Developping grapevine FAIR data

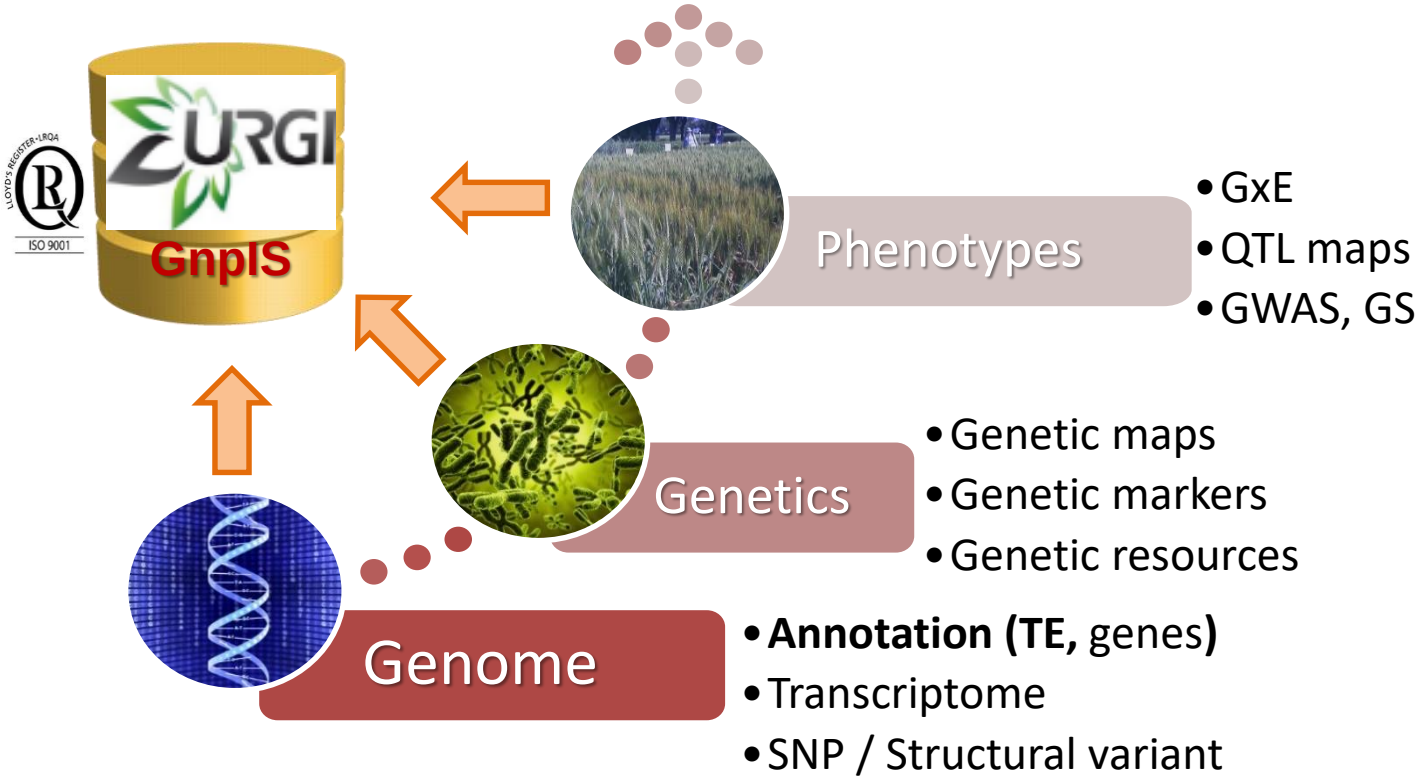
Anne-Francoise Adam-Blondon

Developing Grapevine FAIR data

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

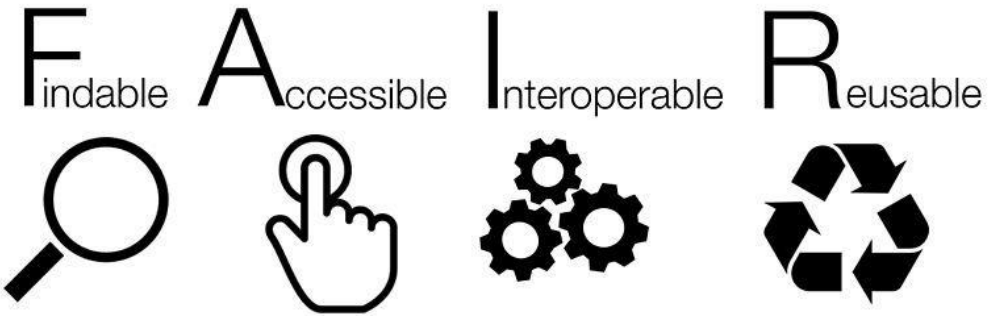
Anne-Françoise Adam-Blondon

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 because data is highly dispersed and likely to stay so
- Facilitate knowledge development and data analysis



Acknowledgements

URGI team

H. Quesneville
C. Pommier
M. Alaux
D. Charruaud
G. Cornut
S. Durand
R. Flores
N. Francillonne

C. Guerche
E. Kimmel
M. Lainé
T. Letelllier
M. Loaec
C. Michotey
N. Mohellibi
F. Philippe

Financial supports

International infrastructures /initiatives

National and international crop projects

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

Within and between:

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

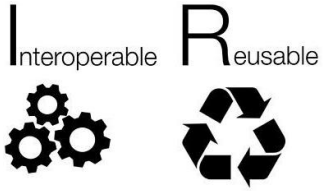
(Global) Infrastructure projects

Crop/biology focused projects

And clarifying what is under the responsability of each community

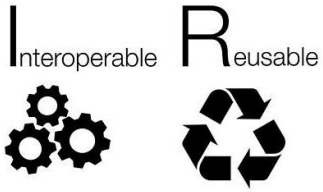
Example of what is currently done to develop international federation of FAIR information systems for :

- the wheat community
- phenotyping data



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 Plateform…
- www.miappe.org
- Steering comittee 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

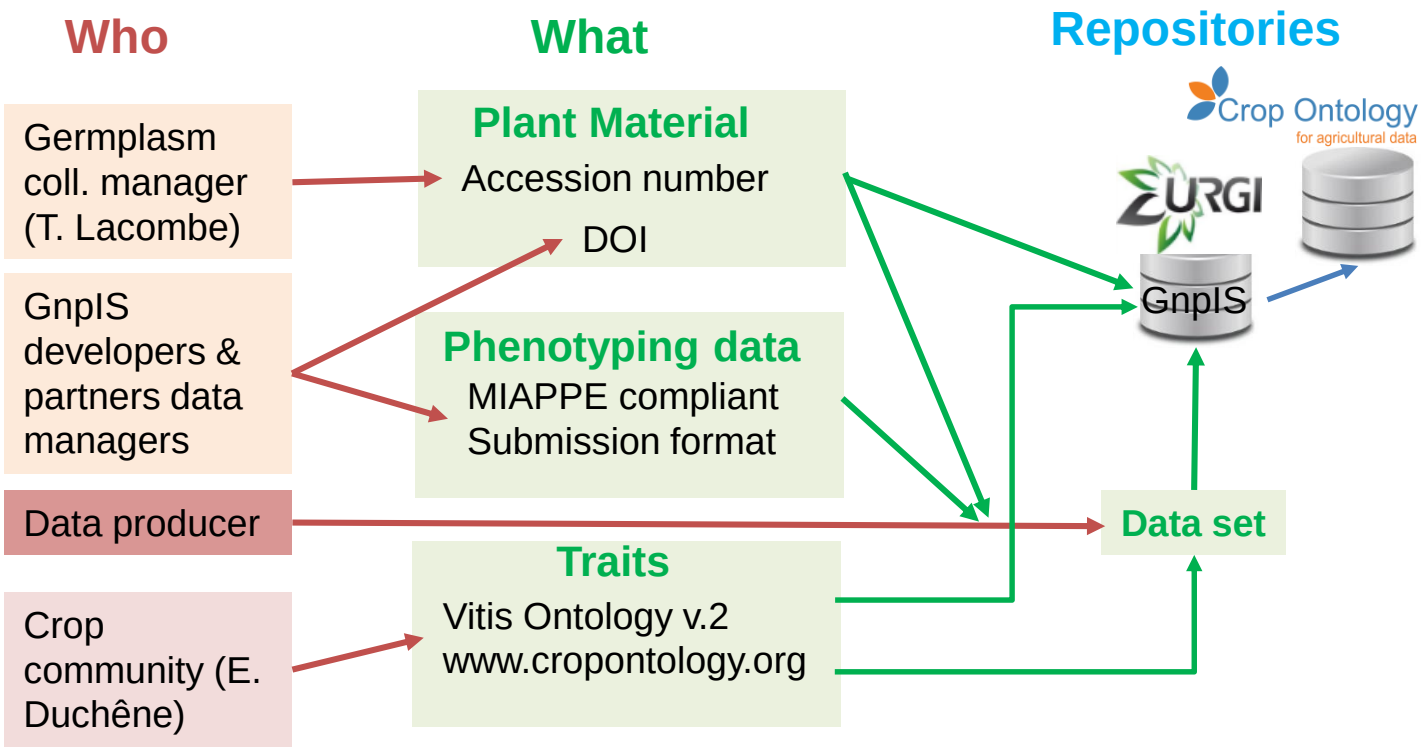
Genotype = observed marker's alleles on a cultivar

Climate 1 = climatic data at GPS1-time1

MIAPPE standard aligned with MCPD and Crop Ontology standards

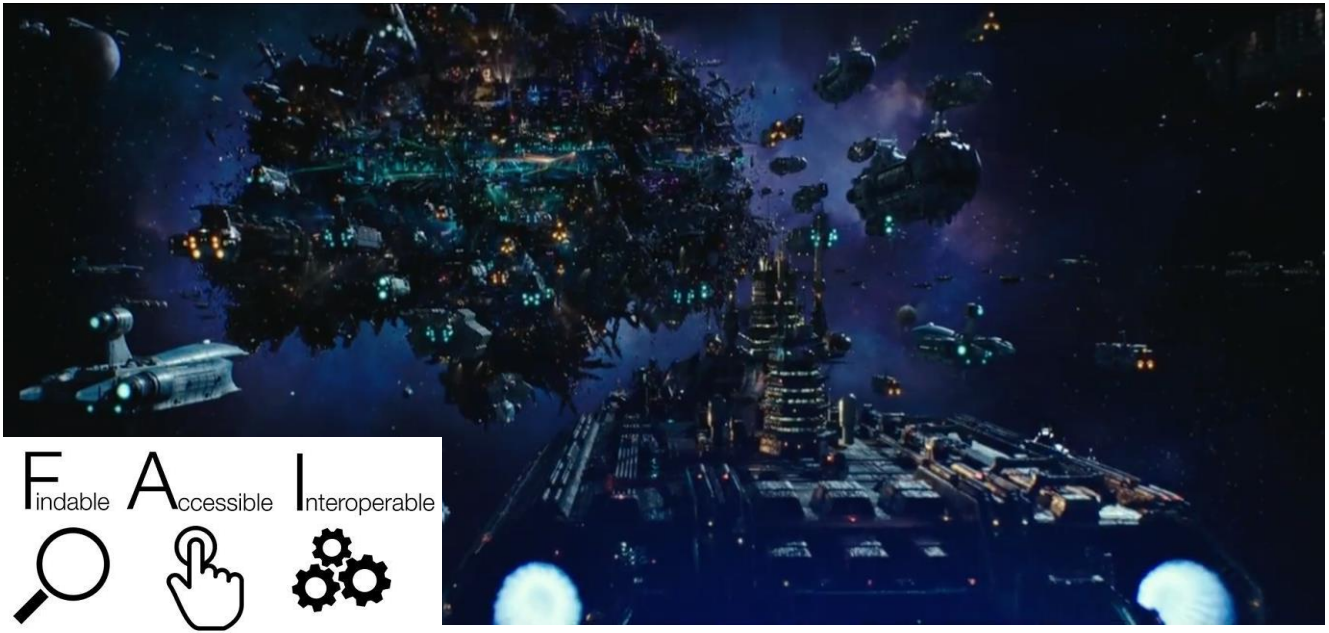
Inspire EU directive?

Who provides the metadata – e.g. phenotyping data at INRA



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

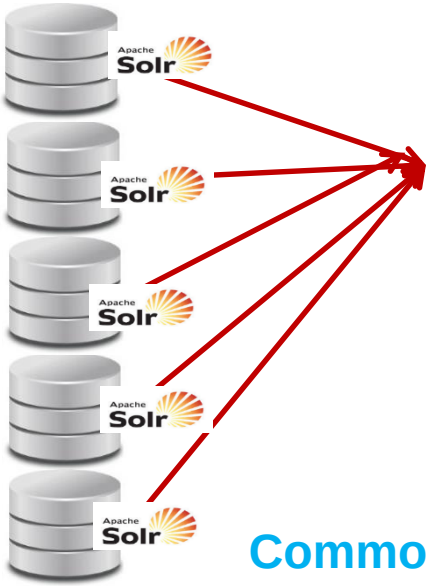
Develop a sustainable and robust federation of plant information systems



Data is scattered in many different information systems

Findability: Data discovery through a common portal

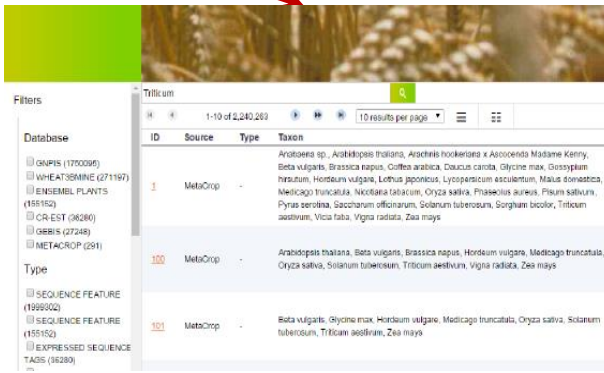
Spannagl et al 2016, doi: 10.3835/plantgenome2015.06.0038



User web interface

http://www.wheatis.org/

Google like list of results



Common Data Model

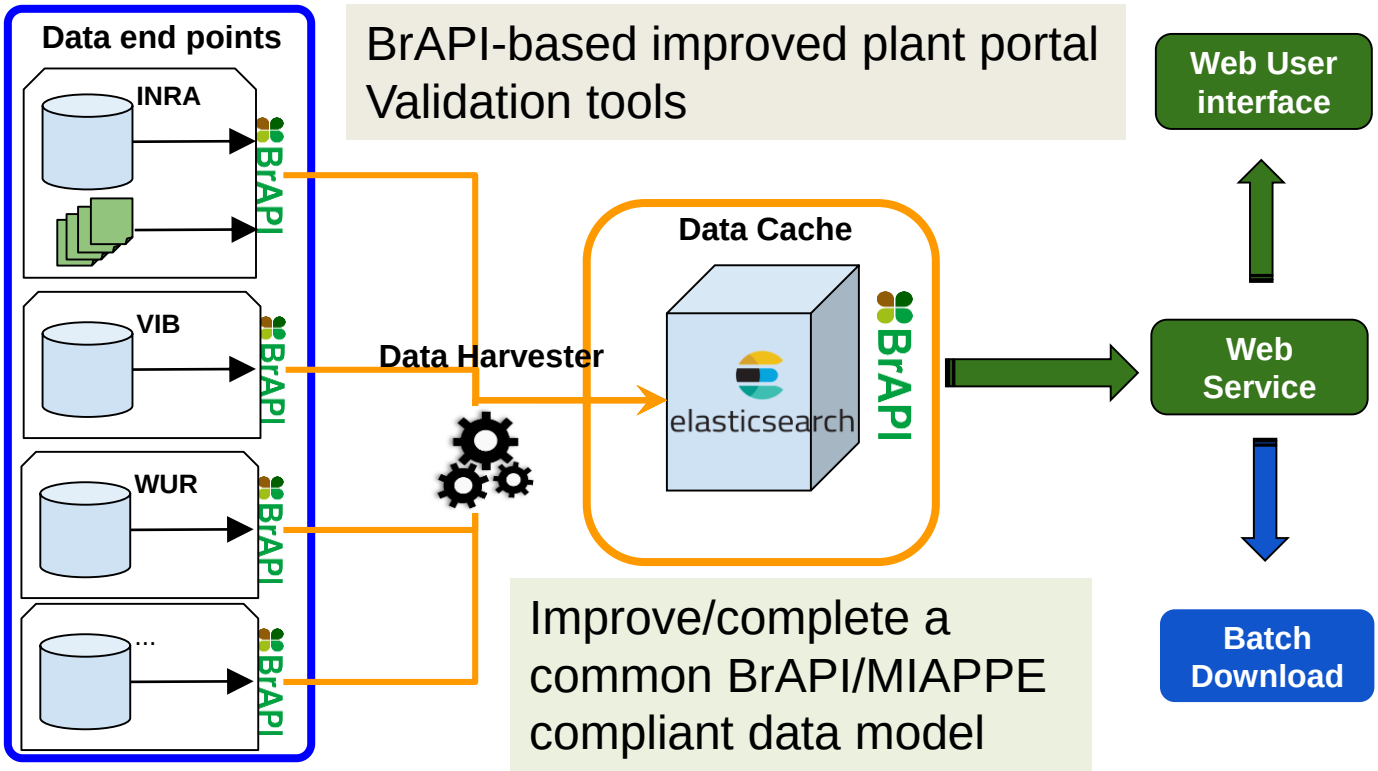
www.wheatis.org

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

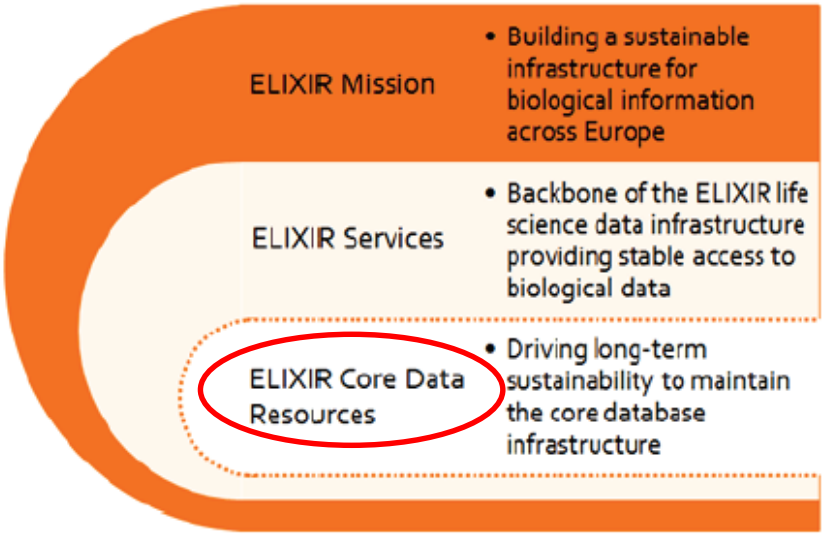
Federation of Plant Information systems



Federation(s) of plant information sytems

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



Conclusions

- Survey of the background in the grapevine community in the frame of the IGGP
  - A-F Adam-Blondon et al. (2016) Towards an open grapevine information system. Hort Res, 3, 16056. https://doi.org/10.1038/hortres.2016.56

One implementation of a plant federation we could start playing with: https://urgi.versailles.inra.fr/ifb/

Thank you!