



Developping grapevine FAIR data

Anne-Francoise Adam-Blondon

► To cite this version:

Anne-Francoise Adam-Blondon. Developping grapevine FAIR data. 27. Plant and Animal Genome Conference (PAG) 2018, Jan 2018, San Diego, CA, United States. pp.16. hal-02786779

HAL Id: hal-02786779

<https://hal.inrae.fr/hal-02786779>

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 4.0 International License



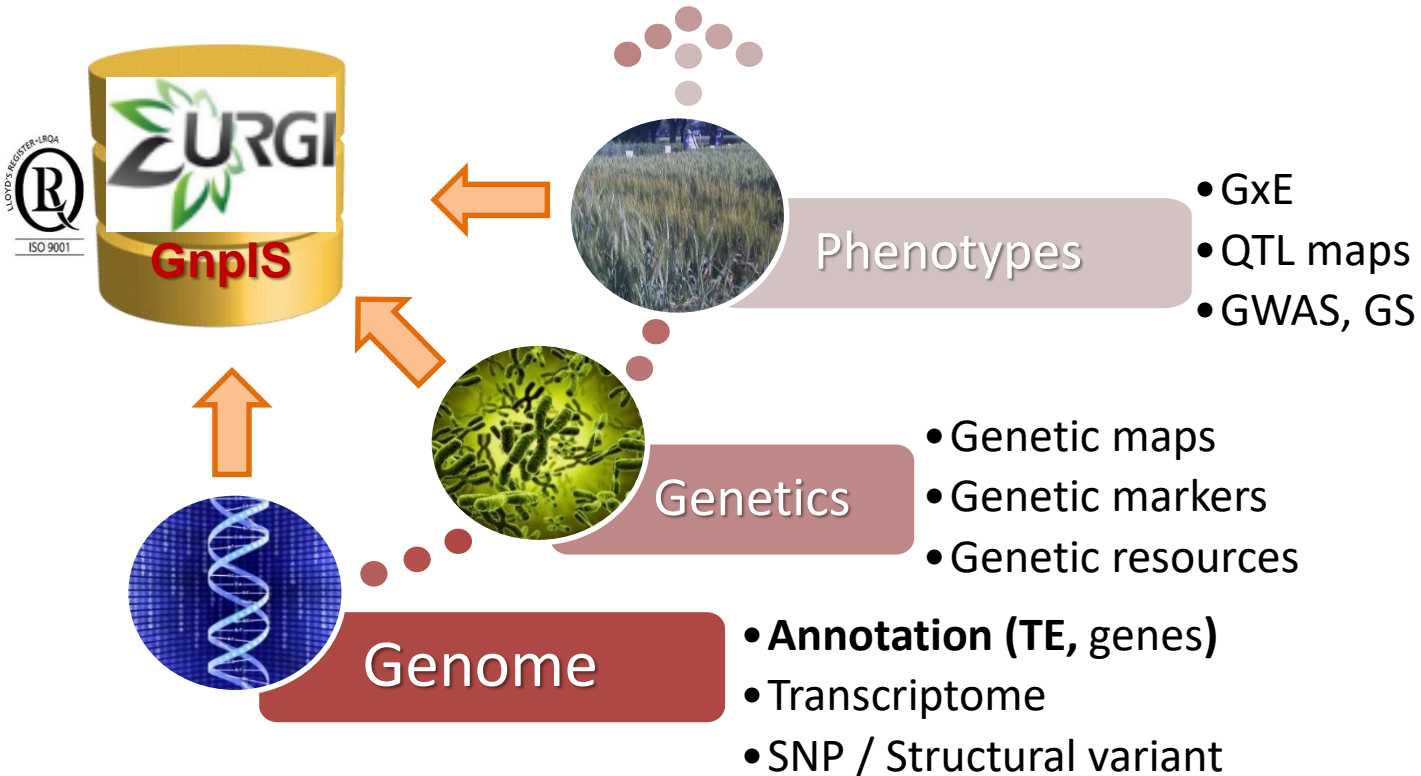
Developing Grapevine FAIR data

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

Anne-Françoise Adam-Blondon

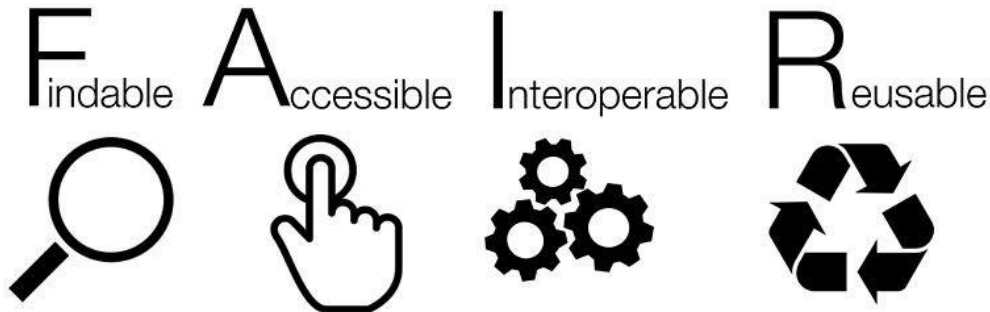


GnplS: INRA IS for crops, forest trees and pathogens



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. Guerche
C. Pommier	E. Kimmel
M. Alaux	M. Lainé
D. Charruaud	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



International
Wheat Genome
Sequencing
Consortium

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

Within and between:

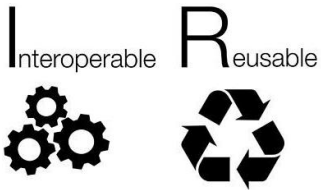
- Developers
- Specialists of ontologies and standards
- Data managers
- Biologists (data producers)

} (Global)
Infrastructure
projects
} Crop/biology
focused
projects

And clarifying what is under the responsibility of each community

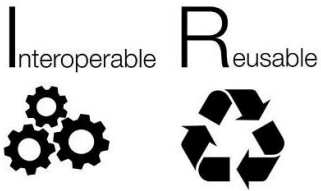
Example of what is currently done to develop international federation of FAIR information systems for :

- 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 Platform...
- 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

MIAPPE standard aligned with
MCPD and Crop Ontology
standards

Inspire EU directive?

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

Who

Germplasm
coll. manager
(T. Lacombe)

GnpIS
developers &
partners data
managers

Data producer

Crop
community (E.
Duchêne)

What

Plant Material

Accession number
DOI

Phenotyping data

MIAPPE compliant
Submission format

Traits

Vitis Ontology v.2
www.cropontology.org

Repositories

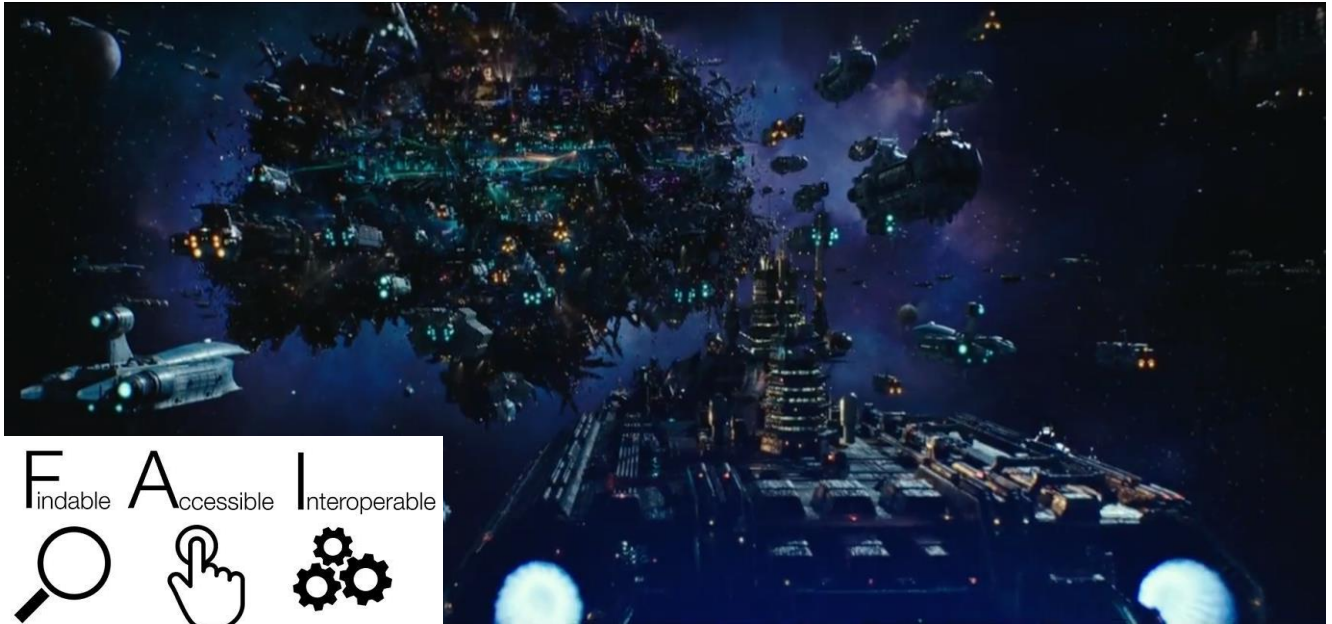
 Crop Ontology
for agricultural data



Data set

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

Filters

Database

- ☐ GENES (170000)
- ☐ WHEATISMINI (271197)
- ☐ ENSEMBL PLANTS (155152)
- ☐ CR EST (26280)
- ☐ GENIS (27248)
- ☐ METACROP (291)

Type

- ☐ SEQUENCE FEATURE (1999002)
- ☐ SEQUENCE FEATURE (155152)
- ☐ EXPRESSED SEQUENCE TAGS (32280)

ID	Source	Type	Taxon
1	MetaCrop	-	Arachis sp., Arachis thaliana, Arachis hookeriana x Acaecidea Madame Kerry, Beta vulgaris, Brassica napus, Coffea arabica, Daucus carota, Glycine max, Gossypium hirsutum, Hordeum vulgare, Lathyrus japonicus, Lycopersicon esculentum, Malus domestica, Medicago truncatula, Nicotiana glauca, Oryza sativa, Phaseolus aureus, Pinus taeda, Pyrus serotina, Saccharum officinarum, Solanum tuberosum, Sorghum bicolor, Triticum aestivum, Vicia faba, Vigna radiata, Zea mays
100	MetaCrop	-	Arachis thaliana, Beta vulgaris, Brassica napus, Hordeum vulgare, Medicago truncatula, Oryza sativa, Solanum tuberosum, Triticum aestivum, Vigna radiata, Zea mays
101	MetaCrop	-	Beta vulgaris, Glycine max, Hordeum vulgare, Medicago truncatula, Oryza sativa, Solanum tuberosum, Triticum aestivum, Zea mays

Common
Data
Model

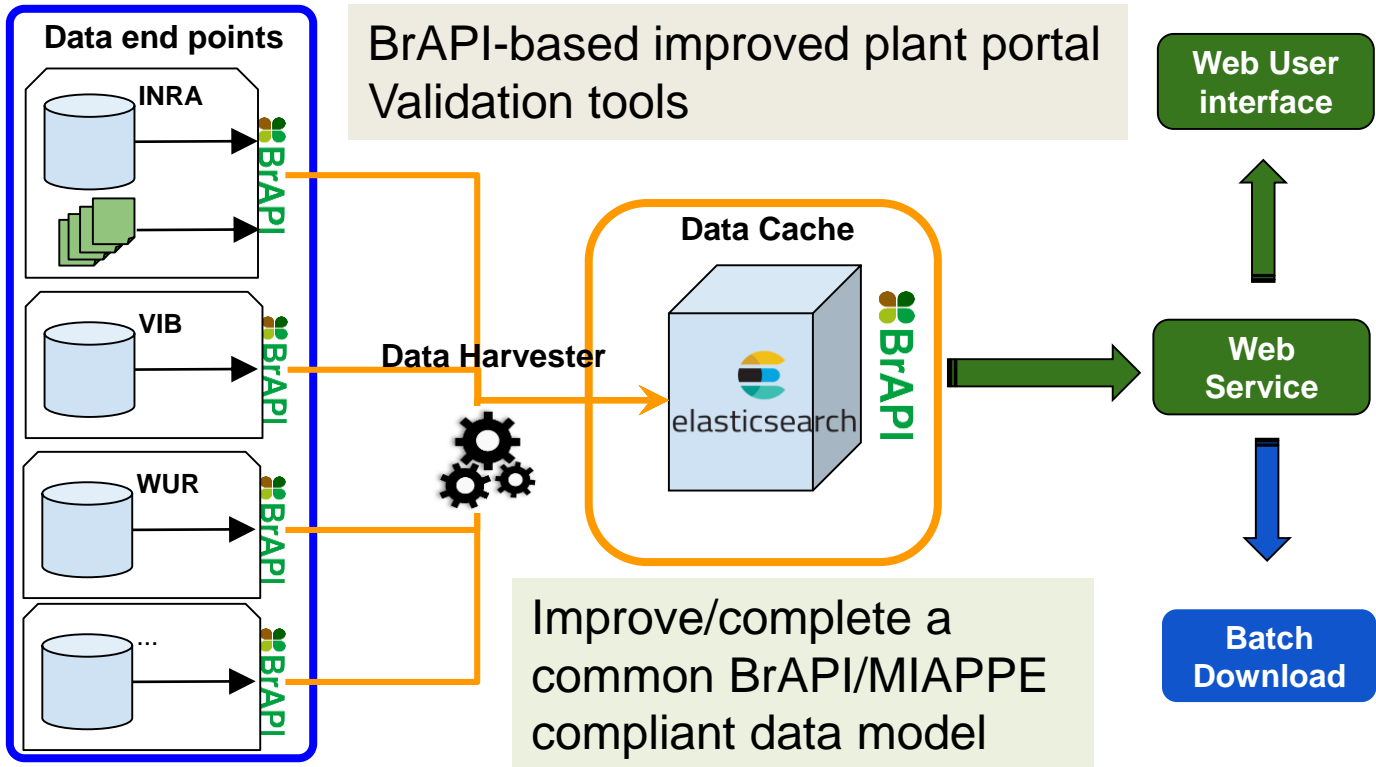


Challenges:

- 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

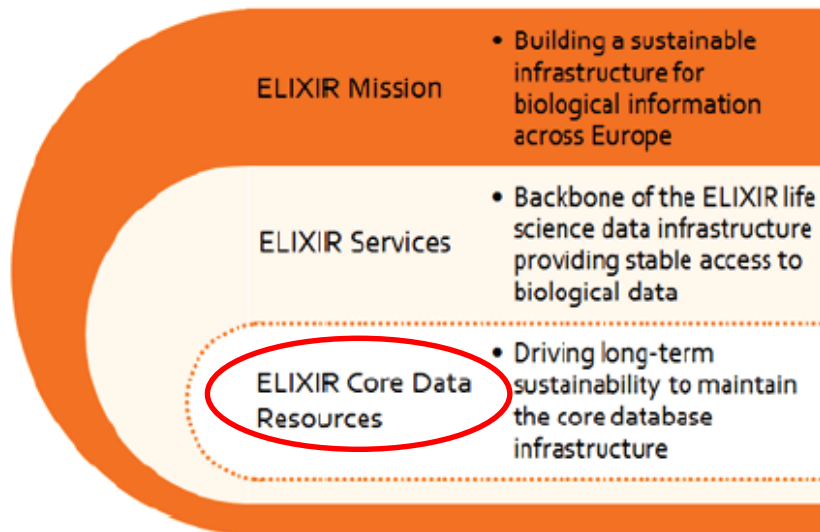
Federation of Plant Information systems



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



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/>

IFB

10 re

1-10 of 407

flowering time

ID	Source	Type	Taxon	Descriptio
GO:0009555	Gramene	GO process	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	GO process specific out formation to
GO:2000028	Gramene	GO process	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	GO process Any proces flowering., r
GO:0048574	Gramene	GO process	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	GO process change from or exposure
GO:0048573	Gramene	GO process	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	GO process the vegetati to, a period
GO:0048575	Gramene	GO process	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	GO process change from exposure to
GO:0009641	Gramene	GO process	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	GO process of response plant. It ofte
IPR007133	Gramene	InterPro Family	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	InterPro Far ll associate (Leo1), VIP!
IPR007149	Gramene	InterPro Family	Triticum aestivum, Triticum urartu, Brachypodium distachyon, Aegilops tauschii, Hordeum vulgare subsp. vulgare	InterPro Far Paf1C relat (CtR) and A

Database

- ☐ GRAMENE (272)
- ☐ GNPIIS (115)
- ☐ TRITICEAE TOOLBOX (10)
- ☐ WHEAT PANGENOME (8)
- ☐ GRAINGENES (2)

Type

- ☐ GENE (259)
- ☐ SEQUENCE FEATURE (120)
- ☐ EXPERIMENT (10)
- ☐ INTERPRO FAMILY (7)
- ☐ GO PROCESS (6)
- ☐ PHENOTYPE (3)
- ☐ QTL (2)

Species

- ☐ TRITICUM AESTIVUM (81)
- ☐ BRACHYPODIUM DISTACHYON (74)
- ☐ TRITICUM URARTU (66)
- ☐ AEGILOPS TAUSCHII (62)
- ☐ HORDEUM VULGARE SUBSP. VULGARE (60)
- ☐ ARABIDOPSIS THALIANA (49)
- ☐ SOLANUM LYCOPERSICUM (25)
- ☐ VITIS VINIFERA (18)
- ☐ BRASSICA NAPUS (15)

© INRA-IRGI 2006-2018 | Legal mentions | Designed by

Thank you!