TRAIT CLASS**
    - Frost susceptibility **TRAIT**
    - Lodging score **TRAIT**
      - LS\_score: Lodging score,Susceptibility to lodging** **VARIABLE**
        - Ontology name: Wheat Inra Phenotype Ontology
        - Identifier: WIPO:0000099
        - Name: LS\_score
        - Synonyms: Lodging score,Susceptibility to lodging
        - Institution: INRA
        - Scientist: Jacques Le Gouis
        - Date: 15/06/2016
        - Crop: Wheat
      - Lodging score **TRAIT**
        - Identifier: WIPO:0010030
        - Name: Lodging score
        - Class: Abiotic stress
      - Direct measure **METHOD**
        - Identifier: WIPO:0020016
        - Name: Direct measure
        - Description: Mesure directe # ModeOp Arvalis Echelle de 1-10 prenant en compte %surface versée et l'angle de vers
      - 1 to 9 note **SCALE**
        - Identifier: WIPO:0030006
        - Name: 1 to 9 note
        - Data type: Ordinal
        - Min: 1
        - Max: 9
        - Categories: 1 = Résistant 0 = susceptible
- Agronomical **TRAIT CLASS**
- Biotic stress **TRAIT CLASS**
- Morphological **TRAIT CLASS**
- Other **TRAIT CLASS**
- Quality **TRAIT CLASS**

## Phenotypes

Genotype Variable Trial Reset Form Results

Reset Tab

Add accession by Genus

Choose a species

Add accession by Panel

Add accession by Collection

Add accession by Name

Paste your Accession Name list

Accession 1-10 of 2,556

| Accession Number | Accession Name  | Taxon Name                 |
|------------------|-----------------|----------------------------|
| TX816            | 02438-6HD-106-1 | Triticum aestivum aestivum |
| TX817            | 02438-6HD-132-1 | Triticum aestivum aestivum |
| TX818            | 02438-6HD-141-1 | Triticum aestivum aestivum |
| TX811            | 02438-6HD-15-1  | Triticum aestivum aestivum |
| TX812            | 02438-6HD-32-1  | Triticum aestivum aestivum |
| TX813            | 02438-6HD-56-1  | Triticum aestivum aestivum |
| TX814            | 02438-6HD-78-1  | Triticum aestivum aestivum |
| TX815            | 02438-6HD-80-1  | Triticum aestivum aestivum |

# MIAPPE databases & FAIR

- Accessible
- Reusable: License, File standard
- Whole Phenotyping Network

## Phenotypes

### Winter wheat (*Triticum aestivum* L) phenotypic data from the multiannual, multilocal field trials of the INRA Small Grain Cereals Network.

François-Xavier Oury, Emmanuel Heumez, Bernard Rolland, Jérôme Auzanneau, Pierre Bérard, Maryse Brancourt-Hulmel, Xavier Charrier, Hubert Chiron, Camille Depatureaux, Laurent Falchetto, Olivier Gardet, Stéphane Gilles, Alex Giraud, Christophe Lecomte, Jean-Yves Morlais, Pierre Pluchard, Didier Tropée, Maxime Trottet, Patrice Walczak, Gérard Doussinault, Michel Rousset, Gilles Charmet

[Query dataset as a semantic graph.](#)

[Or download the dataset as RDF archive.](#)

#### Abstract

Winter wheat phenotypic data from a French experimental network. It includes observations for different characters (agronomic, quality, disease, phenology,...) on more than 10 experimental locations during more than 15 years and for more than 1700 winter wheat genotypes.

Published 2015 by INRA

[Back to Form](#)

#### Search parameter(s):

[Geolocation](#)

#### DATA SETS: 4

##### Network Data Set :

[INRA Wheat Network BRC accession \(A series\)](#)

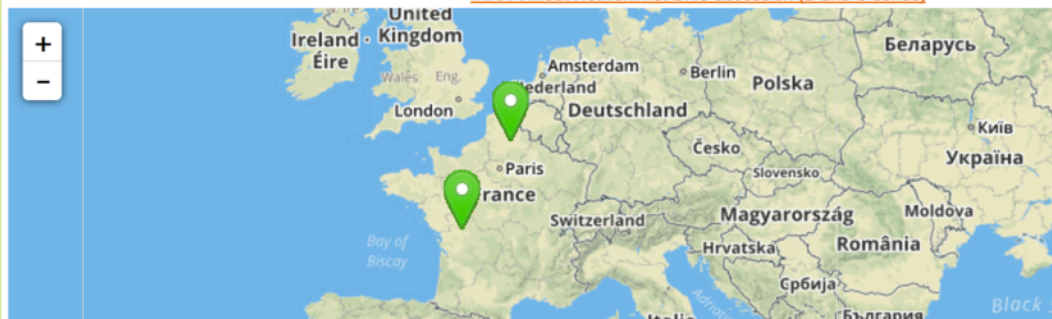
##### Network Data Set :

[INRA Small Grain Cereals Network](#)

DOI:<http://dx.doi.org/10.15454/1.4489666216568333E12>

##### Network Data Set :

[INRA Wheat Network not BRC accession \(B and C series\)](#)



Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

Origin site Collecting site Evaluation site

Phenotyping campaign(s): 2000 × 2001 × 2002 × 2003 × 2004 × 2005 × 2006 × 2007 × 2009 × 2010 × 2011 × 2012 × 2013 × 2014 × 2015 ×

[remove all](#) [add all](#)

**Trial list** **Phenotypic data**

LEVEL: TRIAL

1-10 of 52,875 | Display 10 results per page

| Lot Number | Accession Number        | Accession Name | ltk                       | Trial Name                |
|------------|-------------------------|----------------|---------------------------|---------------------------|
| AO13047    | <a href="#">39291</a>   | AO13047        | t: traite                 | BTH_Champagne-céréale     |
| AO14001    | <a href="#">AO14001</a> | AO14001        | t: traite                 | BTH_Champagne-céréale     |
| Bermude    | <a href="#">TX9</a>     | Bermude        | fi: faible intrant        | BTH_Clermont-Ferrand_2    |
| AO13030    | <a href="#">39286</a>   | AO13030        | t: traite                 | BTH_Clermont-Ferrand_2    |
| AO13031    | <a href="#">AO13031</a> | AO13031        | fi: faible intrant        | BTH_Clermont-Ferrand_2    |
| ARKEOS     | <a href="#">36693</a>   | ARKEOS         | t: traite                 | BTH_Clermont-Ferrand_2    |
| DI13002-2  | <a href="#">39307</a>   | DI13002-2      | fi: faible intrant        | BTH_Dijon_2015_SetA1      |
| ARKEOS     | <a href="#">36693</a>   | ARKEOS         | t: traite                 | BTH_Dijon_2015_SetA1      |
| Bermude    | <a href="#">TX9</a>     | Bermude        | t: traite                 | BTH_Epi-centre_2015_SetA1 |
| AO13022    | <a href="#">AO13022</a> | AO13022        | nt: non traite fongicides | BTH_Epi-centre_2015_SetA1 |

[Ephesis data export](#) [Ephesis MIAPPE ISA-Tab export](#)

LEVEL: REPLICATION

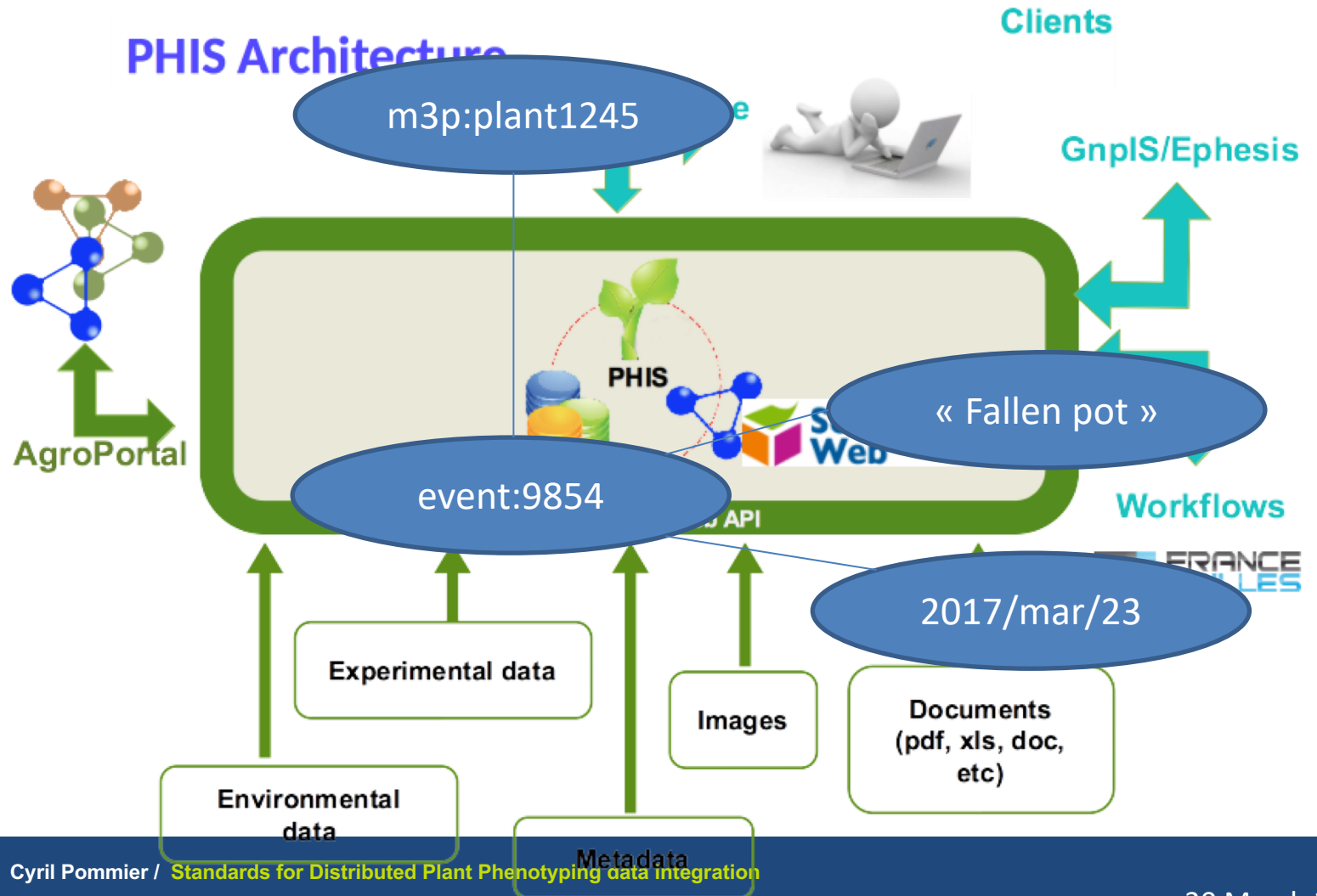
1-10 of 81,434 | Display 10 results per page

| Lot Number | Accession Number        | Accession Name | ltk                       | Trial Name                  |
|------------|-------------------------|----------------|---------------------------|-----------------------------|
| RE12080    | <a href="#">RE12080</a> | RE12080        | fi: faible intrant        | BTH_Clermont-Ferrand_2      |
| EM11537    | <a href="#">37829</a>   | EM11537        | fi: faible intrant        | BTH_Dijon_2015_SetA1        |
| DI13003    | <a href="#">39308</a>   | DI13003        | t: traite                 | BTH_Dijon_2015_SetA1        |
| AO13015    | <a href="#">AO13015</a> | AO13015        | t: traite                 | BTH_Dijon_2015_SetA1        |
| AO13031    | <a href="#">AO13031</a> | AO13031        | t: traite                 | BTH_Dijon_2015_SetA1        |
| AO13100    | <a href="#">AO13100</a> | AO13100        | nt: non traite fongicides | BTH_Epi-centre_2015_SetA1   |
| RUBISKO    | <a href="#">36680</a>   | RUBISKO        | nt: non traite fongicides | BTH_Epi-centre_2015_SetA1   |
| AO11011    | <a href="#">37247</a>   | AO11011        | fi: faible intrant        | BTH_Estrées-Mons_2015_SetA1 |
| DI13003    | <a href="#">39308</a>   | DI13003        | fi: faible intrant        | BTH_Estrées-Mons_2015_SetA1 |
| EM12096    | <a href="#">39321</a>   | EM12096        | fi: faible intrant        | BTH_Le_Moulon_2015_SetA1    |

[Ephesis data export](#) [Ephesis MIAPPE ISA-Tab export](#)

# MIAPPE databases & FAIR

- Reusable: PHS high tracability
- Annotation of Observation Units



# Example : Data set building with GnpIS-Ephesis

- GnpIS-Ephesis: INRA Information System, part of Phenome and Elixir
- François-Xavier Oury, Emmanuel Heumez, Bernard Rolland, Jérôme Auzanneau, Pierre Bérard, Maryse Brancourt-Hulmel, Xavier Charrier, Hubert Chiron, Camille Depatureaux, Laurent Falchetto, et al (2015) Winter wheat (*Triticum aestivum* L) phenotypic data from the multiannual, multilocal field trials of the INRA Small Grain Cereals Network. *doi: 10.15454/1.4489666216568333E12*
  - ◆ DOI : <http://dx.doi.org/10.15454/1.4489666216568333E12>
  - ◆ Winter wheat phenotypic French experimental network.
  - ◆ Observations : agronomic, quality, disease, phenology,...
  - ◆ Integration of 10 experimental locations, 15 years, 1700 genotypes.

→ 813 trials/study, 81 000 replications

# Wheat datasets publications

- Example of multiple datasets integration

- ◆ Shared locations
- ◆ Some common varieties

- Successive publications

- ◆ Initial publication: Agronomical breeding
- ◆ Quality
  - Breadmaking
  - Technological
- ◆ Additional Year: 20

**DATA SETS: 4** Reset

Genotype Variable Trial

**Traits, methods and scales**

Search terms...

- Wheat Inra Phenotype Ontology **ONTOLOGY**
  - Abiotic stress **TRAIT CLASS**
  - Agronomical **TRAIT CLASS**
  - Biotic stress **TRAIT CLASS**
  - Morphological **TRAIT CLASS**
  - Other **TRAIT CLASS**
  - Quality **TRAIT CLASS**
    - Alveograph Chopin **TRAIT**
    - Breadmaking test BIPEA (G) **TRAIT**
      - NmieG: Breadmaking test BIPEA: cr
      - NpainG: Breadmaking test BIPEA: b

**NmieG: Breadmaking test BIPEA: score** **VARIABLE**

|               |                         |
|---------------|-------------------------|
| Ontology name | Wheat Inra Phenotype O  |
| Identifier    | WIPO:0000238            |
| Name          | NmieG                   |
| Synonyms      | Breadmaking test BIPEA  |
| Institution   | INRA                    |
| Scientist     | Arnaud Gauffretau, Fran |
| Date          | 21/09/2017              |

Genotype Variable Trial

**Traits, methods and scales**

Search terms...

- Wheat Inra Phenotype Ontology **ONTOLOGY**
  - Abiotic stress **TRAIT CLASS**
  - Agronomical **TRAIT CLASS**
  - Biotic stress **TRAIT CLASS**
  - Morphological **TRAIT CLASS**
  - Other **TRAIT CLASS**
    - SDS sedimentation test **TRAIT**
      - sds: SDS sedimentation tes
    - Test weight **TRAIT**
    - Zeleny sedimentation test **TRAIT**

Collecting site  Evaluation site

2000 x 2001 x 2002 x 2003 x 2004 x 2005 x 2006 x  
2009 x 2010 x 2011 x 2012 x 2013 x 2014 x 2015 x

[remove all](#) [add all](#)

# Scientific Question

- Evolution of wheat mean yield for genebank accessions from 2003 to 2007, including disease interaction.

# Wheat Yield dataset building

- MIAPPE Plant Material Selection
- Optional Panel or Collections

## Phenotypes

Reset Form
Results

Genotype
Variable
Trial

Reset Tab

Add accession by Genus

↳ Choose a species

Add accession by Panel

Add accession by Collection

Add accession by Name

Paste your Accession Name list

### Accession

⏪
⏴
1-10 of 1,729
⏵
⏩

| Accession Number | Accession Name | Taxon Name                 |
|------------------|----------------|----------------------------|
| Alberic          | Alberic        | Triticum aestivum aestivum |
| 29843            | ALTIGO         | Triticum aestivum aestivum |
| 37771            | AO00001        | Triticum aestivum aestivum |
| AO01001          | AO01001        | Triticum aestivum aestivum |
| AO03001          | AO03001        | Triticum aestivum aestivum |
| AO03002          | AO03002        | Triticum aestivum aestivum |
| AO03003          | AO03003        | Triticum aestivum aestivum |
| AO03004          | AO03004        | Triticum aestivum aestivum |
| AO04001          | AO04001        | Triticum aestivum aestivum |
| AO05001          | AO05001        | Triticum aestivum aestivum |

Plant Material Export



# Wheat Yield dataset building

- MIAPPE investigation : Genebank/BRC accessions dataset selection

## Phenotypes

[Reset Form](#)

[Genotype](#)
[Variable](#)
[Trial](#)

[Reset Tab](#)

**Add trial by Name**

**Add trial by Number**

**Add trial by Campaign**

**Add trial by Site**

**Add trial by Data Set**

**Add trial by Project**

- INRA Small Grain Cereals Network
- INRA Wheat Network not BRC accession (B and C series)

---

**Selected Trials**

⏪
⏩
1-10 of 258
⏪
⏩

| Trial Number                      | Trial Name                        | Trial Site         |
|-----------------------------------|-----------------------------------|--------------------|
| BTH_Champagne-céréales_2004_SetA1 | BTH_Champagne-céréales_2004_SetA1 | Champagne-céréales |
| BTH_Champagne-céréales_2005_SetA1 | BTH_Champagne-céréales_2005_SetA1 | Champagne-céréales |

# Wheat Yield dataset building

- MIAPPE Assay/Observation variable:
  - ◆ Yield and stress variables selection
  - ◆ See BrAPPS

Reset Form Results

Genotype Variable Trial

rust

- Wheat Inra Phenotype Ontology **ONTOLOGY**
  - Biotic stress **TRAIT CLASS**
    - Disease score Yellow Rust* **TRAIT**
      - YR-Score\_score: Yellow rust score* **VARIABLE**
      - Susceptibility to leaf rust* **TRAIT**
        - rb: Susceptibility to leaf rust* **VARIABLE**

**YR-Score\_score: Yellow rust score** **VARIABLE**

|                      |  |
|----------------------|--|
| <b>Ontology name</b> | Wheat Inra Phenotype Ontology                      |
| <b>Identifier</b>    | WIPO:0000162                                       |
| <b>Name</b>          | YR-Score_score                                     |
| <b>Synonyms</b>      | Yellow rust score<br>Susceptibility to stripe rust |
| <b>Institution</b>   | INRA   |
| <b>Scientist</b>     | Jacques Le Gouis                                   |
| <b>Date</b>          | 15/06/2016   |
| <b>Crop</b>          | Wheat  |

**Disease score yellow rust** **TRAIT**

|                   |                           |
|-------------------|---------------------------|
| <b>Identifier</b> | WIPO:0010065              |
| <b>Name</b>       | Disease score Yellow Rust |
| <b>Class</b>      | Biotic stress             |

**1 to 9 note** **SCALE**

|                   |                                  |
|-------------------|----------------------------------|
| <b>Identifier</b> | WIPO:0030006                     |
| <b>Name</b>       | 1 to 9 note                      |
| <b>Data type</b>  | Ordinal                          |
| <b>Min</b>        | 1                                |
| <b>Max</b>        | 9                                |
| <b>Categories</b> | 1 = Resistant<br>9 = susceptible |

# Wheat Yield dataset building

- Easy Subsetting  
~ Third of 81 000 replications
- Reuse through MIAPPE ISA Tab File export

**Back to Form**

Search parameter(s):

Genus:   
 Number of Variable(s): 34

Geolocation

**DATA SETS: 7**

**Network Data Set :**  
[INRA Wheat Network BRC accession \(A series\)](#)

**Network Data Set :**  
[INRA Small Grain Cereals Network](#)  
 DOI:http://dx.doi.org/10.15454/1.4489666216568333E12

**Network Data Set :**  
[INRA Wheat Network not BRC accession \(B and C series\)](#)

Phenotyping campaign(s):

[remove all](#) [add all](#)

**Trial list** **Phenotypic data**

Data table view:

**LEVEL: REPLICATION**

1-10 of 26,030 replication | Display  results per page

| GENOTYPE ID |                         |                | TREATMENT               |   |                  |
|-------------|-------------------------|----------------|-------------------------|---|------------------|
| Lot Number  | Accession Number        | Accession Name | itk                     | Trial Name                                      | Trial Site       |
| CF99105     | <a href="#">20364</a>   | CF99105        | untreated, no fungicide | <a href="#">BTH Estrées-Mons 2003_SetA1</a>     | Estrées-Mons     |
| Recital     | <a href="#">6027</a>    | RECITAL        | low inputs              | <a href="#">BTH Le Moulon 2003_SetA1</a>        | Le Moulon        |
| Koreli      | <a href="#">29814</a>   | KORELI         | untreated, no fungicide | <a href="#">BTH Orgeval 2003_SetA1</a>          | Orgeval          |
| CF00108     | <a href="#">23825</a>   | CF00108        | untreated, no fungicide | <a href="#">BTH Rennes 2003_SetA1</a>           | Rennes           |
| Isengrain   | <a href="#">13433</a>   | ISENGRAIN      | treated                 | <a href="#">BTH Dijon 2003_SetA2</a>            | Dijon            |
| RE02101     | <a href="#">23821</a>   | RE02101        | treated                 | <a href="#">BTH Le Moulon 2003_SetA2</a>        | Le Moulon        |
| EM01288     | <a href="#">EM01288</a> | EM01288        | treated                 | <a href="#">BTH Lusignan 2003_SetA2</a>         | Lusignan         |
| DI03006     | <a href="#">DI03006</a> | DI03006        | treated                 | <a href="#">BTH Clermont-Ferrand 2003_SetB1</a> | Clermont-Ferrand |
| DI03001     | <a href="#">DI03001</a> | DI03001        | treated                 | <a href="#">BTH Dijon 2003_SetB1</a>            | Dijon            |
| Caphorn     | <a href="#">20443</a>   | CAPHORN        | untreated, no fungicide | <a href="#">BTH Orgeval 2003_SetB1</a>          | Orgeval          |

Ephesis data export Ephesis MIAPPE ISA-Tab export

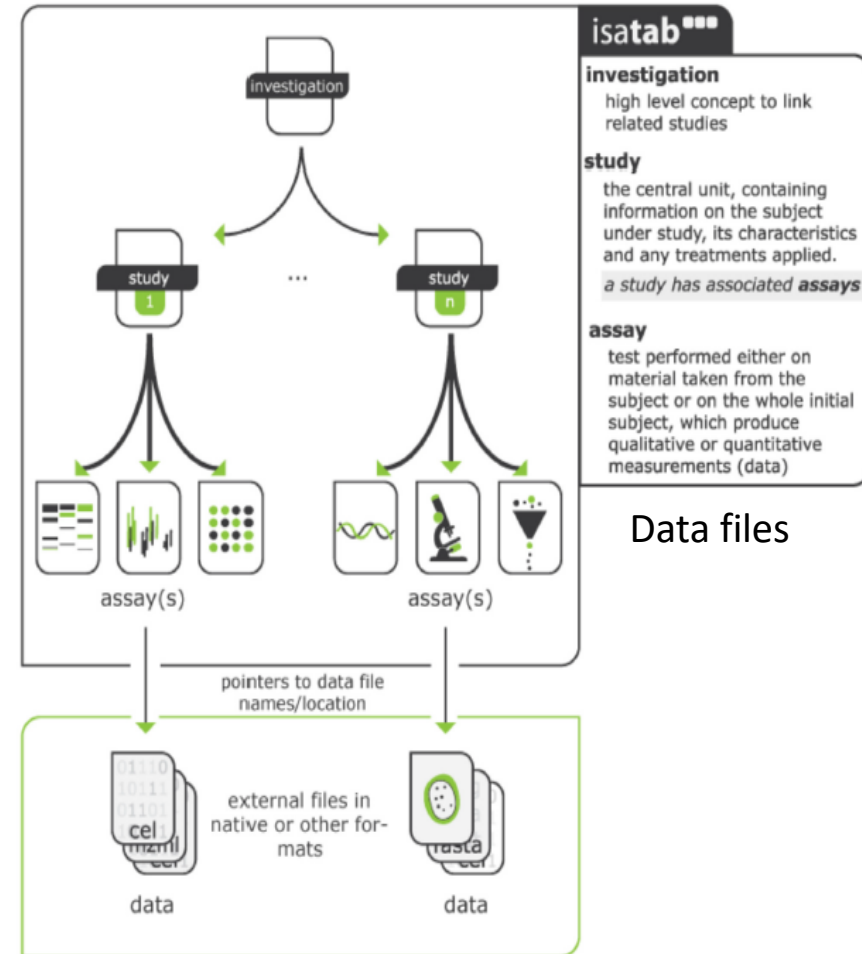
## USE CASE: FILE ARCHIVE



MIAPPE ISA Tab

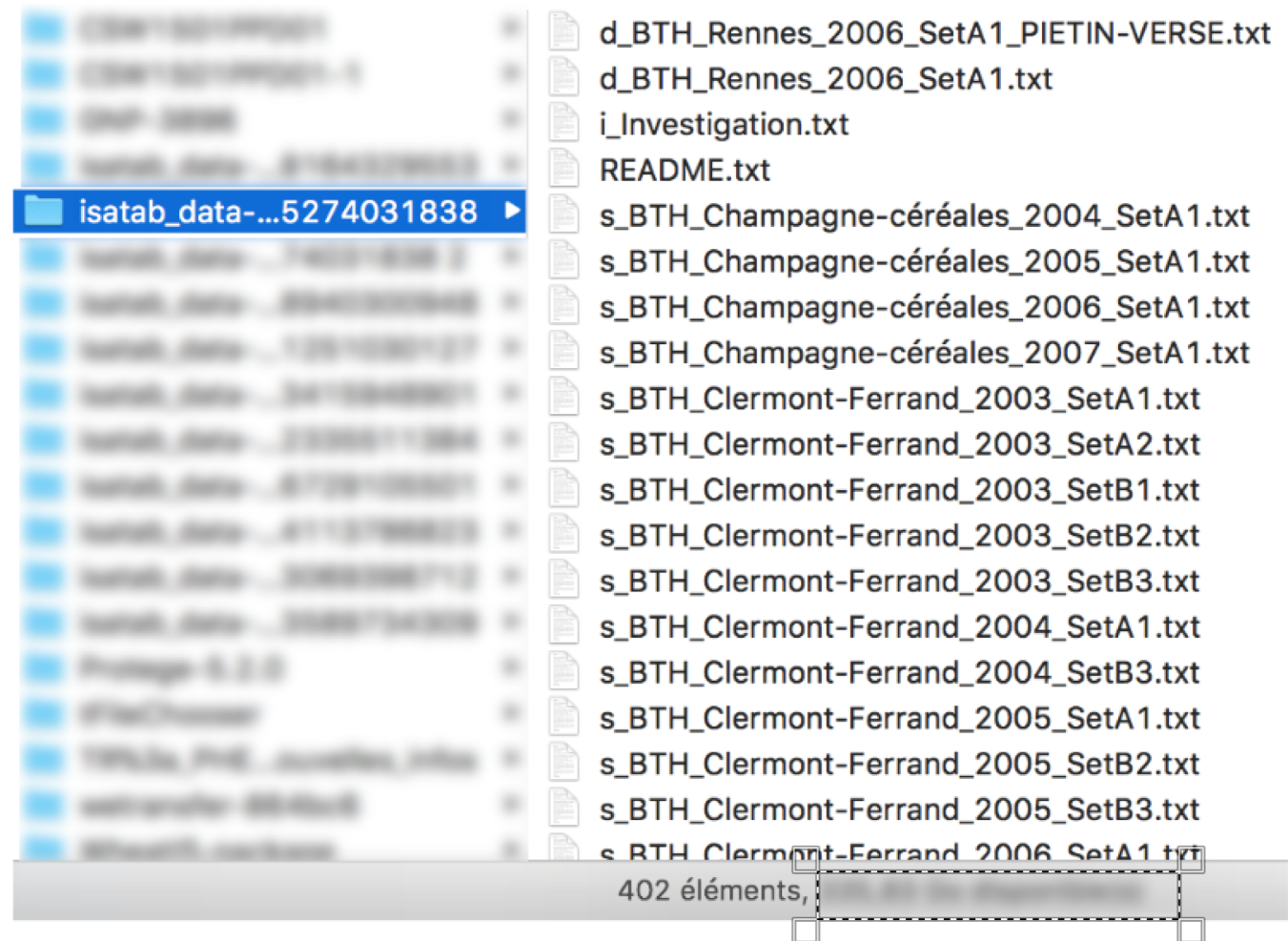
# MIAPPE File Archive

- 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
    - Frederik Coppens & H. Parkinson & M. Courtot + Elixir-fr-nl, ISA Framework, ...



# MIAPPE File: Wheat Yield dataset

- Zip archive
- 402 files
- README
- i\_Investigation.txt :  
initial file, bootstrap
- **MIAPPE v1.0**
  - ◆ Not MIAPPE V1.1



# MIAPPE File

- Derived data file : d\_BTH\_Rennes\_2003\_SetA1.txt

| Observation Unit                                       | A | B               | C                 | D                | E                  | F                  | G                 | H                 |
|--|---|-----------------|-------------------|------------------|--------------------|--------------------|-------------------|-------------------|
| Assay Name   |   | Trait Value [r] | Trait Value [rdt] | Trait Value [ht] | Trait Value [prec] | Trait Value [prot] | Trait Value [oid] | Trait Value [knt] |
| Triticum_aestivum_aestivum-CF00108-CF00108-itk:t-2003  |   | 96,6            | 96,4              | 80               | 138                | 10,2               | 1                 |                   |
| Triticum_aestivum_aestivum-AO03001-AO03001-itk:nt-2003 |   | 103,3           | 86,2              | 93               | 141                |                    | 1                 |                   |

- Metadata files

| Sample Name                                    | Measurement Factor Value [itk] | Assay Name   | Characteristics [knt] |
|--|--------------------------------|--|-----------------------|
| Triticum_aestivum_aestivum-FARANDOLE-Farandole | untreated, no fungicide        | Triticum_aestivum_aestivum-FARANDOLE-Farandole-itk:nt-2003 | Farandoleitknt        |
| Triticum_aestivum_aestivum-Ressor-Ressor       | low inputs                     | Triticum_aestivum_aestivum-Ressor-Ressor-itk:fi-2003       | Ressoritkfi           |

| A                          | B                          | C                                | D                                  | E                                  |
|----------------------------|----------------------------|----------------------------------|------------------------------------|------------------------------------|
| Source Name                | Characteristics [Organism] | Characteristics [Infra-specific] | Characteristics [accession number] | Characteristics [accession number] |
| Triticum_aestivum_aestivum | Triticum_aestivum_aestivum | Orvantis                         | 14088                              | ORVANTIS                           |
| Triticum_aestivum_aestivum | Triticum_aestivum_aestivum | Recital                          | 6027                               | RECITAL                            |

- Raw data file: analysis input

| Accession Name [itk] | Trial Name              | Trial Site | X(m)   | Y(m) | Campaign | GDD | prot: Grain prod [prot] | Date [prot] | rdt: Grain yield [rdt] | Date [rdt] | sept: Sept [knt] |
|----------------------|-------------------------|------------|--------|------|----------|-----|-------------------------|-------------|------------------------|------------|------------------|
| RECITAL              | treated                 | BTH_Rennes | Rennes |      | 2003     |     | 10,4                    |             | 88,4                   |            |                  |
| TAMARO               | untreated, no fungicide | BTH_Rennes | Rennes |      | 2003     |     |                         |             | 60,7                   |            |                  |
| DI01022              | untreated, no fungicide | BTH_Rennes | Rennes |      | 2003     |     |                         |             | 95,2                   |            |                  |

## USE CASE: WEB SERVICES





- Breeding API
- International collaboration
  - ◆ Excellence in Breeding platform (CGIAR)
  - ◆ Coordinator : Peter Selby
  - ◆ Lead: Lukas Mueller, Jan Erik Backlund, Kelly Robbins
- Vision :
  - ◆ Standard Open API
  - ◆ Easy, secure, and efficient
  - ◆ Information Exchange
  - ◆ Between systems and applications
  - ◆ Breeding is the main target

Bill & Melinda Gates Foundation

CassavaBase

T3

IBP

JHI

Bioversity

CIRAD

INRA

IRRI

GOBII

Wageningen

CIP

DaRT

Cornell

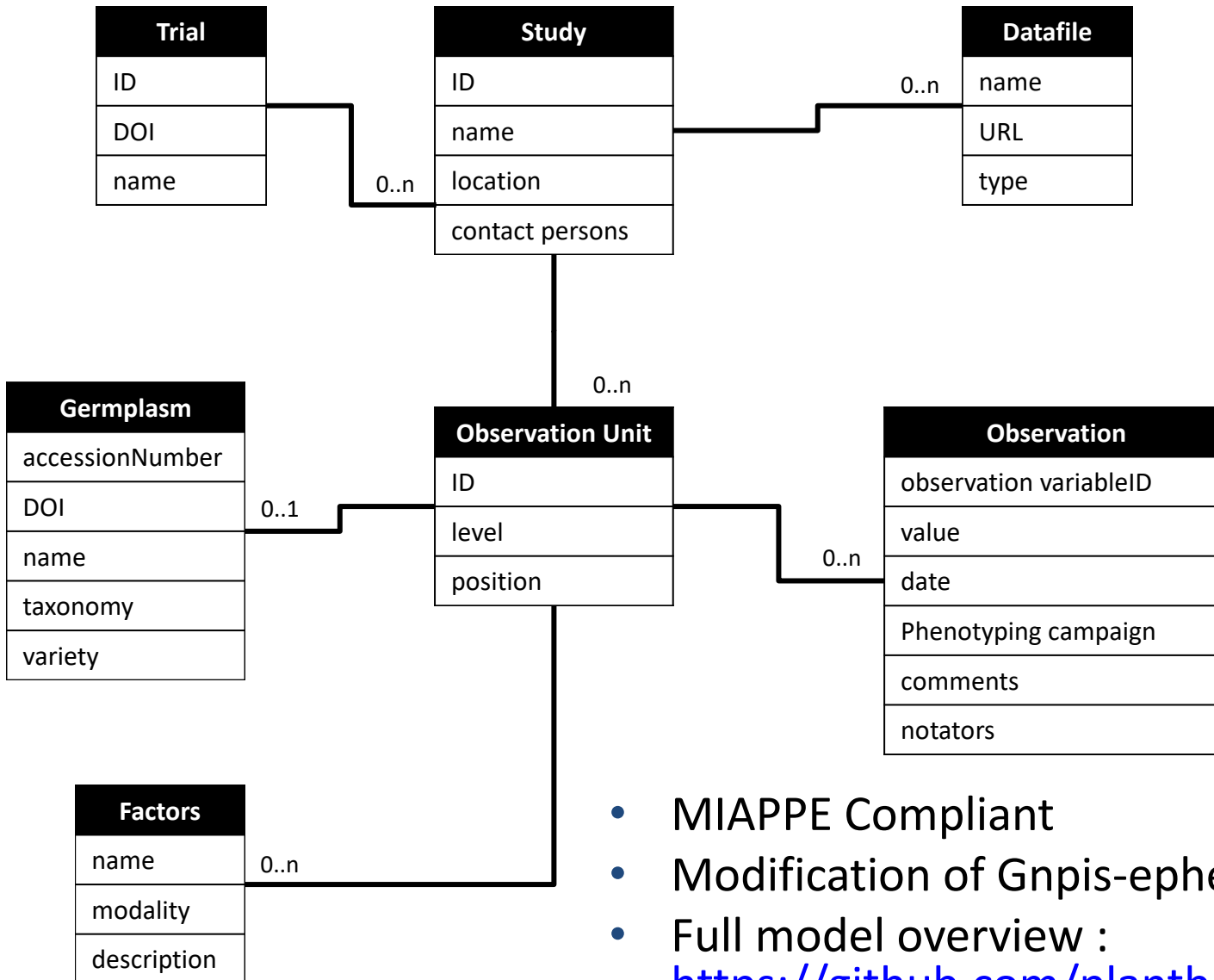
iPlant

BILL & MELINDA  
GATES foundation



- Resources
  - ◆ <http://brapi.org/>
  - ◆ Specifications
    - <https://brapi.docs.apiary.io/#reference>
    - <https://github.com/orgs/plantbreeding>
- Servers implementations
  - ◆ CGIARs international network
    - Integrated Breeding Platform, Breeding Management System
  - ◆ Elixir Excelerate
  - ◆ Emphasis: PHIS
  - ◆ Germinate
  - ◆ GnpIS
- Clients implementations
  - ◆ Flapjack : genotyping data visualization
  - ◆ Phenome-fppn portal, GnpIS
  - ◆ R analysis pipelines
- BrAPPS
  - <https://www.brapi.org/brapps.php>
  - ◆ Ontology Widget
    - <https://github.com/gnpis/trait-ontology-widget>

# Implicit Phenotyping Model overview

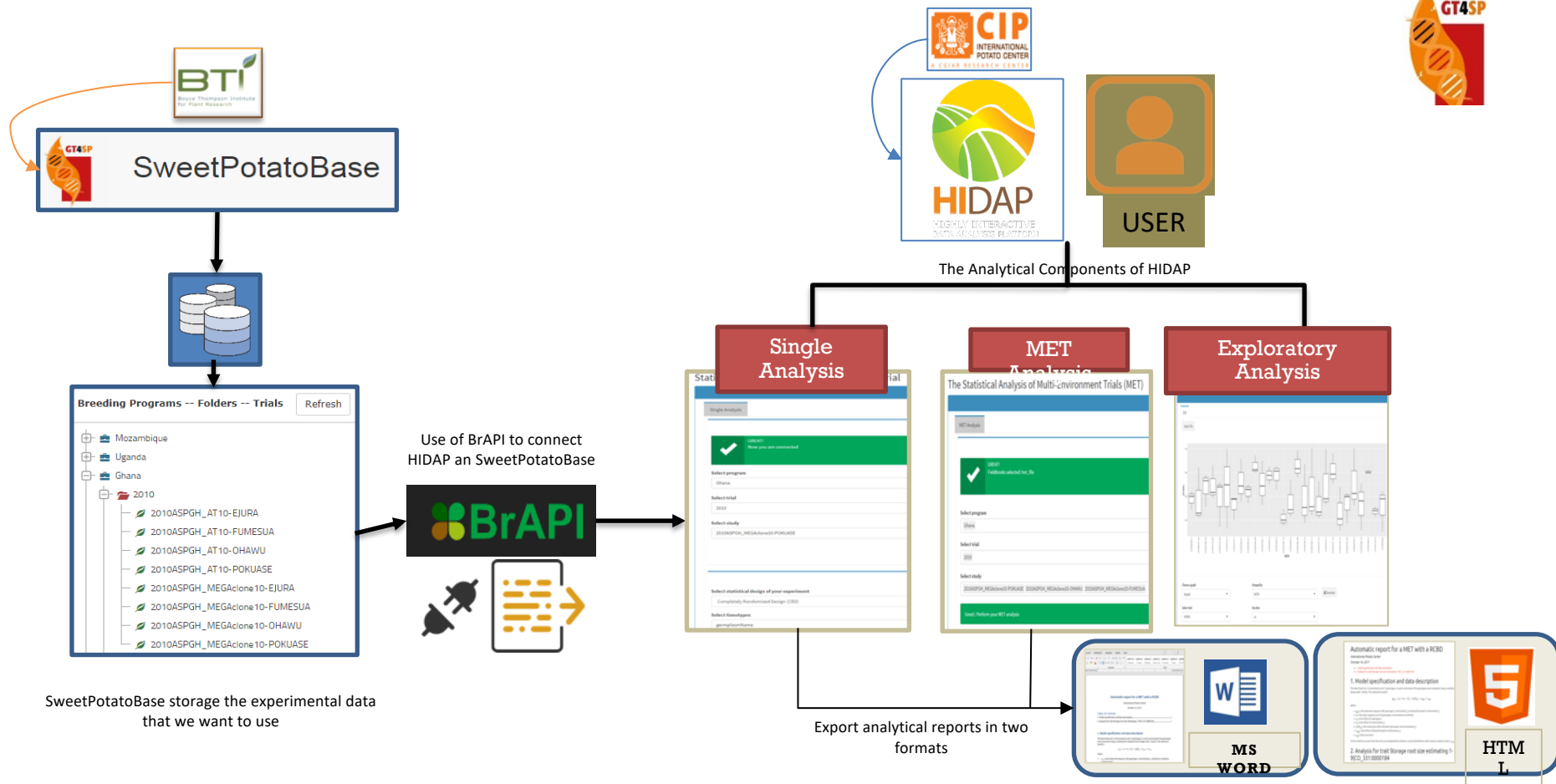


- MIAPPE Compliant
- Modification of Gnpis-ephephesis publication
- Full model overview : <https://github.com/plantbreeding>

# USE case WS1: BrAPI for R analysis



## Workflow HIDAP-SweetPotatoBase



[Brapi/v1/study/table](http://Brapi/v1/study/table) web service → ISA Raw data file

# R brapi libraries: on github

Hub, Inc. [US] | <https://github.com/CIP-RIU/brapi>



README.md

## BrAPI R package

build passing build passing coverage 51%

This version is still under development. The implementation sometimes changes minor details.

## README

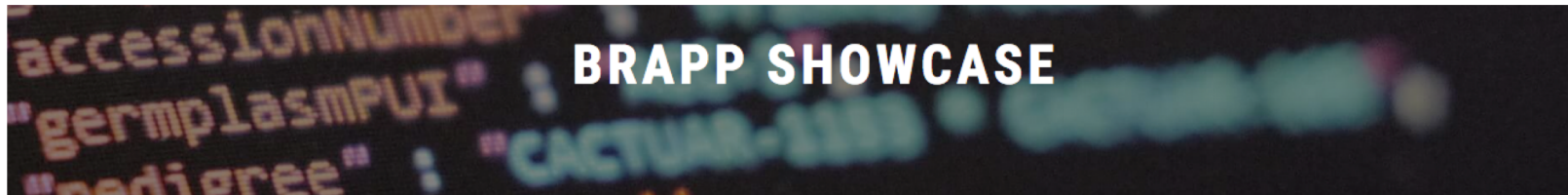
An R package to use the [Breeding API \(BrAPI\)](#) for accessing plant breeding data.

It can be installed using:

```
install.packages("devtools")  
devtools::install_github("CIP-RIU/brapi")
```

# USE case WS2: BrAPPS

- Tools integrable in any BrAPI compliant System
  - ◆ <https://www.brapi.org/brapps.php>



Welcome to the BrAPI Application Showcase, or BrAPPS for short. Each of these BrAPPS is an independent tool which can run on its own or can be easily integrated with a larger system. And of course, each BrAPP is completely BrAPI driven and can be used with any BrAPI compliant system. The goal is to keep the BrAPI community informed about what tools are available and encourage functionality sharing. If you have a BrAPP you would like to share with the community please reach out to a [BrAPI Coordinator](#)



## GRAPHICAL FILTERING TOOL

FROM [CASSAVABASE](#)

DEMO SOURCE DOCS

Built by CassavaBase, this is a fully BrAPI compliant tool for graphically building complex search and filter queries of Observational Data.



## VARIABLE ONTOLOGY WIDGET

FROM [URGI - INRA](#)

DEMO SOURCE DOCS

Built by URGI - INRA, this is a JavaScript widget displaying all the observation variables of a BrAPI endpoint.



## STUDY COMPARISON TOOL

FROM [CASSAVABASE](#)

DEMO SOURCE DOCS

Built by CassavaBase, this is a fully BrAPI compliant tool for comparing the performance of germplasm across studies.



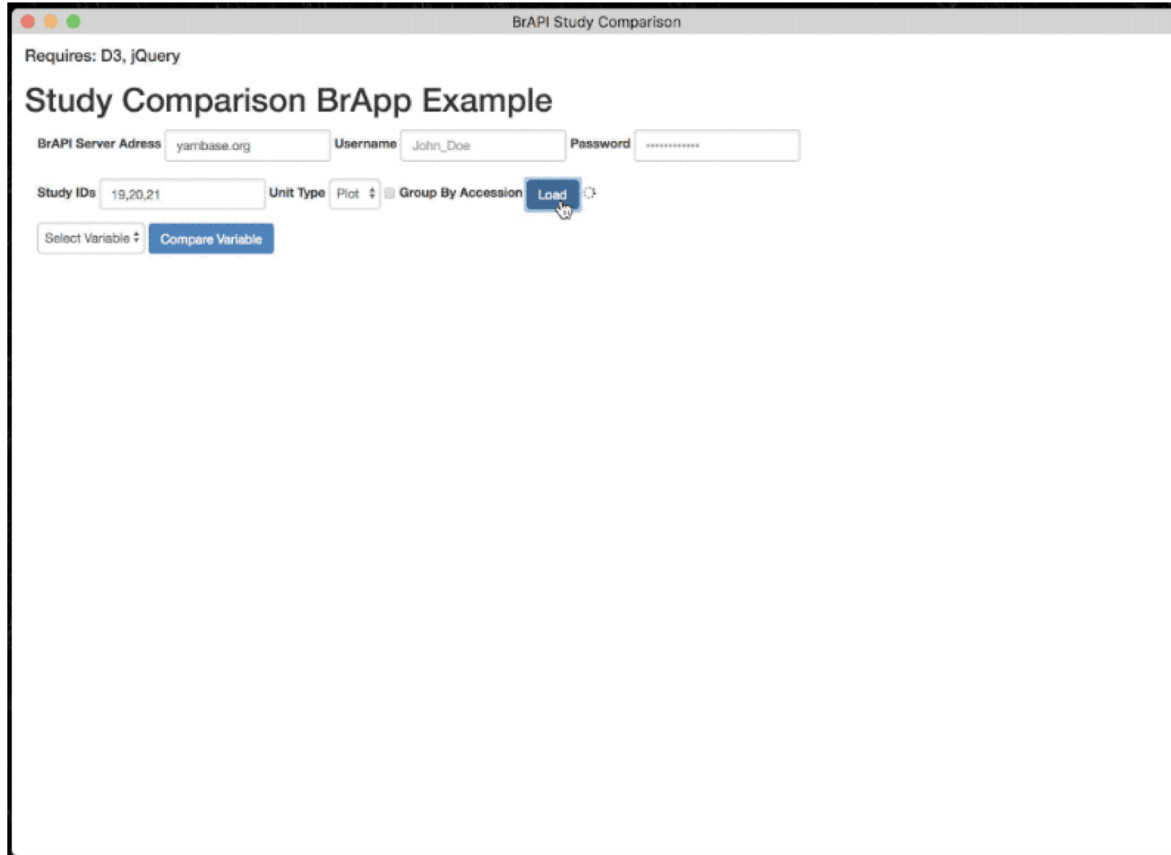
## PEDIGREE VIEWER

FROM [CASSAVABASE](#)

DEMO SOURCE DOCS

Built by CassavaBase, this is a fully BrAPI compliant tool for viewing and exploring germplasm pedigrees.

- Study Comparison tool



Requires: D3, jQuery

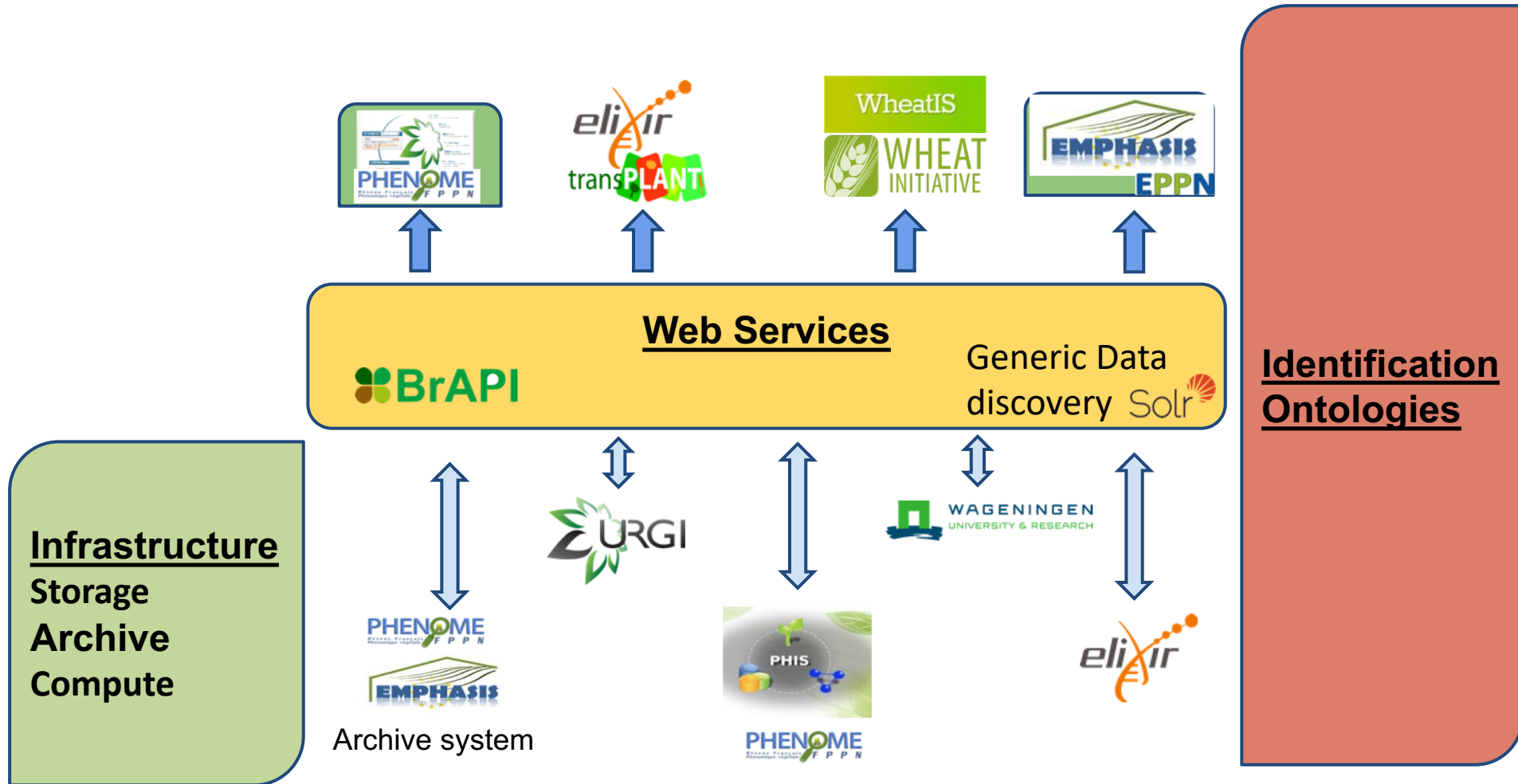
## Study Comparison BrApp Example

BrAPI Server Address:  Username:  Password:

Study IDs:  Unit Type:  Group By Accession:

Select Variable:

# USE case WS3: Distributed Information System





# Generic data discovery

- Full text
- Facets

WheatIS

Filters

[Clear](#)

**Database**

- TRITICEAE TOOLBOX (64)
- CR-EST (7)
- GNPIS (3)
- ENSEMBL PLANTS (1)
- GNPIS JBROWSE (1)
- PLANTPHENODB (1)

**Type**

- ACCESSION (42)
- PHENOTYPE (14)
- EXPERIMENT (9)
- EXPRESSED SEQUENCE TAGS (7)
- SEQUENCE FEATURE (2)
- PHENOTYPE (1)
- QTL (1)
- SEQUENCE FEATURE (1)

**Species**

- TRITICUM AESTIVUM (69)
- HORDEUM VULGARE (6)
- TRITICUM AESTIVUM L. (1)
- TRITICUM DURUM (1)

**Search**

---

**About**

**WheatIS nodes:**

transPlant-MIPS (UP):

- CrowsNest: 13324

transPlant-IPK (UP):

- CR-EST: 199220
- GEBIS: 52878
- MetaCrop: 355

fhb

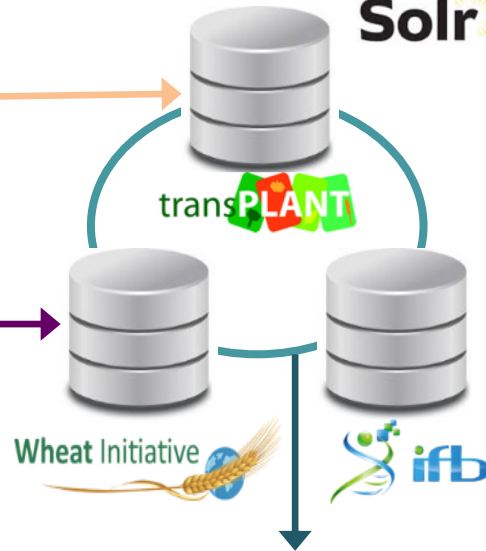
1-10 of 77
10 results per page

| ID                                    | Source            | Type       | Taxon             | Description   |
|---------------------------------------|-------------------|------------|-------------------|---|
| <a href="#">Traes_5DL_E12C501B4</a>   | Ensembl Plants    | -          | Triticum aestivum | Sequence feature, Ensembl Plants, Traes_5DL_E12C501B4, Traes_5DL_Phylla1 [Source:UniProtKB/TrEMBL;Acc:A0FHB0], Triticum aestivum, prot          |
| <a href="#">HDP14M22T</a>             | CR-EST            | -          | Hordeum vulgare   | HDP14M22T, expressed sequence tags, CR-EST, Hordeum vulgare, gij26 coli CFT073] Hypothetical protei; gi 28951047 gb AAO63447.1  At2g37930       |
| <a href="#">HDP20D01w</a>             | CR-EST            | -          | Hordeum vulgare   | HDP20D01w, expressed sequence tags, CR-EST, Hordeum vulgare, gij15 [Rickettsia prowazekii str. Madrid E] SFHB ; gi 34906406 ref NP_914550.1     |
| <a href="#">HDP20D01T</a>             | CR-EST            | -          | Hordeum vulgare   | HDP20D01T, expressed sequence tags, CR-EST, Hordeum vulgare, gij15 [Rickettsia prowazekii str. Madrid E] SFHB ; gi 34906406 ref NP_914550.1     |
| <a href="#">HDP21C08T</a>             | CR-EST            | -          | Hordeum vulgare   | HDP21C08T, expressed sequence tags, CR-EST, Hordeum vulgare, gij15 [Rickettsia prowazekii str. Madrid E] SFHB ; gi 31979237 gb AAP68831.1       |
| <a href="#">HDP31N10w</a>             | CR-EST            | -          | Hordeum vulgare   | HDP31N10w, expressed sequence tags, CR-EST, Hordeum vulgare, gij15 [Rickettsia prowazekii str. Madrid E] SFHB ; gi 34906406 ref NP_914550.1     |
| <a href="#">HDP35A10T</a>             | CR-EST            | -          | Hordeum vulgare   | HDP35A10T, expressed sequence tags, CR-EST, Hordeum vulgare, gij26 coli CFT073] Hypothetical protei; gi 28951047 gb AAO63447.1  At2g37930       |
| <a href="#">TS034O07u</a>             | CR-EST            | -          | Triticum aestivum | TS034O07u, expressed sequence tags, CR-EST, Triticum aestivum, Gij15 thaliana] gi 7487460 pir T01820 hypo; Gi 15604676 ref NP_221194.1  SFH     |
| <a href="#">HWW FHB</a>               | Triticeae Toolbox | Experiment | Triticum aestivum | Experiment, Triticeae Toolbox, HWW FHB, phenotype experiment, Include HWWFHB_2014_Fargo, Triticum aestivum, phenotype                           |
| <a href="#">URSN_2012_BrookingsSD</a> | Triticeae Toolbox | Experiment | Triticum aestivum | Experiment, Triticeae Toolbox, URSN_2012_BrookingsSD, phenotype trial severity, Fusarium head blight disease index, visually scabby kernels URS |

# Generic data discovery



EnsemblPlants  
 EMBL-EBI  
 HELMHOLTZ GEMEINSCHAFT  
 PLANTS DB  
 IPK GATERSLEBEN  
 IGR pan  
 ZURGI  
  
 TGAC The Genome Analysis Centre™  
 CerealsDB  
 GrainGenes A Database for Triticeae and Avena  
 ROTHAMSTED RESEARCH  
  
 Gramene  
 CIMMYT International Maize and Wheat Improvement Center  
 UCW  
  
 THE UNIVERSITY OF WESTERN AUSTRALIA  
 USDA | NIFA | Triticaceae CAP



ZURGI  
 South Green bioinformatics platform  
 PHENOME Réseau Français Phénomique végétale F P P N  
 arcad

WheatIS

transPLANT

Search: About

terpene

More restrictions

Show 10 entries

| ID  | Source         | Type | Taxon           | Description   |
|-----|----------------|------|-----------------|---|
| 101 | Ensembl Plants | -    | Triticum urartu | TRUES38F0911001TCFD_11 is a mRNA-GDEC of Triticum aestivum located between positions 87822747 and 87827718 on trites3pseudomolecule1 and which properties are: loc1_01TAE358F0911001TCFD_11; OGDW443_0111_AUGUSTUS_wheat_060306_06A250_gene_512_mRNA_001; Non-redundant; function - 8802173; MAIZE_TREMBL; database Terepene synthase 7 OS Zea mays PE 2 BV 1; locus agw443_0111_AUGUSTUS_wheat_51; OGDW443_0111_AUGUSTUS_wheat_060306_06A250_gene_012; bestblasthit685V73_MAZE_TREMBL; database Terepene synthase 7 - 1220c; 74.23 - hour: 08:52 - gene: 06.33; blast-evalue: Confidence; function: terpenoid synthase; gene: 060306_06A250; locus: agw443_0111_AUGUSTUS_wheat_060306_06A250_gene_512_mRNA_001; CDS: 8bp; function: identity:74.23; parent: loc1_01TAE358F0911001TCFD_11; function: coverage:98.02 |
| 102 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 103 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 104 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 105 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 106 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 107 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 108 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 109 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |
| 110 | CR-EST         | -    | Hordeum vulgare | Hordeum vulgare   |

Showing 1 to 10 of 87 entries

rubrique < appui & projets > ci-destim

APPUI & PROJETS

PARTENAIRES

EVENEMENTS

CONTACT

SUPPORT REQUEST

ALL FOR PROJETS

GROUP

LOUD

OFFERS

SOURCES

**Data types**

Phenotypes  
 Markers,  
 accessions,  
 GWAS,  
 Experiments  
 Journal articles  
 Seeds of  
 discovery data  
 Genomic  
 Features

# Generic Search Growth

## Indexing Package

- Community Building
- Indexing node deployment support
- <http://wheatis.org/Tools.php>
- <http://wheat-urgi.versailles.inra.fr/Projects/Wheat-Information-System/SolR-tool-package>

To ease the WheatIS nodes to expose their data to the [WheatIS search](#) , we developed a package to help the installation and configuration of the SolR tool.

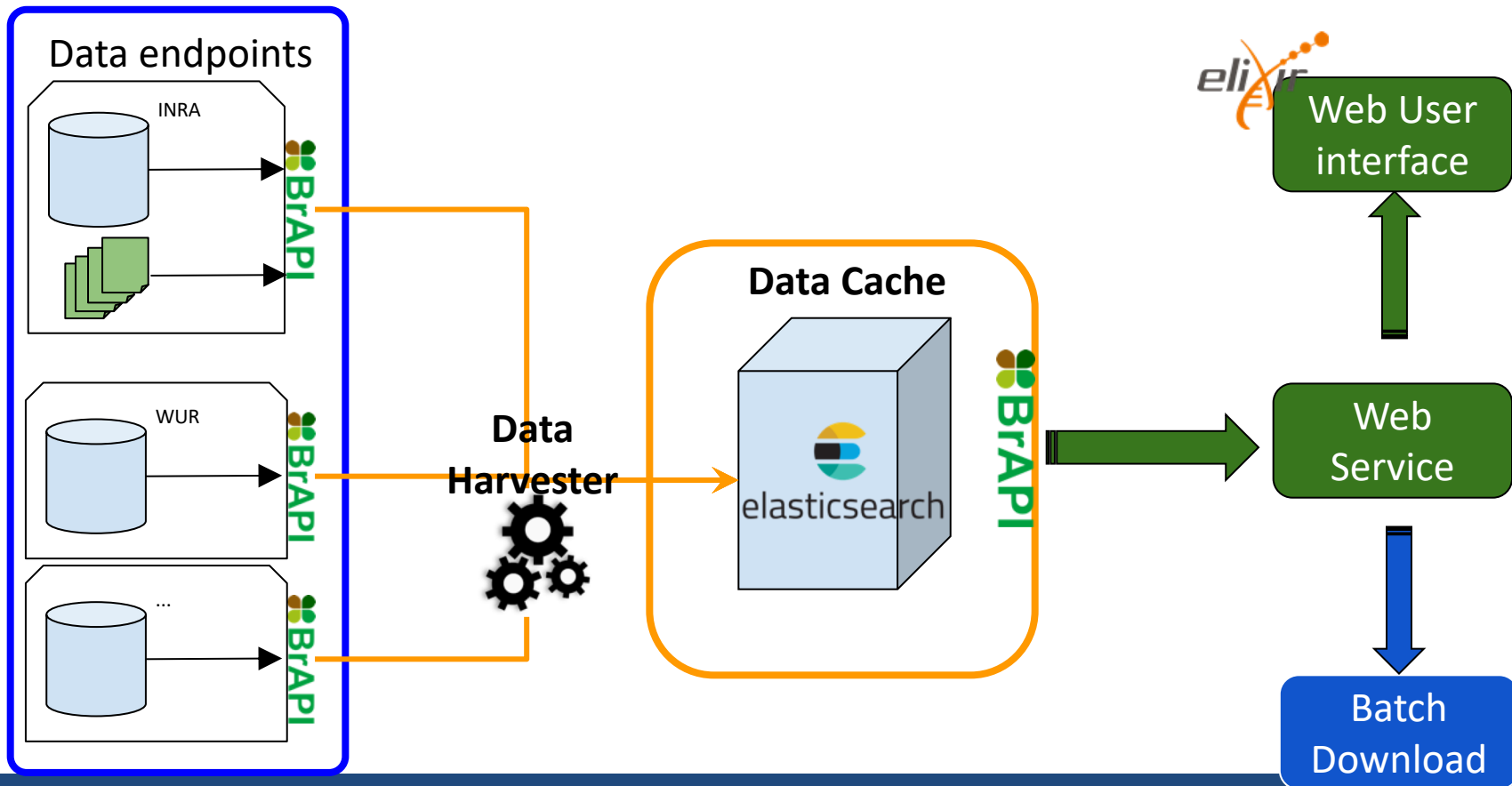
Download the package: [SolR-tool-package-WheatIS.tar.gz](#) (157.33 kB)

Please [contact us](#) to have support and to add your SolR indexes to the WheatIS search.



# Elixir Plant Data Lookup Service

- Phenotype & Study
- Integrated in Elixir collaboration
- Complementarity and convergence with WheatIS datadiscovery network
- Central Index
- GnpIS-Ephesis like data integration over databases network



## USE CASE: MIAPPE ONTOLOGY ( PPEO )



Plant Phenotyping Experiment Ontology

# MIAPPE Semantic: PPEO

- Joint initiative : Elixir (Fr, Nl, Pt), Emphasis, CropOntology, RDA, INRA
- MIAPPE OWL Ontology
  - ◆ First draft
  - ◆ <https://github.com/MIAPPE/MIAPPE-ontology>
- BrAPI 2 MIAPPE RDF workflow
  - ◆ <http://www.wheatis.org> data standard page
  - ◆ <http://ist.blogs.inra.fr/wdi/phenotypes-as-rdf/>
- Agroportal
  - ◆ Agronomy bioportal
  - ◆ <http://agroportal.lirmm.fr/ontologies/PPEO>
- Wheat dataset
  - ◆ <http://dx.doi.org/10.15454/1.4489666216568333E12>
  - ◆ Future Query: Impact of summer temperature on yield

**Wheat Data Interoperability Guidelines**

Home Guidelines Ontologies & Vocabularies Use cases Getting involved About

Home / Phenotypes as RDF

## Phenotypes as RDF

### Phenotypes to Semantic web publication

The sharing and especially the reusability of complex data like phenotypes is a good use case for a publication following the semantic web principles. This page describes how to generate a Phenotype RDF dataset from a [Breeding API](#) MIAPPE compliant endpoint. The growing adoption of this API will facilitate the reusability of such a transformation process.

### Recommendations

The RDF publication relies among other things on the correct identification of different resources through Permanent

**Virtuoso SPARQL Query Editor**

Default Data Set Name (Graph IRI)

Query Text

```
PREFIX brapi: <https://brapi.org/rdf/>
SELECT ?network ?study ?attribute ?value
FROM <urn:urgi:pheno-brapi-inra-small-grain-cereals-network>
WHERE {?study a brapi:Study;
        ?attribute ?value.
        ?trial a brapi:Trial;
        brapi:hasName ?network.}
ORDER BY ?network ?study
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format:

Execution timeout:  milliseconds (values less than 1

Options:  Strict checking of void variables  Log debug info

(The result can only be sent back to browser, not saved on the server, see [details](#))

Run Query Reset

## MIAPPE STATUS AND PERSPECTIVES



# Adoption

- Plant community involved
  - ◆ Elixir (European bioinformatic infrastructure)
  - ◆ Emphasis (European Phenotyping infrastructure)
  - ◆ Bioversity international CGIAR
- Breeding API is Elixir official Phenotyping standard web service
- MIAPPE and BrAPI high collaboration
  - ◆ adoption and compliance
- Data repositories and management tools
  - ◆ GnpIS <https://urgi.versailles.inra.fr/gnpis/>
  - ◆ eDale <https://edal.ipk-gatersleben.de/>
  - ◆ PlantPhenoDB at IPGPAS <http://cropnet.pl/plantphenodb/>
  - ◆ In progress: COPO, Elixir plant databases (iBet, WUR, VIB, ...), Brassica Information Portal, ...





# Perspectives

- Version 2 : Emphasis, Elixir, ...
  - ◆ Environment
  - ◆ Sensor traceability
  - ◆ Phenoharmonis Workshop Montpellier may 2018
- Elixir data lookup service, MIAPPE enabled.
  - ◆ Data discovery
  - ◆ Elixir
  - ◆ WheatIS & Emphasis ?
  - ◆ Open source software
- Dataset Validation
  - ◆ Elixir
  - ◆ File archive (ISA Tab) and BrAPI based



# Perspectives: Emphasis Ontology & MIAPPE

- Emphasis Ontology in construction
  - ◆ High level organization of phenotyping infrastructures
  - ◆ Complementary of MIAPPE
  - ◆ For instance: Installation  $\approx$  Assay
    - elementary level for data acquisition
    - Set of sensor + vectors

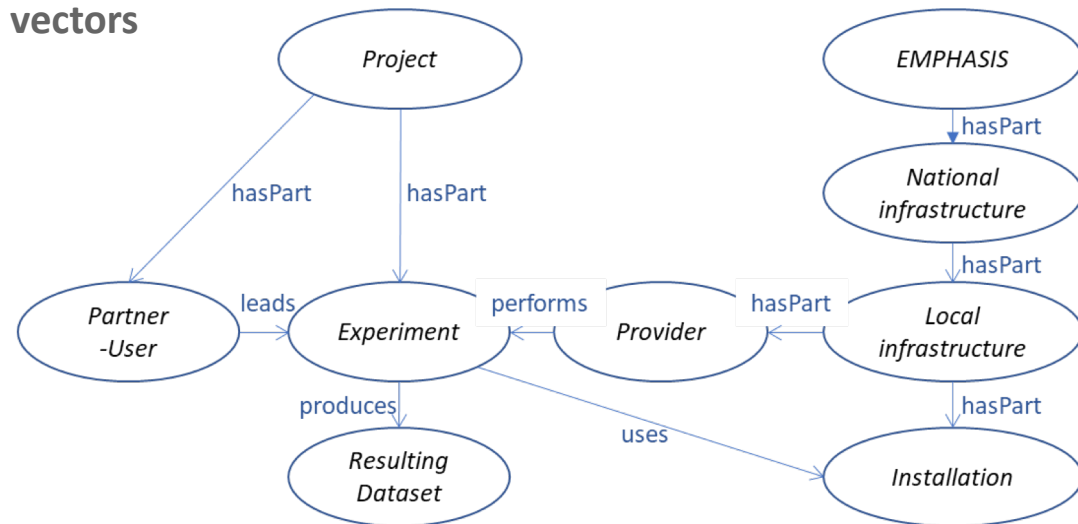


Fig. 1. Main terms used in this document and their functional relationships. Each instance has its own attributes, not represented here for better legibility. *Provider*: group of people involved in a *local infrastructure*, usually involving subgroups at *installation* level. *Project*: any project, (e.g. UE or national) that involves *partners* that perform experiments, themselves part of the *project*. The terms *Partner* (in a project) and *User* (in a phenotyping infrastructure) map to each other. All these terms are implemented in the information system.

# Plant Phenotyping Data & FAIR

- FAIR allows handling big data on the long term (decades)
  - ◆ **Findability**
    - in a lot of datasets
    - With different approaches
    - Without relying on any Archive logic/Folder mechanisms
  - ◆ **Accessibility**
    - Avoid storing high volumes for nothing
  - ◆ **Interoperability**
    - Processing lot of data
    - Aggregating several datasets into new ones
  - ◆ **Reusability**
    - Understandability of the data
    - Reducing data format heterogeneity



# Acknowledgment & Questions



- ◆ Paul Kersey



- Bioversity international CGIAR
  - ◆ Elizabeth Arnaud
  - ◆ Marie Angélique Laporte

IPG PAS



- ◆ Hanna Cwiek-Kupczynska
- ◆ Pawel Krajewski



- ◆ Matthias Lange



- ◆ Cyril Pommier
- ◆ Anne Françoise Adam Blondon
- ◆ Guillaume Cornut
- ◆ Thomas Letellier
- ◆ Célia Michotey
- ◆ François Tardieu
- ◆ Pascal Neveu
- ◆ Manuel Ruiz
- ◆ Pierre Larmande



- ◆ Frederik Coppens



iBet

- ◆ Bruno Costa
- ◆ Inês Chaves
- ◆ Célia M. Miguel



WAGENINGEN  
UNIVERSITY & RESEARCH

- ◆ Richard Finkers

IGC

- ◆ Daniel Faria



- ◆ Bjorn Usadel