



**HAL**  
open science

## Improving the “ FAIRness ” of Inra’s data for plant biology and breeding

Michael Alaux, Cyril Pommier, Esther Dzale Yeumo, Sophie Durand, Raphaël Flores, Erik Kimmel, Thomas Letellier, Célia Michotey, Nacer Mohellibi, Hadi Quesneville, et al.

### ► To cite this version:

Michael Alaux, Cyril Pommier, Esther Dzale Yeumo, Sophie Durand, Raphaël Flores, et al.. Improving the “ FAIRness ” of Inra’s data for plant biology and breeding. PAG XXVI - Plant and Animal Genome Conference, Jan 2018, San Diego, United States. pp.14 slides. hal-02790227

**HAL Id: hal-02790227**

**<https://hal.inrae.fr/hal-02790227>**

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.



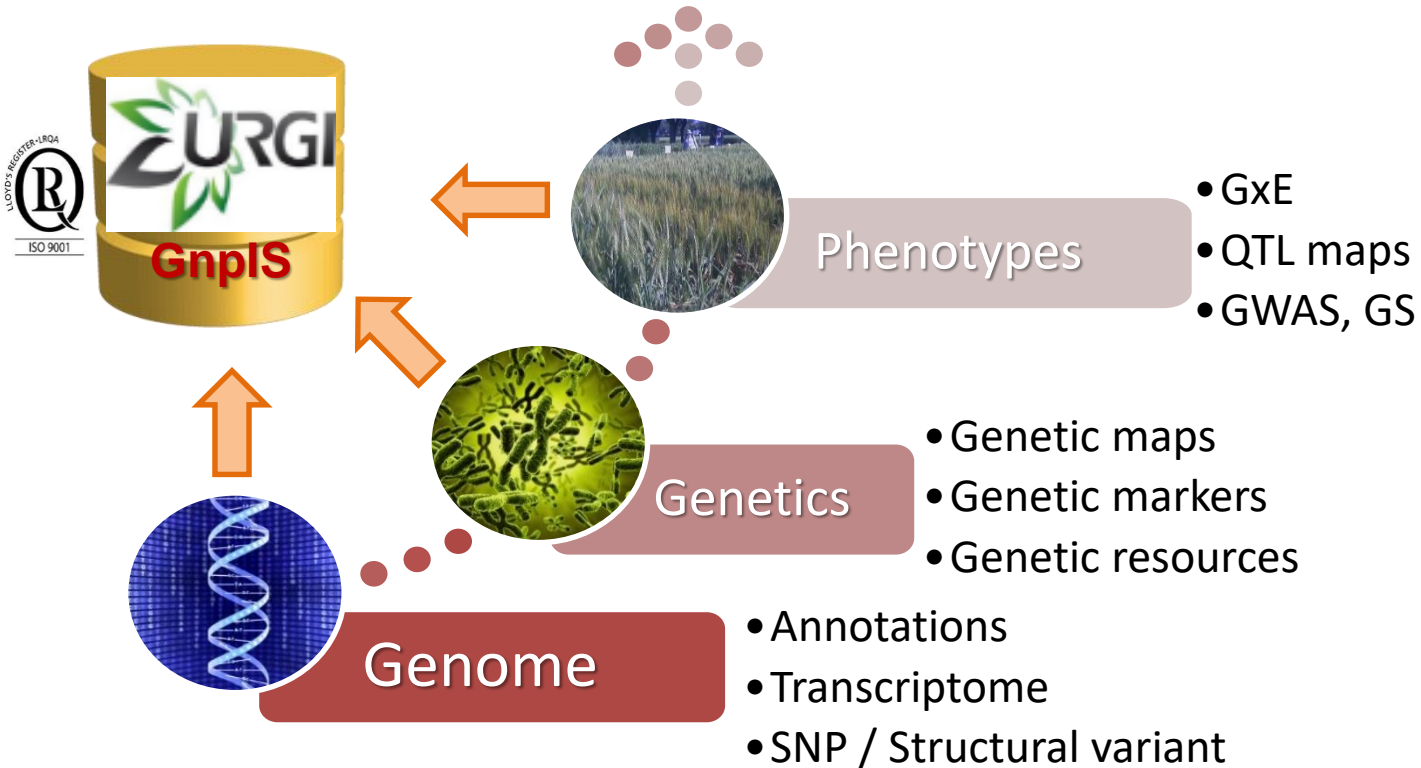
# Improving the « FAIRness » of Inra's Data for Plant Biology and Breeding

Unit of Research in Genomic-Info (URGI), INRA

**Anne-Françoise Adam-Blondon**



# GnpIS: INRA IS for crops, forest trees and pathogens



# Main global objectives

- Be a robust and sustainable repository of FAIR data (Findable, Accessible, Interoperable, Reusable)
- Integrate GnpIS in sustainable and robust federations of information systems
- Facilitate knowledge development and data analysis

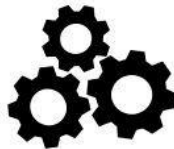
F  
indable



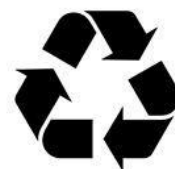
A  
ccessible



I  
nteroperable



R  
eusable



# Acknowledgements



URGI team



H. Quesneville	C. Guerche
C. Pommier	E. Kimmel
M. Alaux	M. Lainé
D. Charrnaud	T. Letellier
G. Cornut	M. Loaec
S. Durand	C. Michotey
R. Flores	N. Mohellibi
N. Francillonne	F. Philippe

## Financial supports



## International infrastructures /initiatives



## National and international crop projects



Betterave2020



International  
Wheat Genome  
Sequencing  
Consortium

# Making data and Information systems FAIR has a lot to do with community management

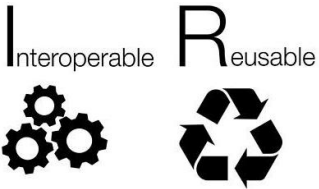
Within and between:

- Developers
- Specialists of ontologies and standards
- Data managers
- Biologists (data producers)

} (Global)  
Infrastructure  
projects

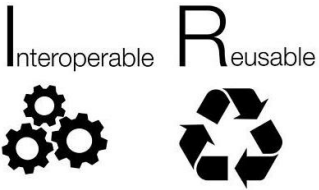
} Crop/biology  
focused  
projects

And clarifying what is under the responsibility of each community



# Metadata : data about the phenotyping experiment

- MIAPPE: Minimum Information About Phenotyping Experiment
- Developed and maintained by an international community interested in plant phenotyping: large community of breeders and biologists, European infrastructure for Plant Phenotyping (EPPN/EMPHASIS), European infrastructure of Bioinformatics (ELIXIR), Planteome, Excellence in Breeding Platform...
- [www.miappe.org](http://www.miappe.org)
- Steering committee Emphasis, Elixir CGIARs



# Metadata : data about the phenotyping experiment

Crop Ontology  
Variable=trait + method + scale

Identification : MultiCrop  
Passport Data  
Standard

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

MIPPE standard aligned with  
MCPD and Crop Ontology  
standards

*Inspire EU directive?*

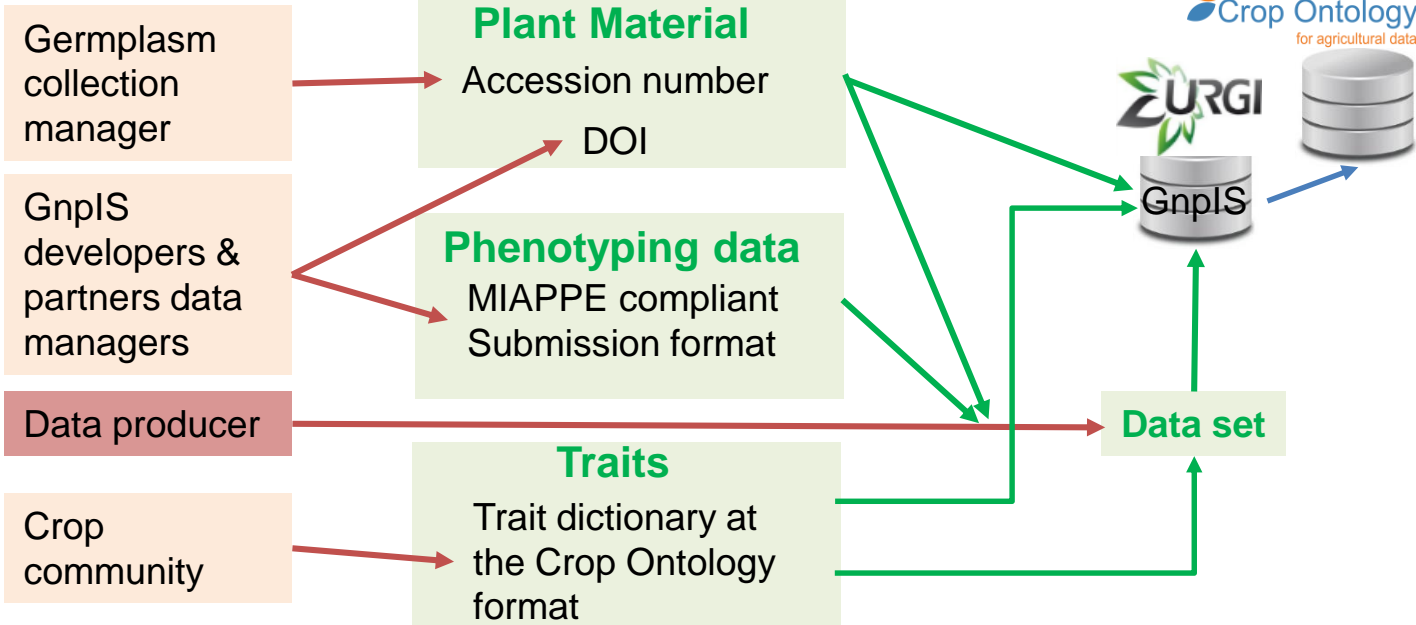


# Who provides the metadata – e.g. phenotyping data

## Who

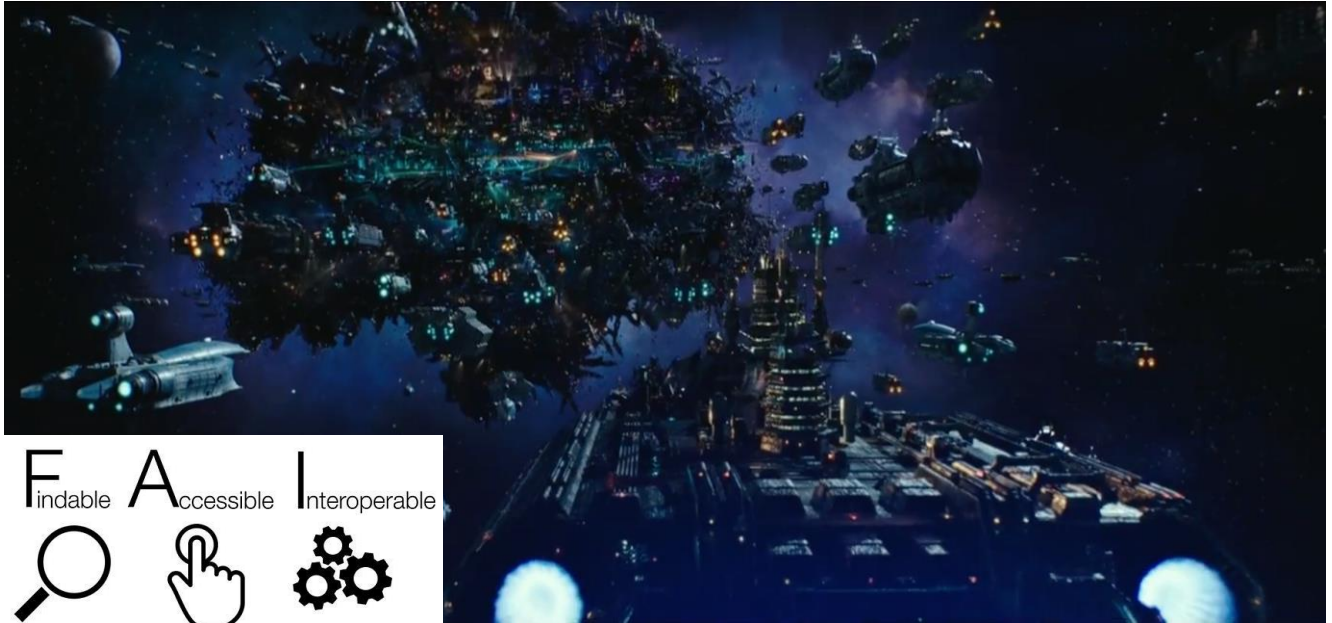
## What

## Repositories



**Data managers:** important role in facilitating flows between registries of identifiers, data set repositories and data producers

# Develop a sustainable and robust federation of plant information systems



Data is scattered in many information systems

# Findability: Data discovery through a common portal

Spannagl et al 2016, doi: 10.3835/plantgenome2015.06.0038



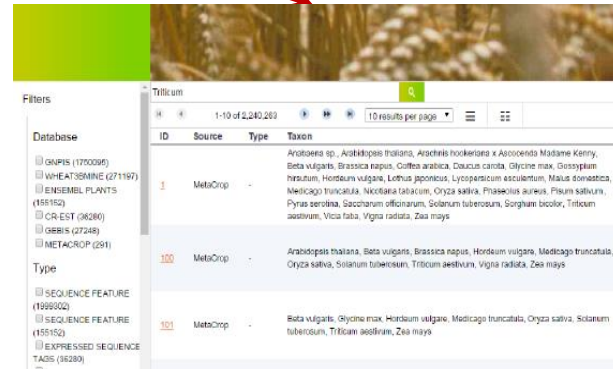
User web interface



<http://www.wheatis.org/>

Google like list of results

Common  
Data  
Model



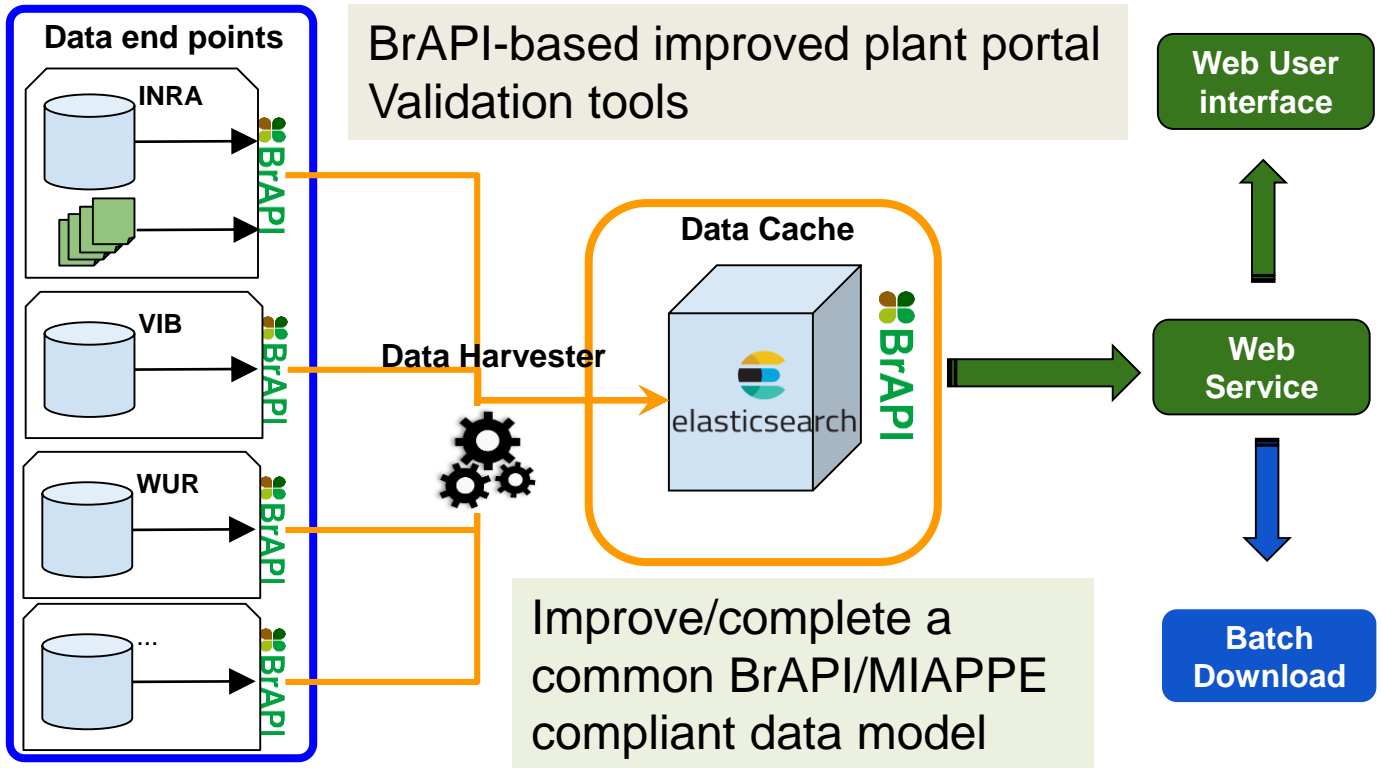


## Challenges:

- Synchronize technical updates of the infrastructure
- Synchronize improvements of the data model
- Searching with increasingly natural language (e.g. for traits)

Opportunity: great tool to build a community of developers, data managers and specialist of ontologies that work together

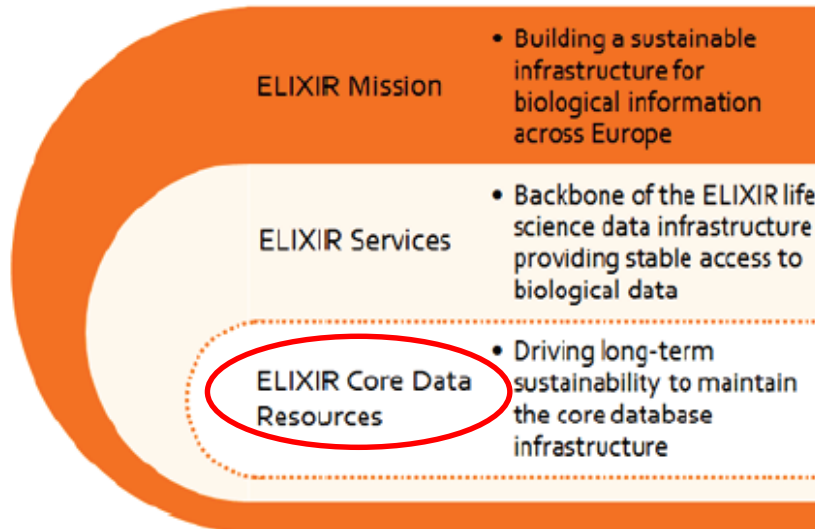
# Federation of Plant Information systems



# Federation(s) of plant information systems

Development of a consistent and robust suite of open source tools based on common internationally agreed standards : one of the key elements of a sustainable federation

Durinx C, *et al.* **Identifying ELIXIR Core Data Resources**  
*F1000Research* 2016, doi: 10.12688/f1000research.9656.1



**Thank you!**