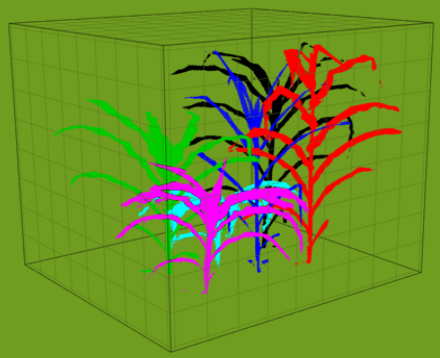
Phenome-Emphasis-Elixir: phenotyping data management and integration

Anne-Francoise Adam-Blondon, Cyril Pommier, Llorenç Cabrera Bosquet

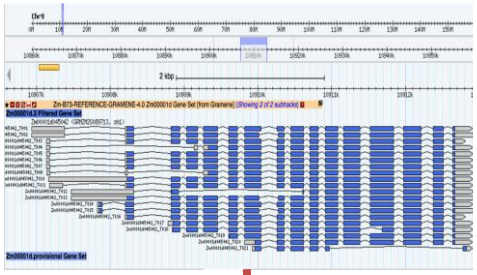


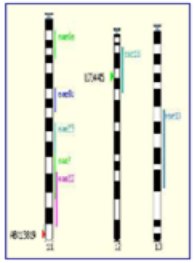
Phenome-Emphasis-Elixir: phenotyping data management and integration

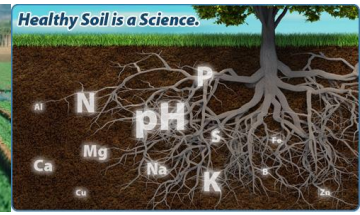
A-F Adam-Blondon, C Pommier, L Cabrera-Bosquet

Grape Genetics and Breeding Symposium, Workshop on grapevine phenotyping, France, Bordeaux, 2018 July 16-20th

Genomic X Environment X Phenomic







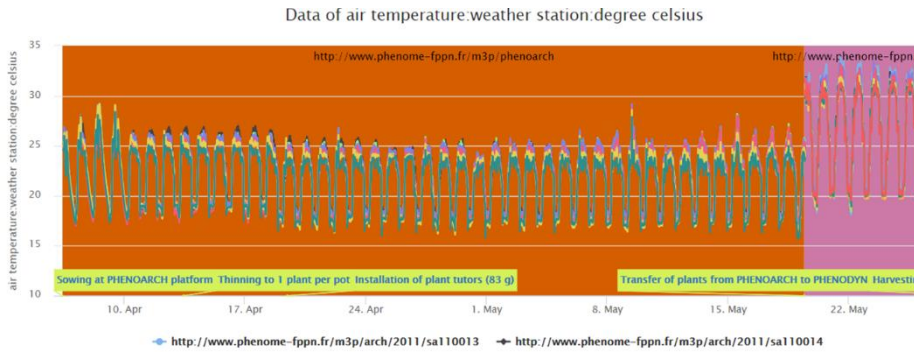
Findable **A**ccessible **I**nteroperable **R**eusable

- Climate change adaptation
- Variety improvement
- Cultural Practices adaptations
- …

Phenomic multidimensional data

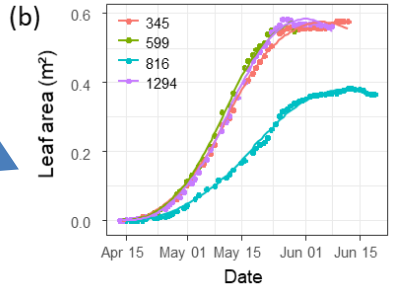
PHENOME Infrastructure Française de Phénomique végétale EMPHASIS

PHIS



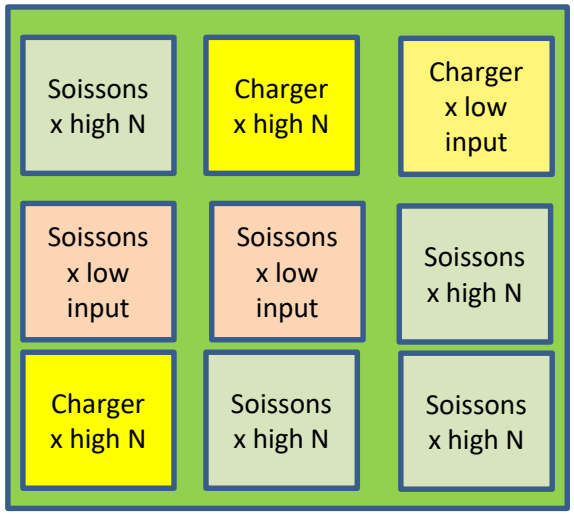
Soissons x high N	Charger x high N	Charger x low input
Soissons x low input	Soissons x low input	Soissons x high N
Charger x high N	Soissons x high N	Soissons x high N

Genotype	traitement	Fusariose
Soissons	low input	5
Soissons	high N	7

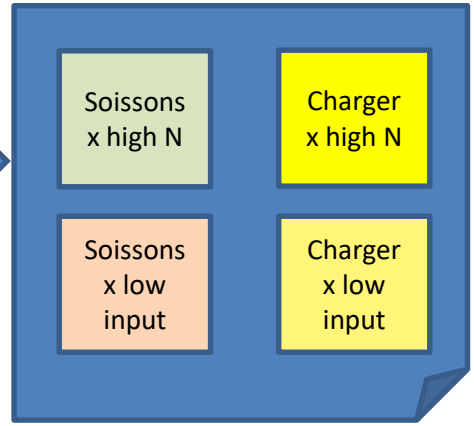


Phenotyping Data in matrix, link with other types of data

Raw data



Elaborated data

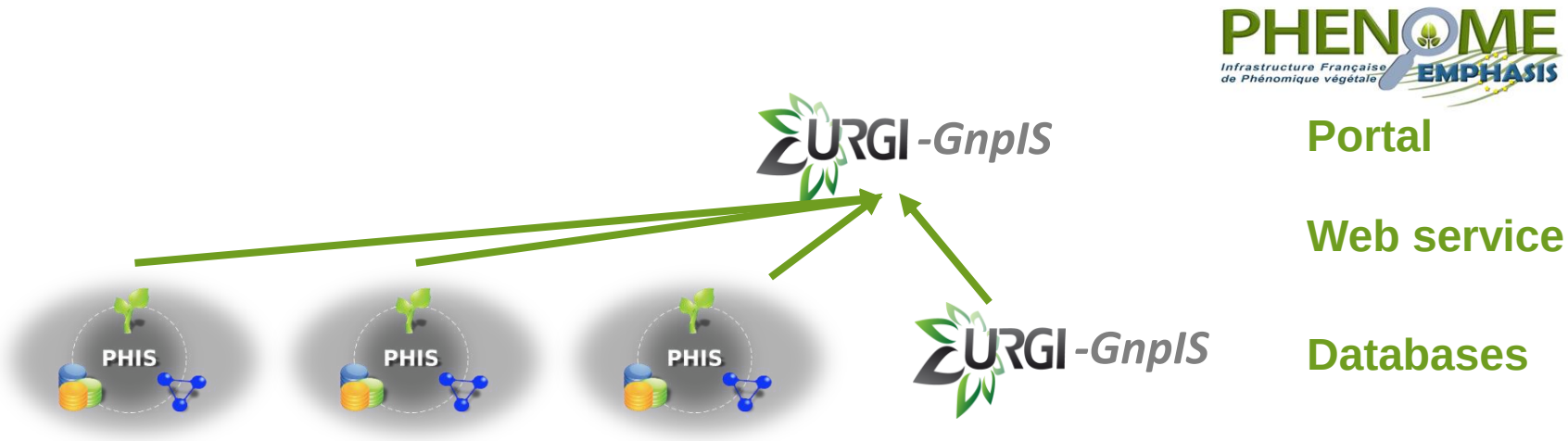


Genotype	treatment	rep	LAI
Soisson	fi	1	5
Soisson	fi	2	7

Genotype	treatment	LAI
Soisson	low input	6

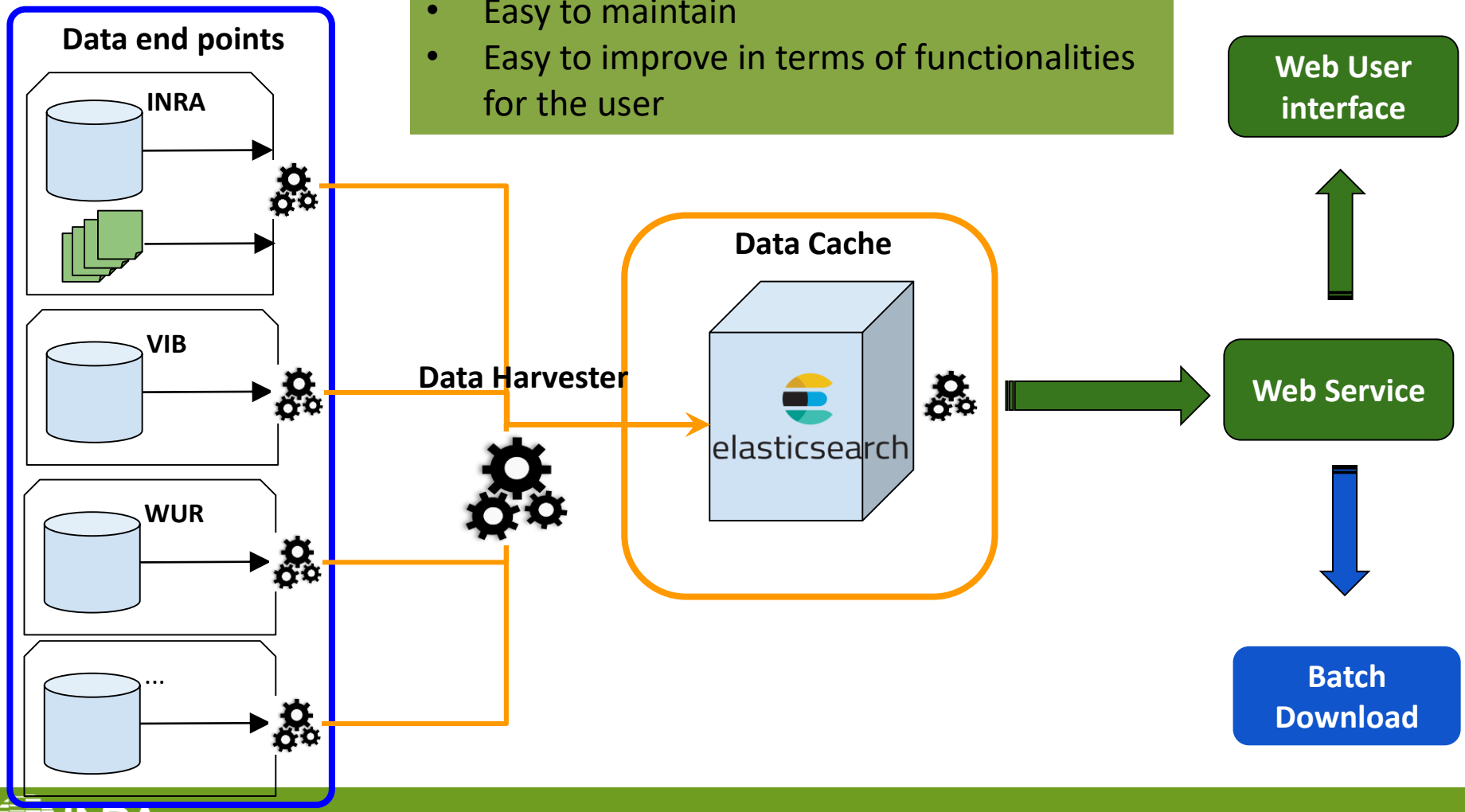
Strategy for data management

- Raw data
  - Big in High through put platforms: Terabyte by experiment
  - Difficult to obtain again and LT interest: long term storage
- No central Phenomic data archive
- Strategy: develop a federation of information system with central services
  - Need for standards to support the federation
  - Need for standard to support automatic data integration



Federation of Plant phenotyping information systems

- Easy to maintain
- Easy to improve in terms of functionalities for the user



Plant Phenotyping Data Life cycle

EMPHASIS

Elixir

Following experiments → **Phenotypic analysis** → **Data sharing and Reuse**

TRAITS

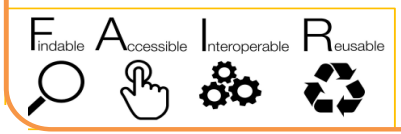
- Tracking
  - objects, events, images
- Following
  - plants, sensors, cameras
- Environmental conditions Mapping/time course
- Sensor calibration

- Data cleaning
- Calculated traits
  - Time courses
  - Spatial analysis
  - Ratios (growth per ...)
- Tools for data quality
  - Reproducible?
  - Cross scale ?
  - Modelling?

- From measurements to traits
- Tools for trait quality
  - heritability
  - genetic correlations

- Enable reproducible analyses (GxE, GWAS, Genomic Prediction, Evolution)
- Data integration and linking
- Data Publication
- Data discovery
- Pipeline Hosting

Interoperable Reusable | Interoperable Reusable | Reusable | Findable Accessible Interoperable



Crop Ontology
for agricultural data
www.cropontology.org

Elizabeth Arnaud et al.

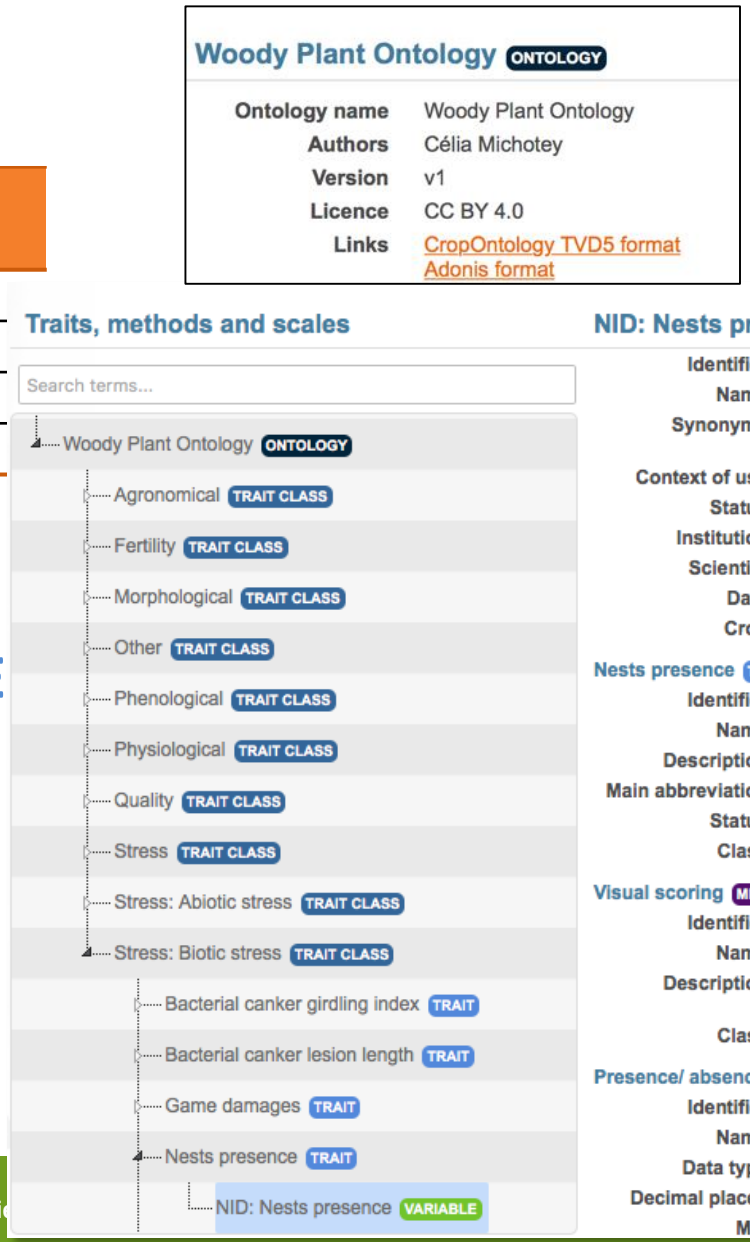
Phenotype data in breeding

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

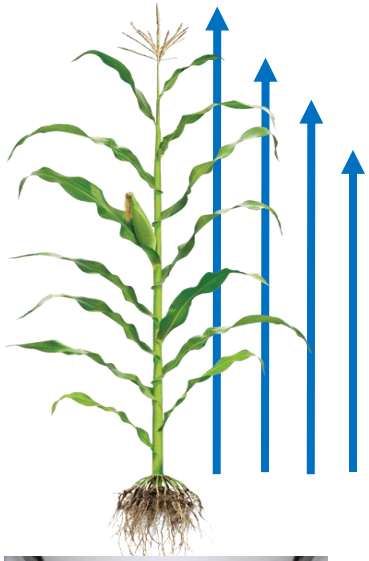
=

1 TRAIT + 1 METHOD + 1 SCALE



Variable identification: Plant height example

Trait + Method + Unit



T1: Plant Height

M1: Total height

M2: First tassel branch

M3: Last expanded leaf

M4: Youngest growing leaf

U1: cm

U2: mm

M5: Highest pixel corresponding to plant

U3: pixel

…Allows as many combination as needed…

…Each trait, method and unit has to be identified

Object identification (URIs)

URIs (Uniform Resource Identifiers)

⇒ standardized, unique, unambiguous identification of objects

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

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

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

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

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

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

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

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

URI of plot
<diaphen:2017/o1700029>

URI of plant:
<diaphen:2017/17000147>

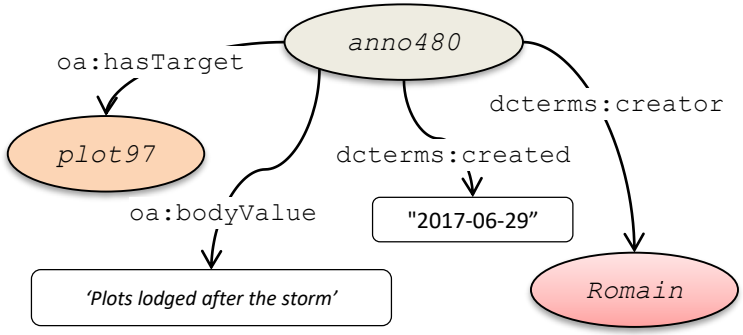
URI of leaf:
<diaphen:2017/l17000590>

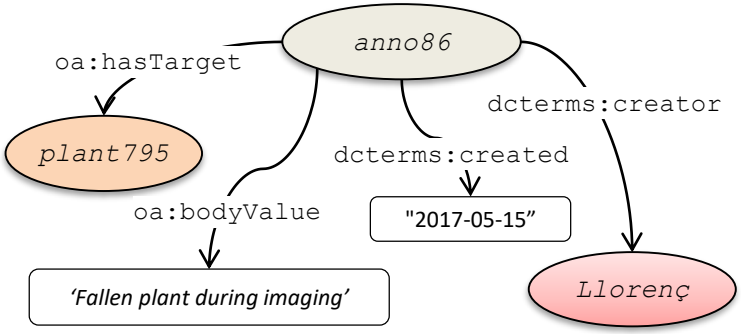
URI of camera:
<diaphen:2018/ac180002>

URI of image:
<diaphen:2017/ic14001480237>

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

Annotation of events: plot and plant accidents





Events or follow the **Web Annotation Data Model** that allows assigning motivation and purpose attributes to annotations (e.g. oa:describing, oa:identifying, oa:linking, oa: replying, etc.). **Dublin Core** properties such as dcterms:created or dcterms:creator are also used.

Minimum Information About Plant Phenotyping Experiment

- Standard for metadata : data about the data set
- www.miappe.org
- Recently improved: detailed specification of the checklist and alignment with other existing standards, include needs of the forest trees community, …

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

+ Author Affiliations

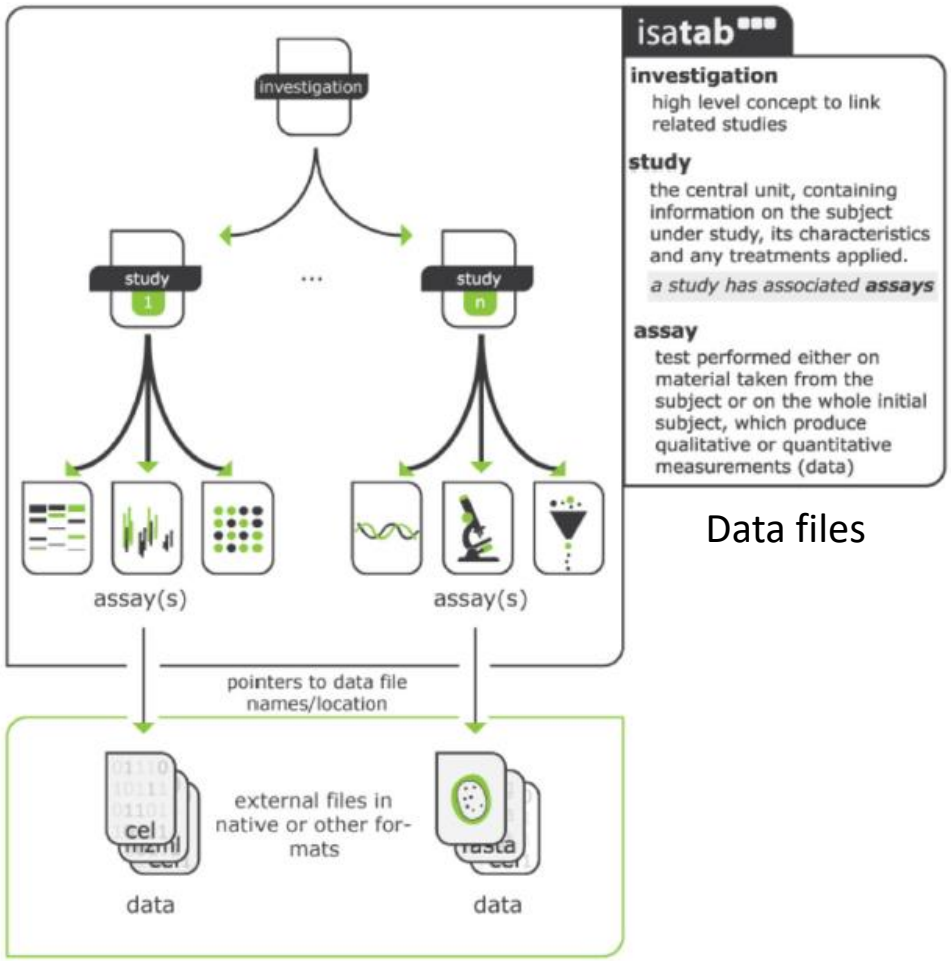
« Previous | Next Article »
Table of Contents

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 Issue: Phenotyping in Plants

MIAPPE implementation: File Archive standard format

- ISA Tab for Phenotyping
  - Investigation/Study/Assay
  - Zip Archive
    - MIAPPE Metadata
    - Raw data
    - Elaborated data
    - Linked data / Metadata / Discovery data



Data files

MIAPPE implementation: Web Service

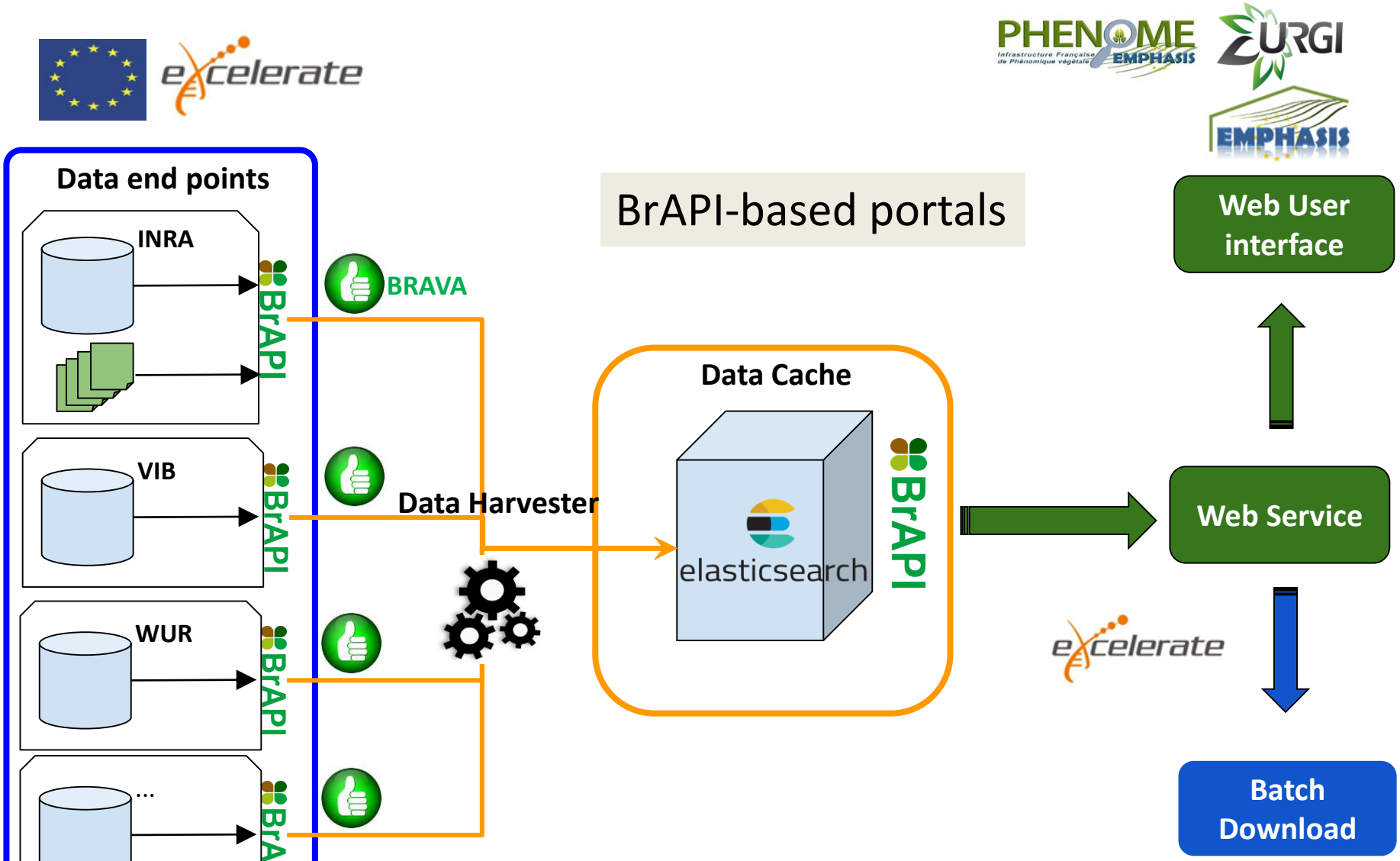
- Breeding API: International collaboration
- Servers implementations => standard data exports
  - CGIARs international network: Integrated Breeding Platform
  - Elixir Plant Community (GnpIS, CIRAD, …)
  - Emphasis Plant Community: (+PHIS, …)
  - Germinate (James Hutton Institute, UK)
- Clients implementations => functionalities for users
  - Flapjack (JHI, UK): genotyping data visualization
  - Search data in federations (URGI, EU)
  - R analysis pipelines (CIP, Peru)
  - Ontology Widget (URGI, FR)
  - Diverse BrAPI Apps (USA) currently being developed for data visualisation
- Tools => functionalities for data managers
  - BRAVA (IPK, D): validator of a BrAPI server implementation
  - BrAPI to ISA-TAB; Batch downloads (ELIXIR, EU)

BrAPI

Bill & Melinda Gates Foundation
CassavaBase
T3
IBP
JHI
Bioversity
CIRAD
INRA
IRRI
GOBII
Wageningen
CIP
DaRT
Cornell
iPlant

www.brapi.org

Federation(s) of Plant Information systems



Aknowledgment & Questions