Federations of information systems in the plant community and possible application to grapevine
Anne-Francoise Adam-Blondon

To cite this version:
Anne-Francoise Adam-Blondon. Federations of information systems in the plant community and possible application to grapevine. First Annual conference of COST INTEGRAPE, Mar 2019, Chania, Greece. pp.22. hal-02786993

HAL Id: hal-02786993
https://hal.inrae.fr/hal-02786993
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
Federations of information systems in the Plant community and possible application to Grapevine

A-F Adam-Blondon (INRA, FR)
Enabling interdisciplinary research

Example: modelling the impact of climate change using plant phenology

Global plant phenology data portal: [www.plantphenology.org](http://www.plantphenology.org)
Pan European PEP725 Plant phenology database: [http://www.pep725.eu/](http://www.pep725.eu/)

**Phenology data**

- Date associated to:
  - Geo-reference
  - species
  - event (e.g. 50% of opened flowers)

**Climatic data**

- Date associated to:
  - Geo-reference
  - Measures (e.g. T°C, …)
  - Events (e.g. hail, …)

Interoperability
Example: modelling the impact of climate change using plant phenology

**Phenology data**: different sources, different accuracy in terms of identification of the plant material, scoring methods, record formats, …

- Modelers of the impact of climate change
- Geneticists, Breeders
- Genbank managers
- Experimental station managers
- Producers: vintage dates
- Crowd sourcing
- …
Data is scattered in an “ecosystem” of information systems

Insertion of GnpIS in federations of information systems

2002: Single database

2008: DB Interoperability

2010: set of interoperable databases

2015: Distributed information system

URGI DBs + GnpIS = UGDB

GxE
QTL maps
GWAS; GS

Phenotypes

Genetics

Genome

Annotation; Synteny
Transcriptome
SNP / Structural variants

Information System

Funded by the Horizon 2020 Framework Programme of the European Union
Featuring federations of information systems

- A network of (stable/sustainable) nodes
- A central portal offering services (e.g. search data)

Backbone of good practices enabling such infrastructures

SCIENTIFIC DATA, 3:160018, DOI: 10.1038/sdata.2016.18
Two main examples of federations

- The Wheat Initiative (G20 Initiative) and its Wheat Information System Expert Working group ([www.wheatis.org](http://www.wheatis.org)). Also supported by the Research Data Alliance.

- The European Infrastructure for Multi-scale Plant Phenomics and Simulation (EMPHASIS) and its information system ([https://emphasis.plant-phenotyping.eu/e-Infrastructure](https://emphasis.plant-phenotyping.eu/e-Infrastructure)). In the frame of a strong collaboration with ELIXIR.
Developing a federation of FAIR data repositories and commons
Development of guidelines: e.g. www.wheatis.org

https://ist.blogs.inra.fr/wdi/
Registries of standards and guidelines

http://agroportal.lirmm.fr/

https://www.ebi.ac.uk/ols/ontologies

http://agroportal.lirmm.fr/
Developing search and data access across federations
Generic Data Discovery Tool

User web
http://www.wheatis.org/

Common Data Model

transPLANT data model

**Sequence feature**
- feature type
- [sequence_id]
- [sequence_version]
- [start_position]
- [end position]

**Genetic marker**
- map
- map_position
- [xref] (Seq. feature)

**Accession**
- authority

**Phenotype**
- [trait]
- [trait_id]
- [environment]*
- [environment_id]*
- [statistic]
- [unit]

**Database entry**
- id
- entry_type
- database_name
- db_id
- [db_version]
- [name]
- description
- url
- [species]*
- [xref]

**GWAS result**
- genotype
- xref (Phenotype)

**QTL**
- xref (Seq. feature)
- xref (Phenotype)

**Experiment**
- experiment_type
- [xref]* (Phenotype)
- [xref]* (Accession)

**Reaction**
- no additional fields defined

**Other**
- no additional fields defined

- Full text search
- Facettes / Aggregations
WheatIS data discovery: https://urgi.versailles.inra.fr/wheatis/

Wheat@URGI portal

Sequence Repository

GnplS

File Repository

Graingenes

Gramene

EnsemblPlants

TheTriticeaeToolbox

WheatIS Data Discovery tool

PlantsDB, CIMMYT, …

WheatIS data discovery tool: evolution

Open software, very generic, that can (and is) adapted to any type of federation:
e.g. the federation of information systems for of the French infrastructure of genetic resources for research in agriculture, AgroBRC-RARe (!beta! https://urgi.versailles.inra.fr/rare-beta/)
The GnplS search tool: prefiguration of a search tool for a plant community
Deeper interoperability in federations
The Breeding API (BrAPI) Project is an international effort to create a RESTful specification that enables interoperability among plant breeding databases -> standard web service
Enabling improvements of data services

Data endpoints

INRA, WUR

VIB, Pt

EBI...

Data Harvester

Data Cache

Search interface

Web Service

Batch Download

Data analysis

Funded by the Horizon 2020 Framework Programme of the European Union
Conclusions and perspectives

- Light tool that connect already existing databases to a single search tool
- Builds a community working together towards making data «machine actionnable »: FAIR
Acknowledgements

URGI team

National and International infrastructures/initiatives

National and international Wheat projects

Financial supports

H. Quesneville
C. Pommier
M. Alaux
M. Buy
D. Charruaud
G. Cornut
J. Destin
S. Diagne
S. Durand
B. El-Houdaigui
R. Flores
C. Guerche
E. Kimmel
T. Letellier
C. Michotey
N. Mohellibi

International Wheat Genome Sequencing Consortium

BrAPI

Wheat Initiative

openMINED

PHENOME

transPLANT

elixir

National and International Wheat projects

Bioversity International

Planteome

COST Action

Planteome

INRA Science & Impact

Research Data Alliance

ANR

COST Action

National and International Wheat projects

Cost

COST Action

INTEGRAPE
Thank you!