Interopérabilité des données en sciences de la vie

Contexte, ressources et cas d'utilisation

Michaël Alaux & Anne-Françoise Adam-Blondon

Les données de recherche en sciences de la vie

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'ESPACE
Liberté
Égalité
Fraternité

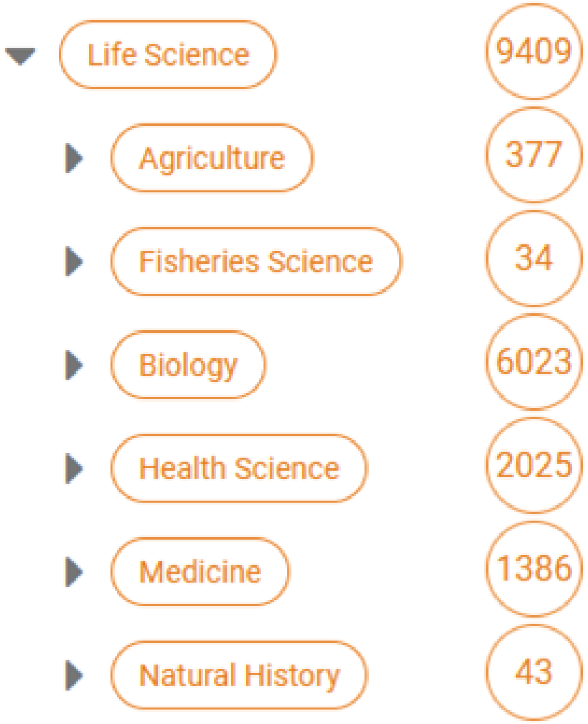
université
PARIS-SACLAY

FRANCE 2030

Des données hétérogènes et dispersées

- Des domaines d'études très divers
  - Nombre de taxons étudiés : santé / biodiversité
  - Type de données collectées : génomiques, phénotypiques, cliniques, climatiques, …
  - Types d'analyses à effectuer
  - Degré de sensibilité des données
- Des sous-communautés partageant des objectifs de recherche
  - silos en termes de vocabulaires, ontologies, sémantique associée
- Ex: FAIRsharing liste actuellement 9 409 standards/bases de données/guidelines en lien avec les sciences de la vie au niveau international

FAIRsharing.org
standards, databases, policies



Life Science	9409
Agriculture	377
Fisheries Science	34
Biology	6023
Health Science	2025
Medicine	1386
Natural History	43

Un grand nombre de « communs »

Développés par des communautés représentant des sous domaines des sciences de la vie et formant un écosystème plus ou moins connecté

- **Des entrepôts centralisés pour les données de séquence** dès la fin des années 80 : *International Nucleotide Sequence Database Collaboration* (INSDC)
  - DDBJ (Japon), ENA (Europe), NCBI (USA)
  - Recommandations et implémentation de formats et standards de données
  - Réplication des données entre les 3 entrepôts
- **Développement de jeux de données de référence curés, d'ontologies, vocabulaires contrôlés**
  - Ex: Catalogue of Life → Agrégation de 16 954 jeux de données de taxonomie grâce au travail de plusieurs centaines d'experts

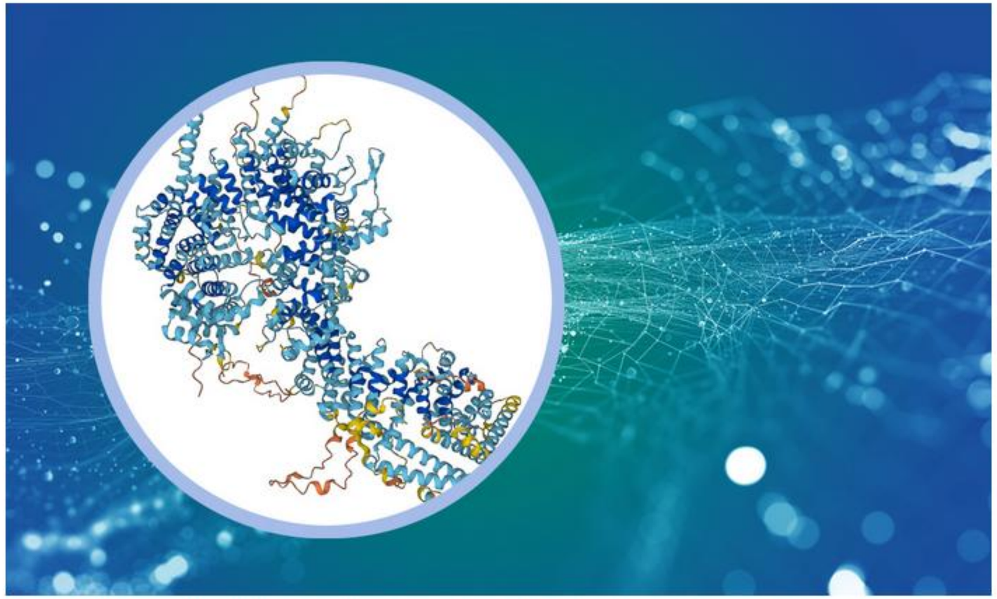
Un grand nombre de « communs »

- **Développement de fédération de données** (données non centralisables): interopérabilité des données et des systèmes d’information
  - Ex: données de phénotypage chez les plantes, données génomiques en santé humaine
- **Développement de portails de connaissances avec des spectres très divers**
  - Un large spectre taxonomique (Ex: InterPRO) ou au contraire spécialisés sur un taxon (Ex: FlyBase ou TAIR)

Ces communs sont un terreau important pour la recherche et l'innovation

- Rapports de l'OCDE en 2015 et 2021
- Rapport Draghi en 2024 dans le contexte de l'IA
  Des données *FAIR* aux données *AI ready*
- En 2025
  - changement de politique des Etats-Unis sur les données publiques → prise de conscience de l'importance d'avoir des ressources en données souveraines en Europe
  - L'essor spectaculaire de l'IA générative fait émerger des questions sur les données ouvertes

Open data stored at EMBL-EBI played a pivotal role in the development of the AlphaFold AI.



Credit: Nuclear pore complex prediction by AlphaFold. Edited by Karen Arnott/EMBL-EBI. Background image from Adobe Stock Images.

Challenge: Predicting how proteins fold

Défis actuels

- Comment identifier et gérer les « communs » de données critiques pour l’Europe ?
- Comment améliorer ces communs : interopérabilité des données et des systèmes d’information

Des infrastructures/initiatives pour les données en sciences de la vie en France et en Europe

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'ESPACE
Liberté
Égalité
Fraternité

université
PARIS-SACLAY

FRANCE 2030

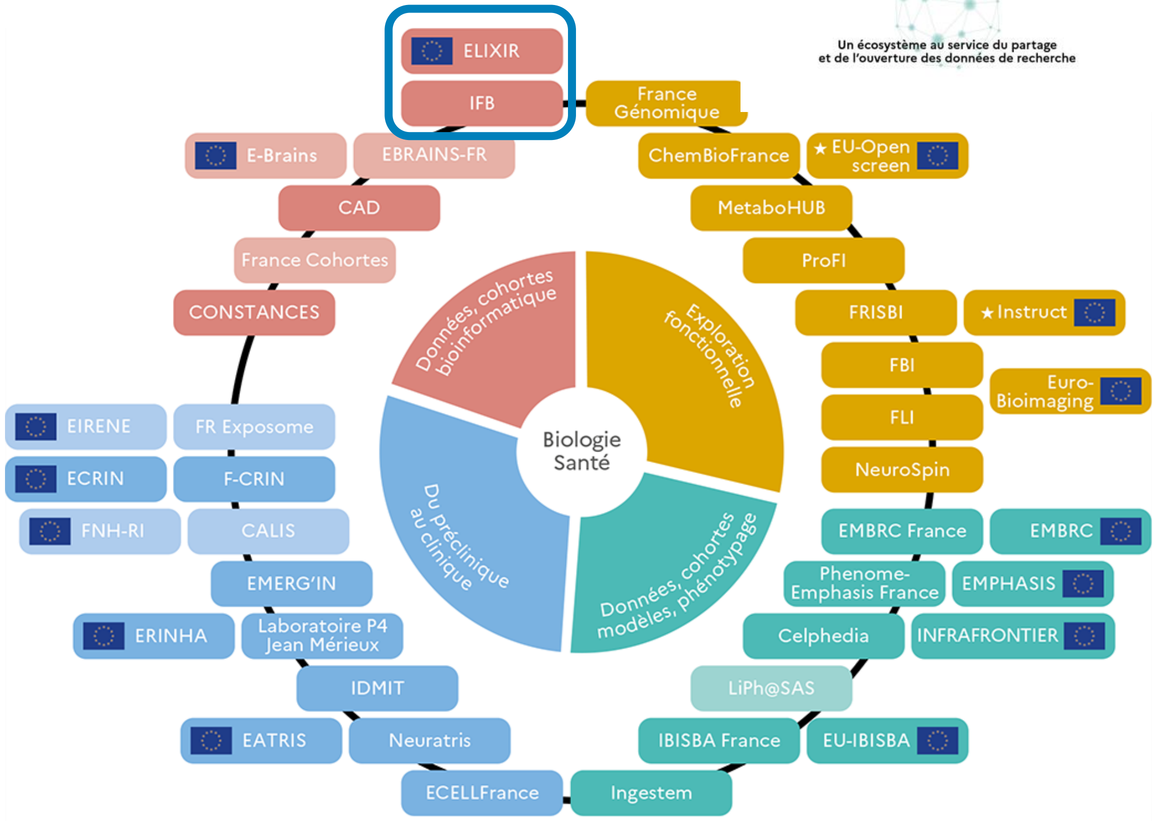
L'Institut National de Bioinformatique

INSTITUT FRANÇAIS DE BIOINFORMATIQUE

- L'IFB/ELIXIR-FR est une **Infrastructure de Recherche Nationale de Biologie-Santé** distribuée (20 plateformes membre)
- 2022 - Attribution du rôle de **Centre de compétence disciplinaire pour la biologie-santé** dans l'écosystème Recherche Data Gouv
- Des collaborations importantes avec l'Infrastructure **Data Terra** (sciences de la terre):
  - **Biodiversité**
  - **Milieux anthropisés** (ex: agriculture)

RÉPUBLIQUE FRANÇAISE recherche.data.gouv.fr

Un écosystème au service du partage et de l'ouverture des données de recherche



ELIXIR: Infrastructure de Recherche Européenne de données en Sciences de la Vie

What?

A European life science data infrastructure

An intergovernmental consortium

25 noeuds nationaux + EMBL-EBI

How?

connecting and developing a network of experts

providing hundreds of interoperable, high-quality services and resources

sustaining a distributed infrastructure, freely available to all

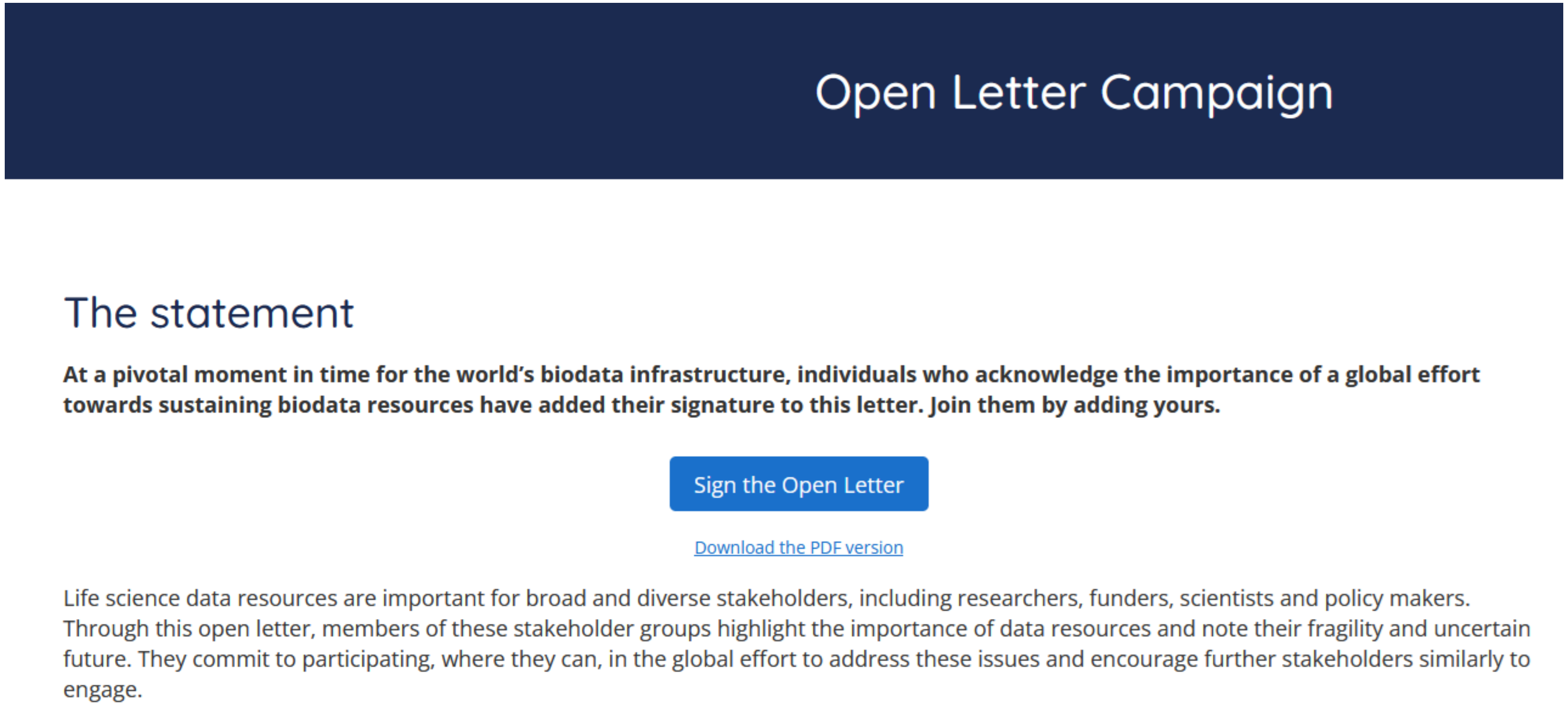
Why?

to accelerate life science research and its impact on society

IFB = elixir FRANCE

Global Biodata Coalition

Un consortium international pour sécuriser les données communes



Open Letter Campaign

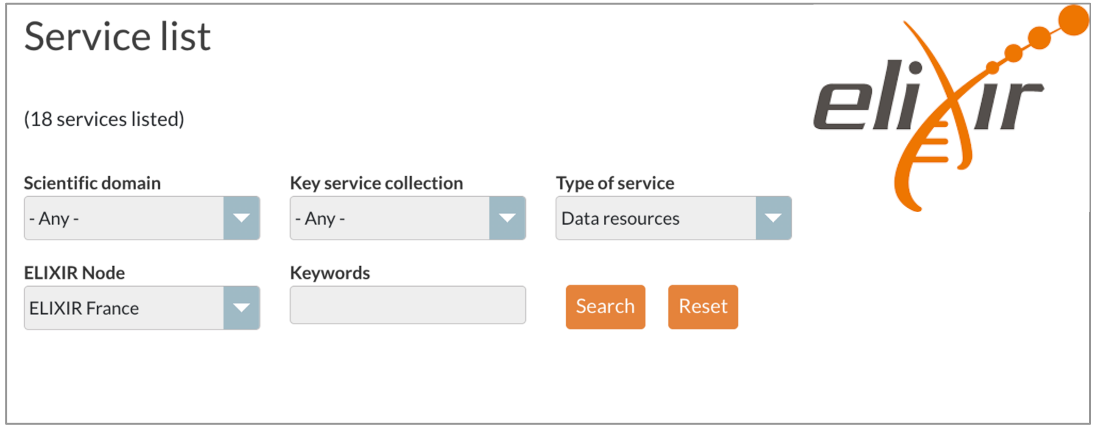
The statement

At a pivotal moment in time for the world's biodata infrastructure, individuals who acknowledge the importance of a global effort towards sustaining biodata resources have added their signature to this letter. Join them by adding yours.

Life science data resources are important for broad and diverse stakeholders, including researchers, funders, scientists and policy makers. Through this open letter, members of these stakeholder groups highlight the importance of data resources and note their fragility and uncertain future. They commit to participating, where they can, in the global effort to address these issues and encourage further stakeholders similarly to engage.

Systèmes de labellisation pour lister et qualifier les ressources importantes

Interopérabilité des données et des systèmes d'information entre CDR = critère important de labellisation



Service list

(18 services listed)

Scientific domain: - Any -

Key service collection: - Any -

Type of service: Data resources

ELIXIR Node: ELIXIR France

Keywords:

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

recherche.data.gouv.fr

Recherche Data Gouv: 14 entrepôts de confiance

IFB/ELIXIR-FR: 14 ressources en données sur son **Service Delivery Plan**

ELIXIR-FR: 2 ressources labellisées **ELIXIR Core Data Ressource (CDR)**

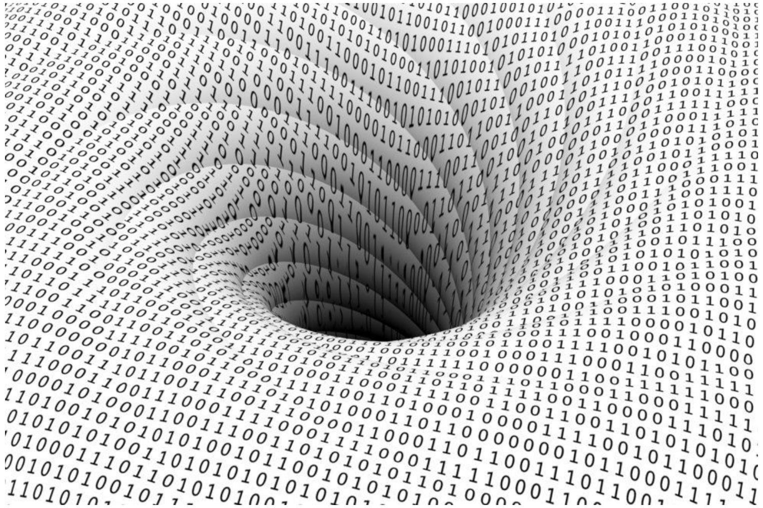
elixir Core Data Resource

elixir Deposition Database

2 Ressources françaises labellisées **Global Core Biodata Resource**

GBC GLOBAL CORE BIODATA RESOURCE

Comment favoriser l’alimentation de ces bases de données avec des données FAIR par l’ensemble des chercheurs Européens ?



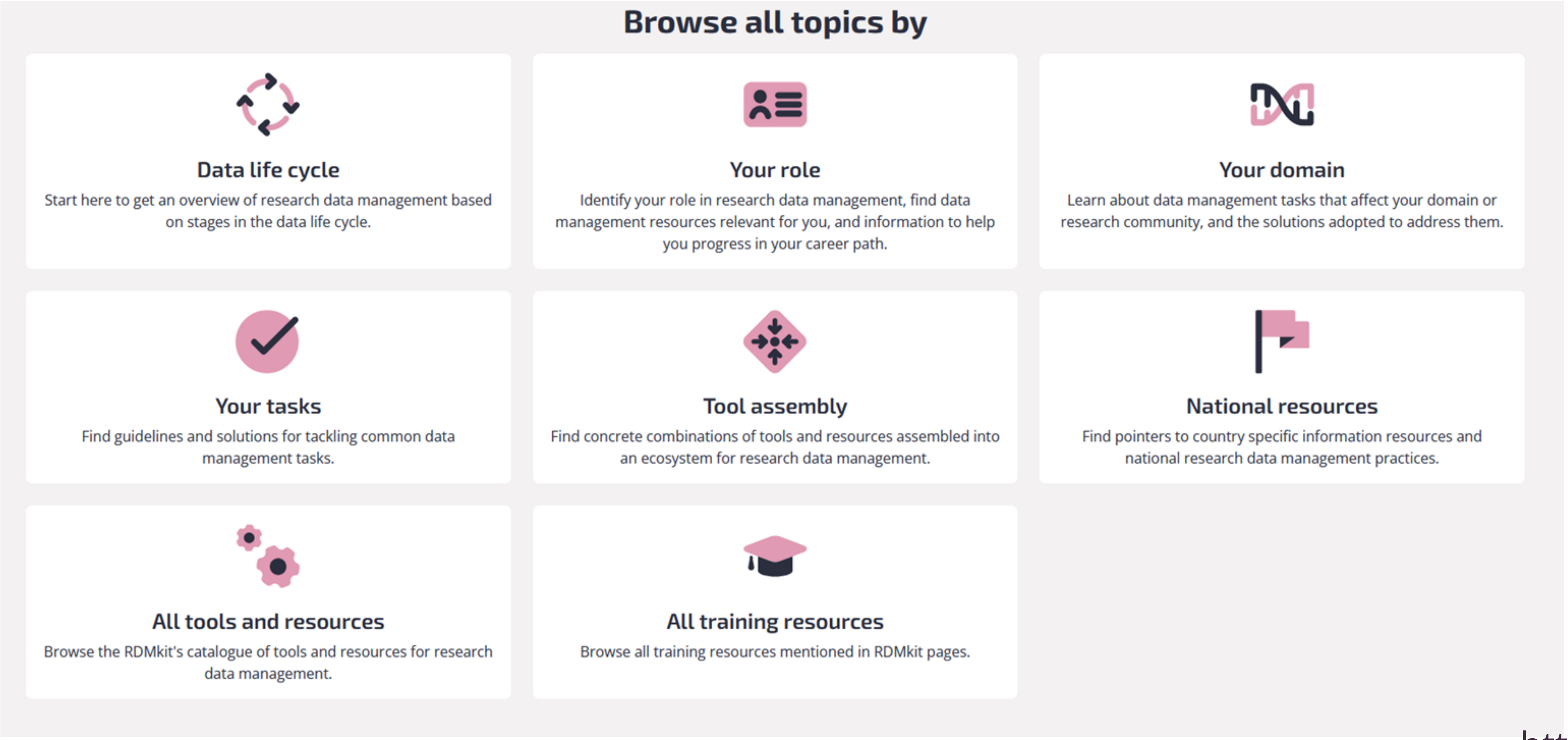
Focus sur les ressources support à la gestion FAIR de données

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'ESPACE
Liberté
Égalité
Fraternité

université
PARIS-SACLAY

FRANCE 2030

RDMkit: *The ELIXIR Research Data Management toolkit for Life Sciences*



Browse all topics by

Data life cycle
Start here to get an overview of research data management based on stages in the data life cycle.

Your role
Identify your role in research data management, find data management resources relevant for you, and information to help you progress in your career path.

Your domain
Learn about data management tasks that affect your domain or research community, and the solutions adopted to address them.

Your tasks
Find guidelines and solutions for tackling common data management tasks.

Tool assembly
Find concrete combinations of tools and resources assembled into an ecosystem for research data management.

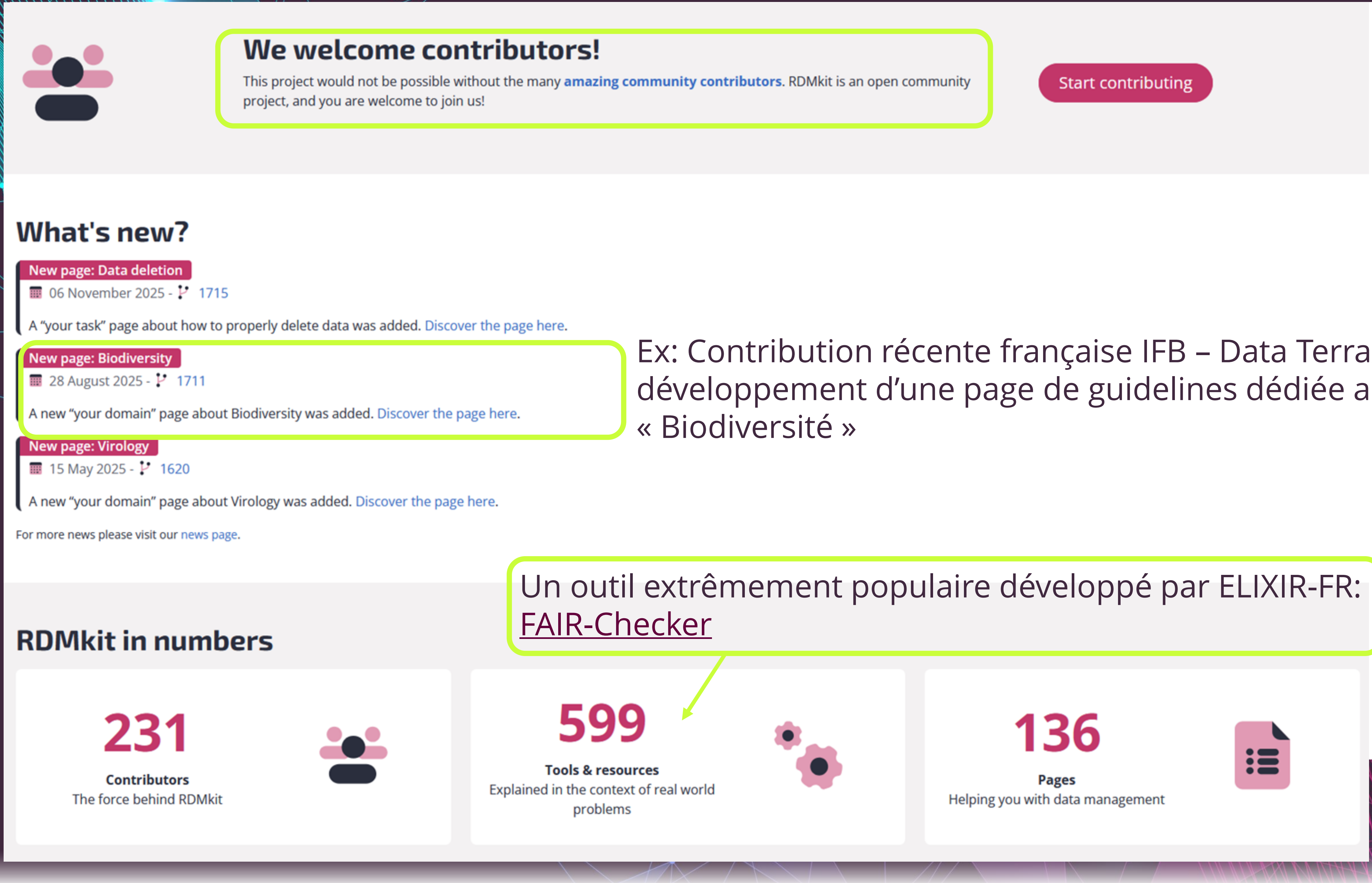
National resources
Find pointers to country specific information resources and national research data management practices.

All tools and resources
Browse the RDMkit's catalogue of tools and resources for research data management.

All training resources
Browse all training resources mentioned in RDMkit pages.

Linked to other registries:
- FAIRCookbook
- FAIRsharing
- TeSS
- Bio.tools
- WorkflowHub
- Data Stewardship Wizard

https://rdmkit.elixir-europe.org



We welcome contributors!

This project would not be possible without the many **amazing community contributors**. RDMkit is an open community project, and you are welcome to join us!

What's new?

New page: Data deletion

06 November 2025 - 1715

A "your task" page about how to properly delete data was added. Discover the page here.

New page: Biodiversity

28 August 2025 - 1711

A new "your domain" page about Biodiversity was added. Discover the page here.

New page: Virology

15 May 2025 - 1620

A new "your domain" page about Virology was added. Discover the page here.

For more news please visit our news page.

RDMkit in numbers

231
Contributors
The force behind RDMkit

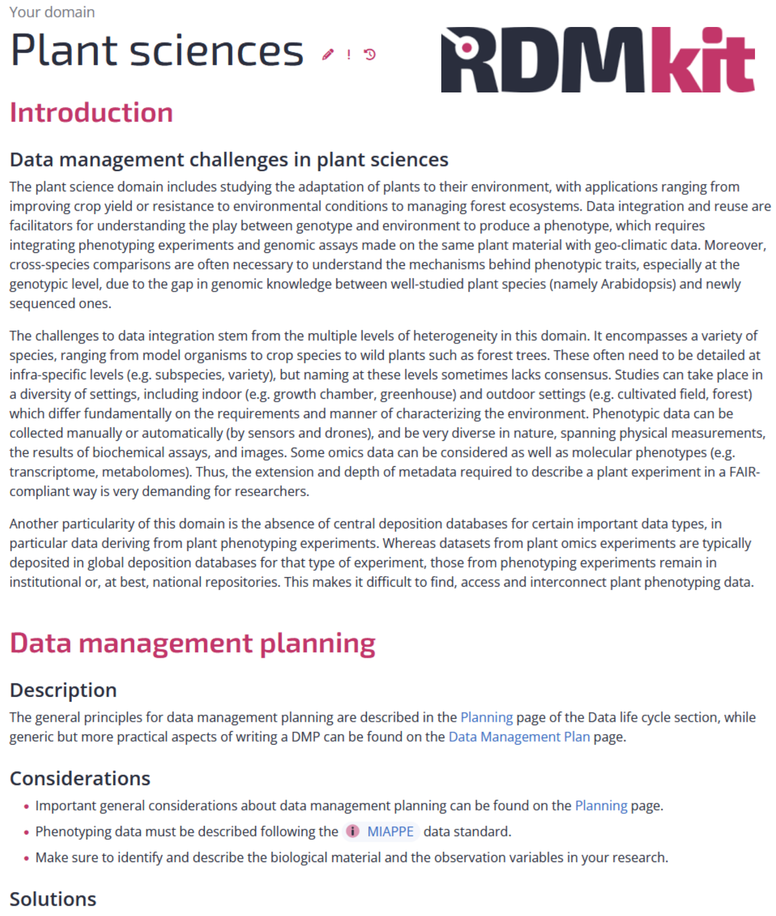
599
Tools & resources
Explained in the context of real world problems

136
Pages
Helping you with data management

Ex: Contribution récente française IFB – Data Terra/PNDB au développement d'une page de guidelines dédiée au domaine « Biodiversité »

Un outil extrêmement populaire développé par ELIXIR-FR: FAIR-Checker

Des ressources support au support de projets



Your domain

Plant sciences

RDMkit

Introduction

Data management challenges in plant sciences

The plant science domain includes studying the adaptation of plants to their environment, with applications ranging from improving crop yield or resistance to environmental conditions to managing forest ecosystems. Data integration and reuse are facilitators for understanding the play between genotype and environment to produce a phenotype, which requires integrating phenotyping experiments and genomic assays made on the same plant material with geo-climatic data. Moreover, cross-species comparisons are often necessary to understand the mechanisms behind phenotypic traits, especially at the genotypic level, due to the gap in genomic knowledge between well-studied plant species (namely Arabidopsis) and newly sequenced ones.

The challenges to data integration stem from the multiple levels of heterogeneity in this domain. It encompasses a variety of species, ranging from model organisms to crop species to wild plants such as forest trees. These often need to be detailed at infra-specific levels (e.g. subspecies, variety), but naming at these levels sometimes lacks consensus. Studies can take place in a diversity of settings, including indoor (e.g. growth chamber, greenhouse) and outdoor settings (e.g. cultivated field, forest) which differ fundamentally on the requirements and manner of characterizing the environment. Phenotypic data can be collected manually or automatically (by sensors and drones), and be very diverse in nature, spanning physical measurements, the results of biochemical assays, and images. Some omics data can be considered as well as molecular phenotypes (e.g. transcriptome, metabolomes). Thus, the extension and depth of metadata required to describe a plant experiment in a FAIR-compliant way is very demanding for researchers.

Another particularity of this domain is the absence of central deposition databases for certain important data types, in particular data deriving from plant phenotyping experiments. Whereas datasets from plant omics experiments are typically deposited in global deposition databases for that type of experiment, those from phenotyping experiments remain in institutional or, at best, national repositories. This makes it difficult to find, access and interconnect plant phenotyping data.

Data management planning

Description

The general principles for data management planning are described in the Planning page of the Data life cycle section, while generic but more practical aspects of writing a DMP can be found on the Data Management Plan page.

Considerations

- Important general considerations about data management planning can be found on the Planning page.
- Phenotyping data must be described following the MIAPPE data standard.
- Make sure to identify and describe the biological material and the observation variables in your research.

Solutions

Exemple du domaine « Plantes »

- Ressources support à la gestion FAIR des données sur les plantes développées depuis 2015 dans un cadre international : Wheat Initiative (collaboration RDA), l'ELIXIR Plant Sciences Community, EMPHASIS, BrAPI consortium, CGIAR, MIAPPE consortium
- Compilation dans la page "Plant Science" du RDMkit
  - ➢ « *Crash test* » et amélioration dans des projets Internationaux, Européens et nationaux

Cas d’utilisation dans le cadre de projets et consortiums

MINISTÈRE
DE L’ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L’ESPACE
Liberté
Égalité
Fraternité

université
PARIS-SACLAY

FRANCE 2030

L’URGI en bref

Unité Ressources Génomique-Info

- Unité de service INRAE (centre Versailles-Saclay) en bioinformatique pour les plantes d’intérêt agronomique et forestières
- Membre de l’Institut Français de Bioinformatique (IR nationale)
- Forte implication à l’Europe et l’international dans le développement de standards et recommandations pour la gestion des données “plantes” selon les principes FAIR
- Partenaire de nombreux projets nationaux et internationaux pour la gestion de données : 3 exemples récents montrés ici

Les services développés ou mobilisés autour de la donnée

- Development and maintenance of metadata standards and ontologies for plant phenotyping data (including forests)
- Development of guidelines for data management according to FAIR principles

RDMkit miappe Crop Ontology for agricultural data

- Data discovery portals: international (FAIDARE, WheatIS) and national (RARe)
- Contribution to the development of an international web services standard: BrAPI

FAIDARE wheatIS WHEAT INFORMATION SYSTEM RARe BrAPI

- GnpIS Information System: INRAE genetic resource data and phenotyping data
- Support for FAIR datasets submission to Recherche Data Gouv

GnpIS

Supporting infrastructures and major projects with their data management plan

DMP OPIDoR HAL science ouverte

- Provision of reference data (genetic resources, identifiers/DOI, consensus of transposable elements, etc.)
- Provision of templates for metadata collection

- Provision of a space for sharing and curating metadata (FAIRDOM-SEEK, Nextcloud)
- Metadata curation and enrichment

- Reproducible annotation of repeated elements in genomes (REPET Factory)
- Provision of containerized pipelines: REPET
- Provision of Virtual Research Environments

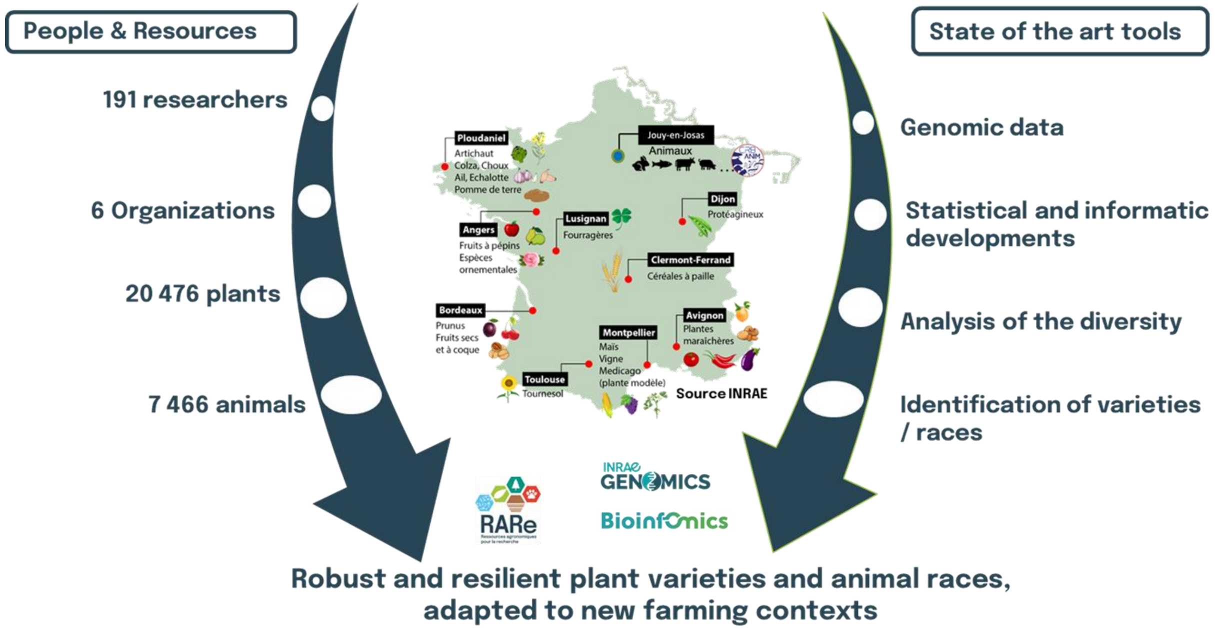
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'ESPACE

FRANCE 2030

Les projets BReIF et AgroDIV du PEPR AgroEcoNum

Principaux objectifs de BReIF:

- Interopérabilité des services numériques des infrastructures de recherche française dans le domaine des sciences du vivants (RARe, IFB, France génomique)
- Amélioration de la gestion FAIR des données



People & Resources

191 researchers

6 Organizations

20 476 plants

7 466 animals

State of the art tools

Genomic data

Statistical and informatic developments

Analysis of the diversity

Identification of varieties / races

Source INRAE

Robust and resilient plant varieties and animal races, adapted to new farming contexts

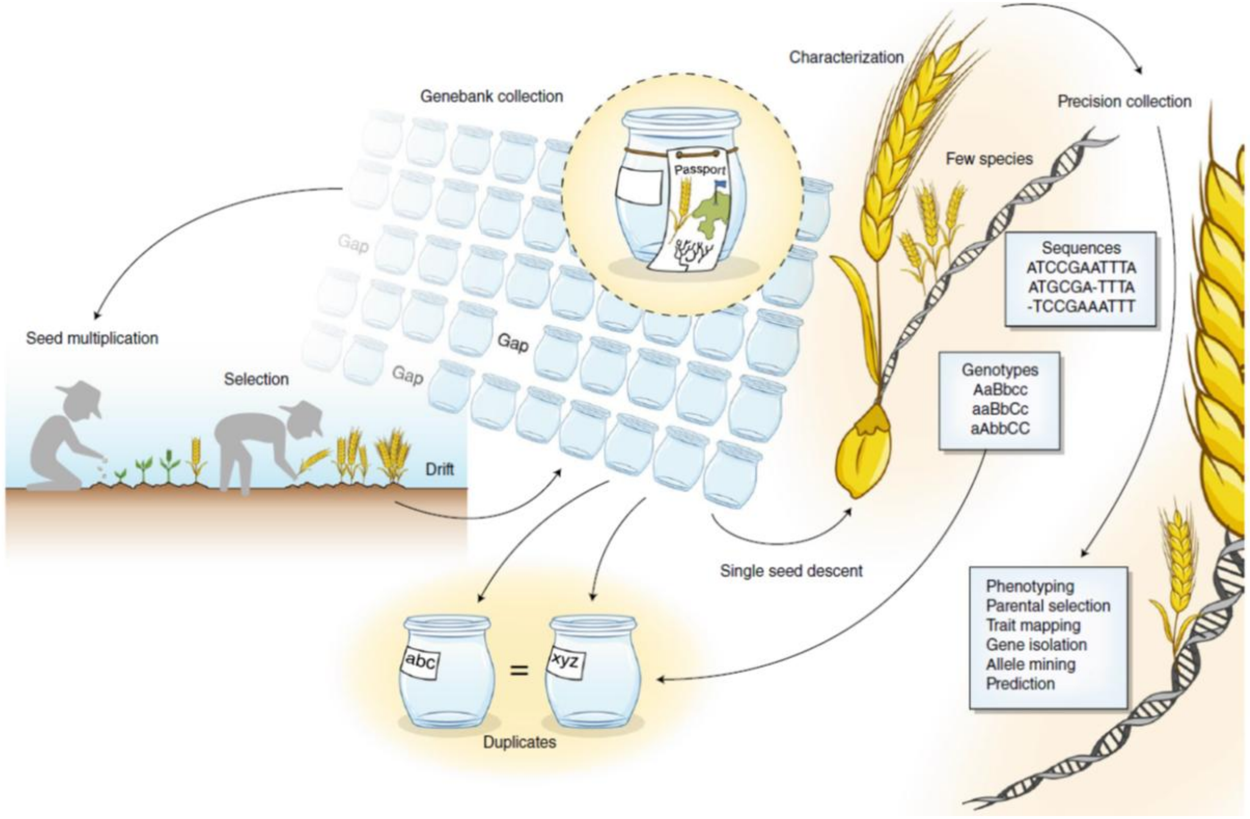
Les projets européens sur le blé

AGENT

caractérisation d’un grand nombre de collection de blé et orge des centres de ressources biologiques européens

PRO-WILD
Protect & Promote Crop Wild Relatives

sur espèces sauvages apparentés aux espèces cultivés

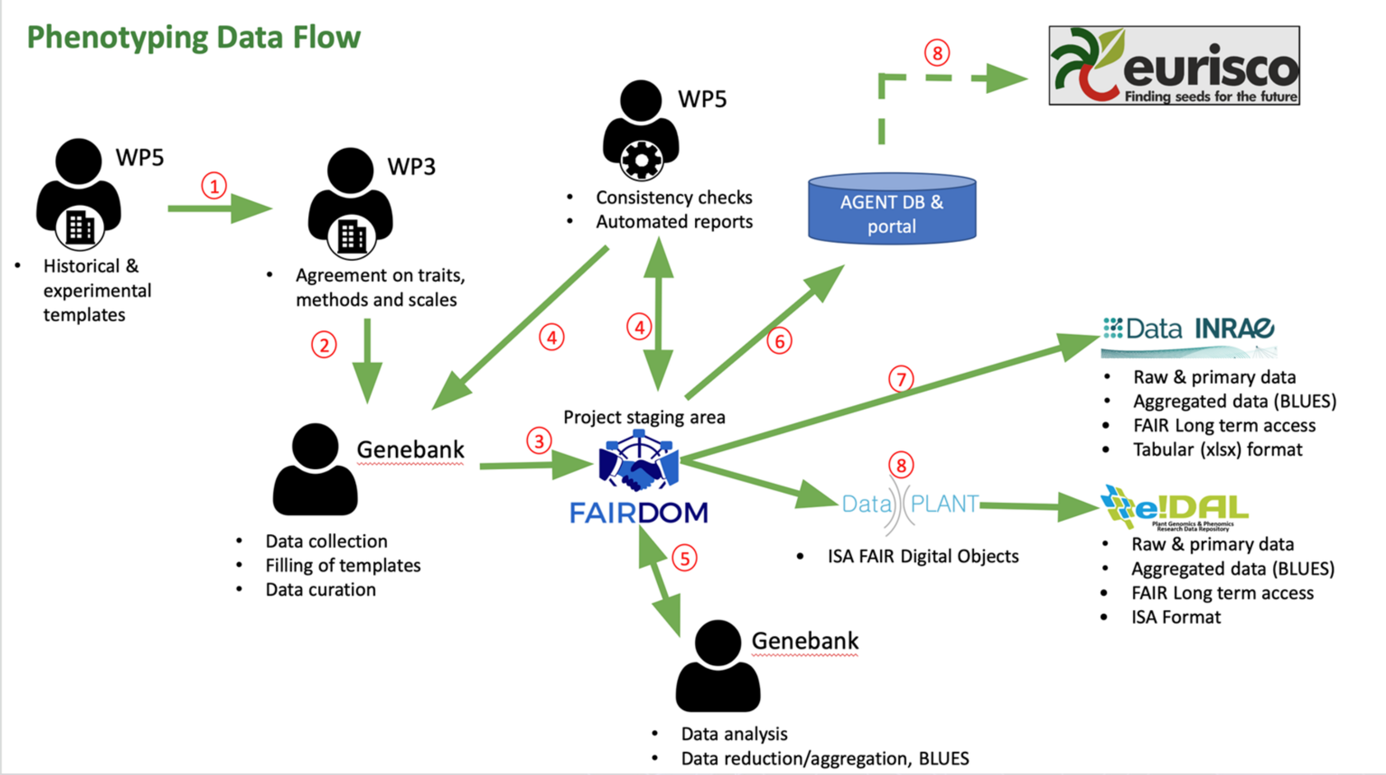


Mascher et al. Nature Genetics

Guidelines

- Mise à disposition de guidelines pour gérer les données du projet
  - Complémentaire du plan de gestion de données
  - Flux de données par type de données
  - Formats, standards de données et outils à utiliser
  - Référence aux sources externes (RDMkit, FAIRCookbook)

elixir
PHENOME EMPHASIS
FRANCE GENOMIQUE
RARe

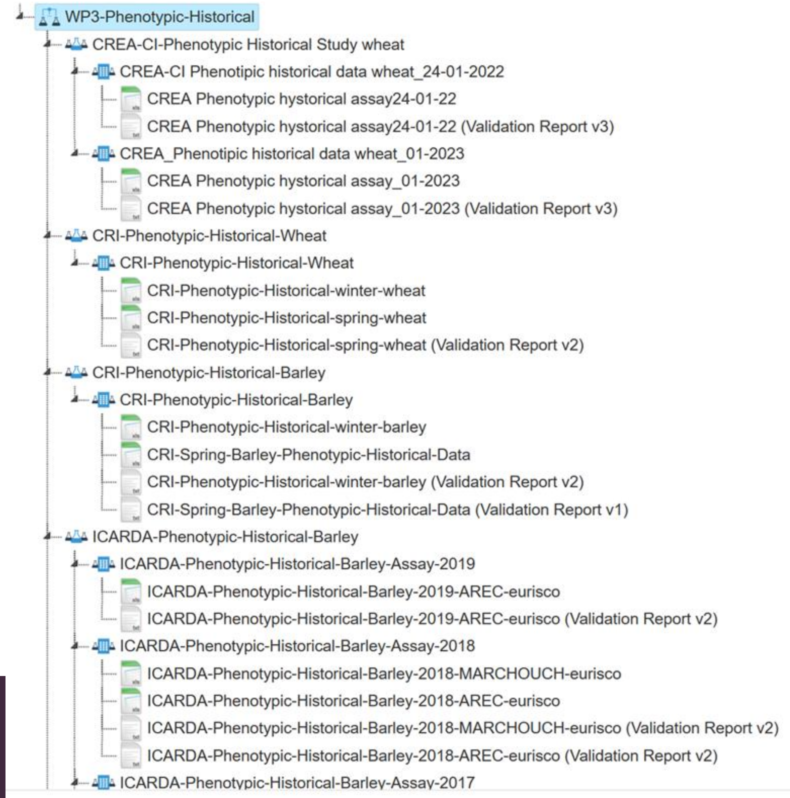


- Historical & experimental templates
- Agreement on traits, methods and scales
- Consistency checks
- Automated reports
- Raw & primary data
- Aggregated data (BLUES)
- FAIR Long term access
- Tabular (xlsx) format
- Data collection
- Filling of templates
- Data curation
- ISA FAIR Digital Objects
- Raw & primary data
- Aggregated data (BLUES)
- FAIR Long term access
- ISA Format
- Data analysis
- Data reduction/aggregation, BLUES

https://doi.org/10.5281/zenodo.126
59

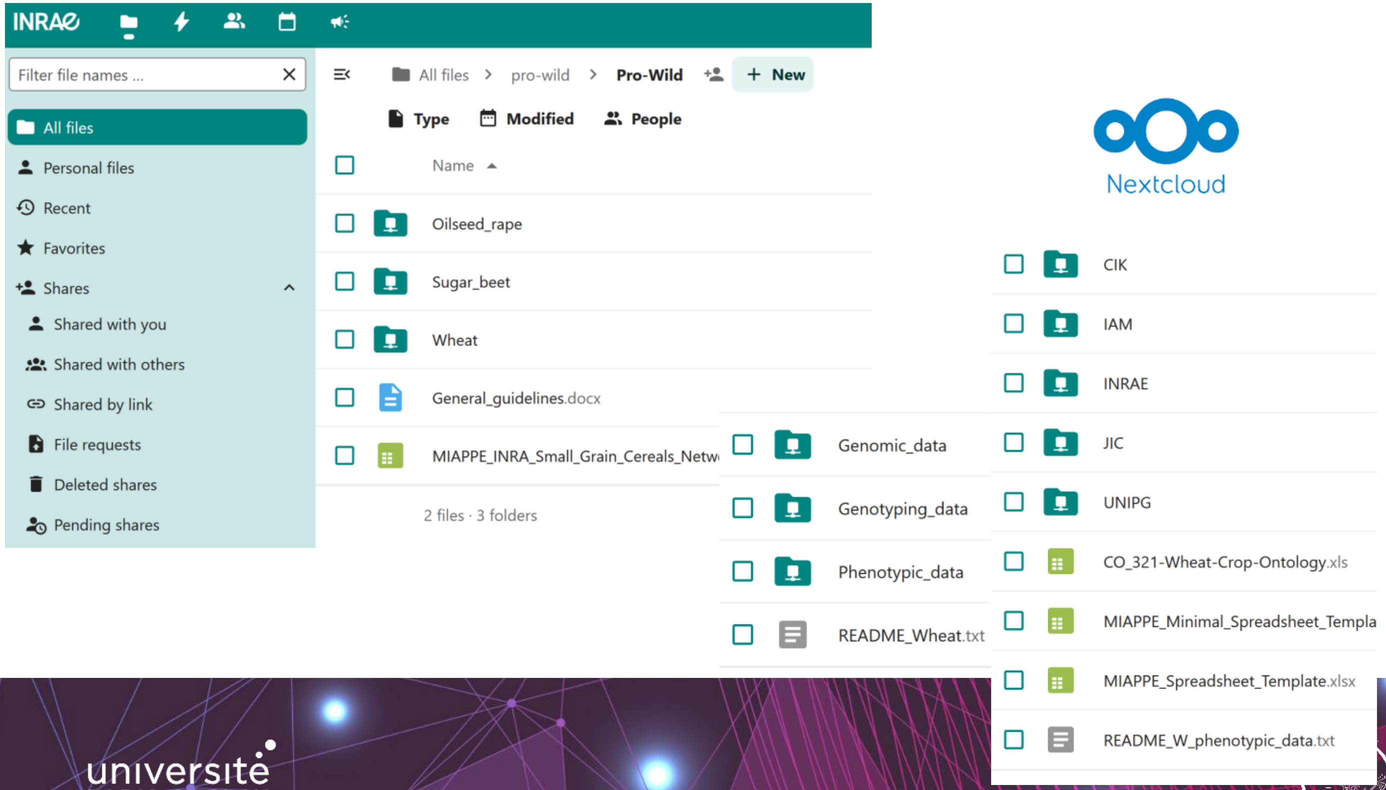
Outils de partage et curation

- Espace de partage, curation et validation de données
  - arborescence
  - contient templates, fichiers d'exemple, guidelines, README



FAIRDOM

- WP3-Phenotypic-Historical
  - CREA-CI-Phenotypic Historical Study wheat
    - CREA-CI Phenotipic historical data wheat_24-01-2022
      - CREA Phenotypic hystorical assay24-01-22
      - CREA Phenotypic hystorical assay24-01-22 (Validation Report v3)
    - CREA_Phenotipic historical data wheat_01-2023
      - CREA Phenotypic hystorical assay_01-2023
      - CREA Phenotypic hystorical assay_01-2023 (Validation Report v3)
  - CRI-Phenotypic-Historical-Wheat
    - CRI-Phenotypic-Historical-Wheat
      - CRI-Phenotypic-Historical-winter-wheat
      - CRI-Phenotypic-Historical-spring-wheat
      - CRI-Phenotypic-Historical-spring-wheat (Validation Report v2)
  - CRI-Phenotypic-Historical-Barley
    - CRI-Phenotypic-Historical-Barley
      - CRI-Phenotypic-Historical-winter-barley
      - CRI-Spring-Barley-Phenotypic-Historical-Data
      - CRI-Phenotypic-Historical-winter-barley (Validation Report v2)
      - CRI-Spring-Barley-Phenotypic-Historical-Data (Validation Report v1)
  - ICARDA-Phenotypic-Historical-Barley
    - ICARDA-Phenotypic-Historical-Barley-Assay-2019
      - ICARDA-Phenotypic-Historical-Barley-2019-AREC-eurisco
      - ICARDA-Phenotypic-Historical-Barley-2019-AREC-eurisco (Validation Report v2)
    - ICARDA-Phenotypic-Historical-Barley-Assay-2018
      - ICARDA-Phenotypic-Historical-Barley-2018-MARCHOUCH-eurisco
      - ICARDA-Phenotypic-Historical-Barley-2018-AREC-eurisco
      - ICARDA-Phenotypic-Historical-Barley-2018-MARCHOUCH-eurisco (Validation Report v2)
      - ICARDA-Phenotypic-Historical-Barley-2018-AREC-eurisco (Validation Report v2)
    - ICARDA-Phenotypic-Historical-Barley-Assay-2017



☐ Oilseed_rape
☐ Sugar_beet
☐ Wheat
☐ General_guidelines.docx
☐ MIAPPE_INRA_Small_Grain_Cereals_Netw

2 files · 3 folders

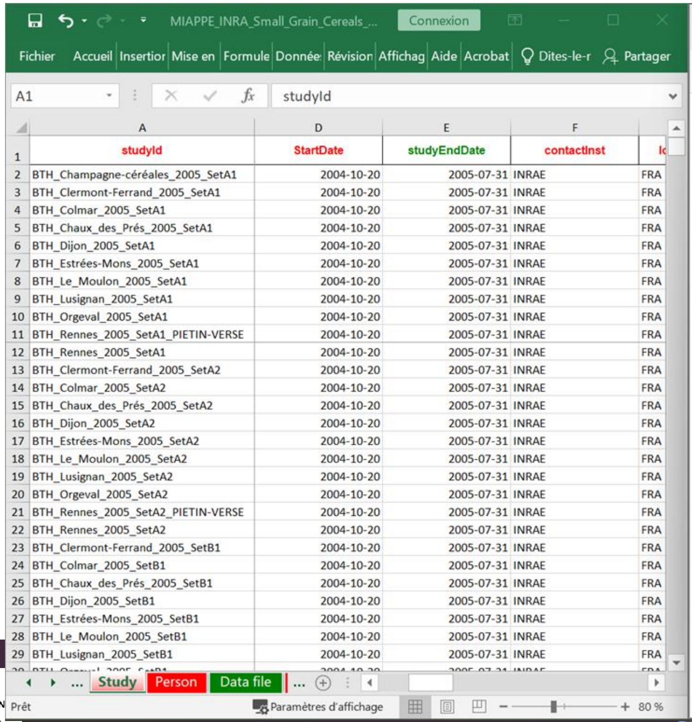
☐ Genomic_data
☐ Genotyping_data
☐ Phenotypic_data
☐ README_Wheat.txt

☐ CIK
☐ IAM
☐ INRAE
☐ JIC
☐ UNIPG
☐ CO_321-Wheat-Crop-Ontology.xls
☐ MIAPPE_Minimal_Spreadsheet_Templa
☐ MIAPPE_Spreadsheet_Template.xlsx
☐ README_W_phenotypic_data.txt

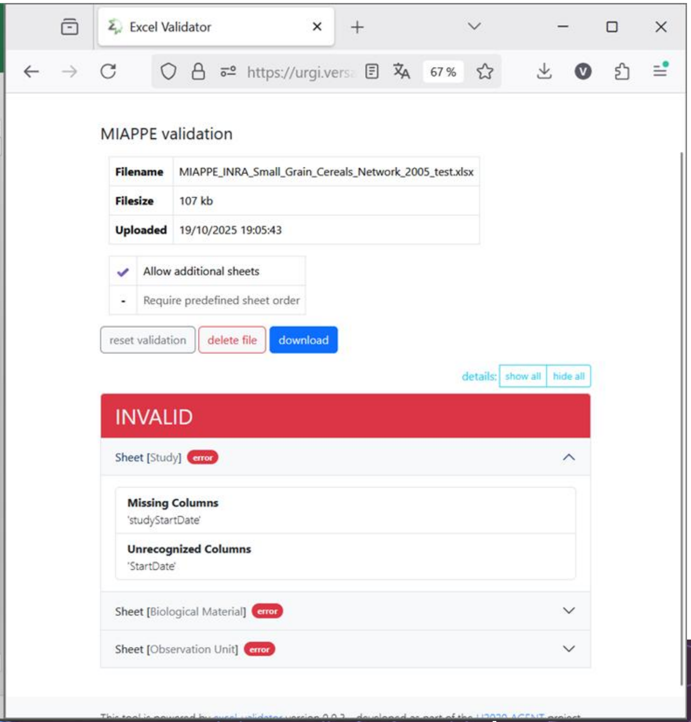
Outils de partage et curation

- Outil d’aide à la curation de métadonnées de phénotypage
  - version en ligne https://urgi.versailles.inrae.fr/excel-validator/miappe
  - lancement automatique dès qu’un nouveau fichier est déposé et écriture de rapport

Frictionless Data



	A	D	E	F	Ic
1	studyId	StartDate	studyEndDate	contactInst	Ic
2	BTH_Champagne-céréales_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
3	BTH_Clermont-Ferrand_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
4	BTH_Colmar_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
5	BTH_Chaux_des_Prés_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
6	BTH_Dijon_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
7	BTH_Estrées-Mons_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
8	BTH_Le_Moulon_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
9	BTH_Lusignan_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
10	BTH_Orgeval_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
11	BTH_Rennes_2005_SetA1_PIETIN-VERSE	2004-10-20	2005-07-31	INRAE	FRA
12	BTH_Rennes_2005_SetA1	2004-10-20	2005-07-31	INRAE	FRA
13	BTH_Clermont-Ferrand_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
14	BTH_Colmar_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
15	BTH_Chaux_des_Prés_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
16	BTH_Dijon_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
17	BTH_Estrées-Mons_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
18	BTH_Le_Moulon_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
19	BTH_Lusignan_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
20	BTH_Orgeval_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
21	BTH_Rennes_2005_SetA2_PIETIN-VERSE	2004-10-20	2005-07-31	INRAE	FRA
22	BTH_Rennes_2005_SetA2	2004-10-20	2005-07-31	INRAE	FRA
23	BTH_Clermont-Ferrand_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA
24	BTH_Colmar_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA
25	BTH_Chaux_des_Prés_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA
26	BTH_Dijon_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA
27	BTH_Estrées-Mons_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA
28	BTH_Le_Moulon_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA
29	BTH_Lusignan_2005_SetB1	2004-10-20	2005-07-31	INRAE	FRA



MIAPPE validation

Filename	MIAPPE_INRA_Small_Grain_Cereals_Network_2005_test.xlsx
Filesize	107 kb
Uploaded	19/10/2025 19:05:43

✓ Allow additional sheets

\- Require predefined sheet order

INVALID

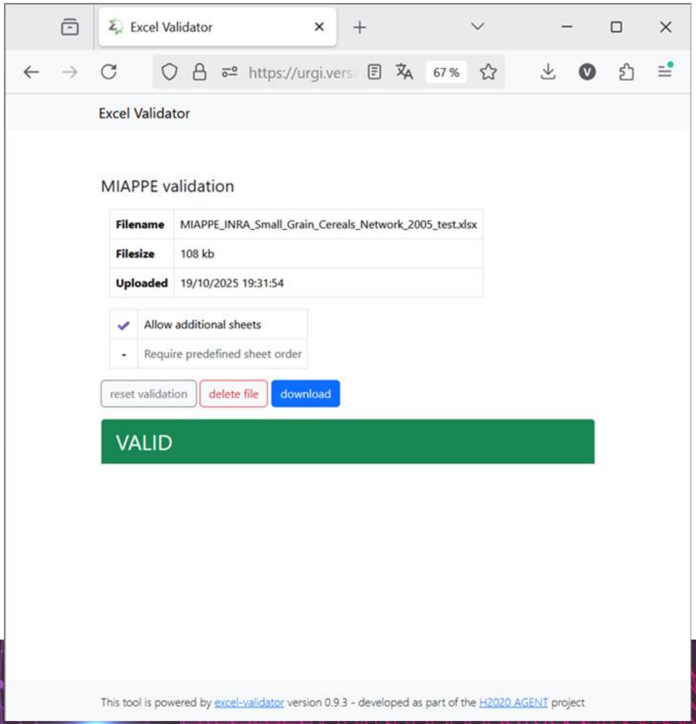
Sheet [Study] error

Missing Columns
'studyStartDate'

Unrecognized Columns
'StartDate'

Sheet [Biological Material] error

Sheet [Observation Unit] error



MIAPPE validation

Filename	MIAPPE_INRA_Small_Grain_Cereals_Network_2005_test.xlsx
Filesize	108 kb
Uploaded	19/10/2025 19:31:54

✓ Allow additional sheets

\- Require predefined sheet order

VALID

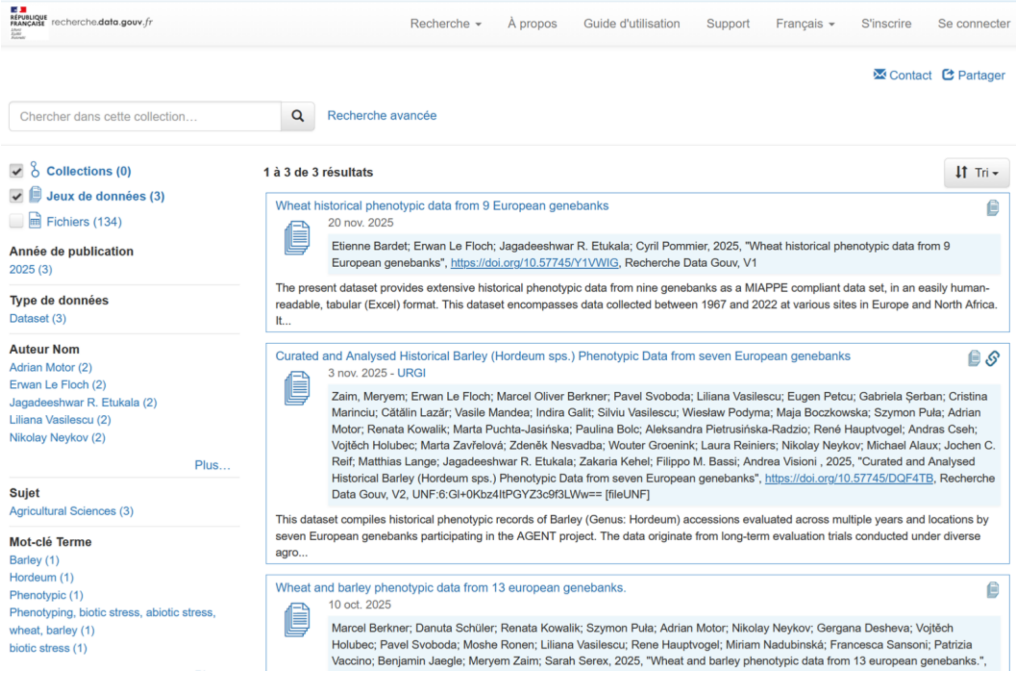
This tool is powered by excel-validator version 0.9.3 - developed as part of the H2020 AGENT project

Aide à la soumission dans des entrepôts

- Disciplinaires: EBI/BioSamples et ENA pour les séquences et échantillons associés → Scripts en accès ouvert pour faciliter la soumission et la récupération d’identifiants

- Recherche Data Gouv: jeux de données de phénotypage curés, validés et analysés

- DataPLANT: FAIR Digital Object (RO-Crate, ARC) → machine actionable



Année de publication
2025 (3)

Type de données
Dataset (3)

Auteur Nom
Adrian Motor (2)
Erwan Le Floch (2)
Jagadeeshwar R. Etukala (2)
Liliana Vasilescu (2)
Nikolay Neykov (2)
Plus…

Sujet
Agricultural Sciences (3)

Mot-clé Terme
Barley (1)
Hordeum (1)
Phenotypic (1)
Phenotyping, biotic stress, abiotic stress, wheat, barley (1)
biotic stress (1)

1 à 3 de 3 résultats

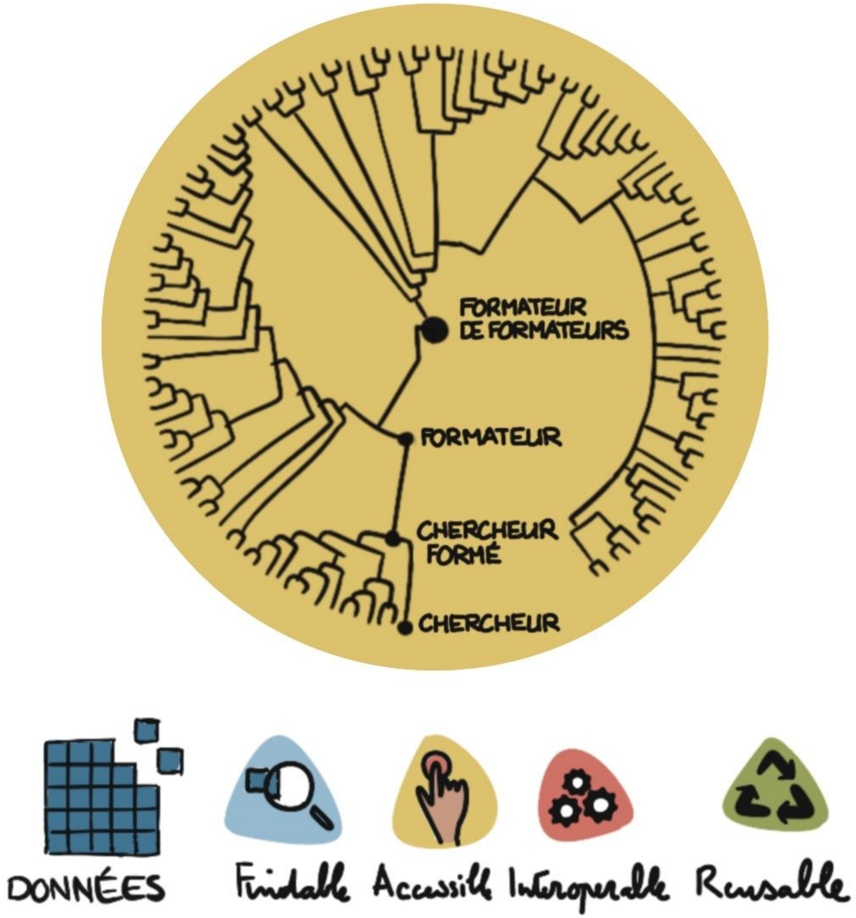
Wheat historical phenotypic data from 9 European genebanks
20 nov. 2025
Etienne Bardet; Erwan Le Floch; Jagadeeshwar R. Etukala; Cyril Pommier, 2025, "Wheat historical phenotypic data from 9 European genebanks", https://doi.org/10.57745/Y1VWIG, Recherche Data Gouv, V1
The present dataset provides extensive historical phenotypic data from nine genebanks as a MIAPPE compliant data set, in an easily human-readable, tabular (Excel) format. This dataset encompasses data collected between 1967 and 2022 at various sites in Europe and North Africa. It...

Curated and Analysed Historical Barley (Hordeum sps.) Phenotypic Data from seven European genebanks
3 nov. 2025 - URGI
Zaim, Meryem; Erwan Le Floch; Marcel Oliver Berkner; Pavel Svoboda; Liliana Vasilescu; Eugen Petcu; Gabriela Şerban; Cristina Marinciu; Cătălin Lazăr; Vasile Mandea; Indira Galit; Silviu Vasilescu; Wiesław Podyma; Maja Boczkowska; Szymon Puła; Adrian Motor; Renata Kowalik; Marta Puchta-Jasińska; Paulina Bolc; Aleksandra Pietrusińska-Radzio; René Hauptvogel; Andras Cseh; Vojtěch Holubec; Marta Zavřelová; Zdeněk Nesvadba; Wouter Groenink; Laura Reiniers; Nikolay Neykov; Michael Alaux; Jochen C. Reif; Matthias Lange; Jagadeeshwar R. Etukala; Zakaria Kehel; Filippo M. Bassi; Andrea Visioni , 2025, "Curated and Analysed Historical Barley (Hordeum sps.) Phenotypic Data from seven European genebanks", https://doi.org/10.57745/DQF4TB, Recherche Data Gouv, V2, UNF:6:GI+0Kbz4ItPGYZ3c9f3LWw== [fileUNF]
This dataset compiles historical phenotypic records of Barley (Genus: Hordeum) accessions evaluated across multiple years and locations by seven European genebanks participating in the AGENT project. The data originate from long-term evaluation trials conducted under diverse agro...

Wheat and barley phenotypic data from 13 european genebanks.
10 oct. 2025
Marcel Berkner; Danuta Schüler; Renata Kowalik; Szymon Puła; Adrian Motor; Nikolay Neykov; Gergana Desheva; Vojtěch Holubec; Pavel Svoboda; Moshe Ronen; Liliana Vasilescu; Rene Hauptvogel; Miriam Nadubinská; Francesca Sansoni; Patrizia Vaccino; Benjamin Jaegle; Meryem Zaim; Sarah Serex, 2025, "Wheat and barley phenotypic data from 13 european genebanks.",

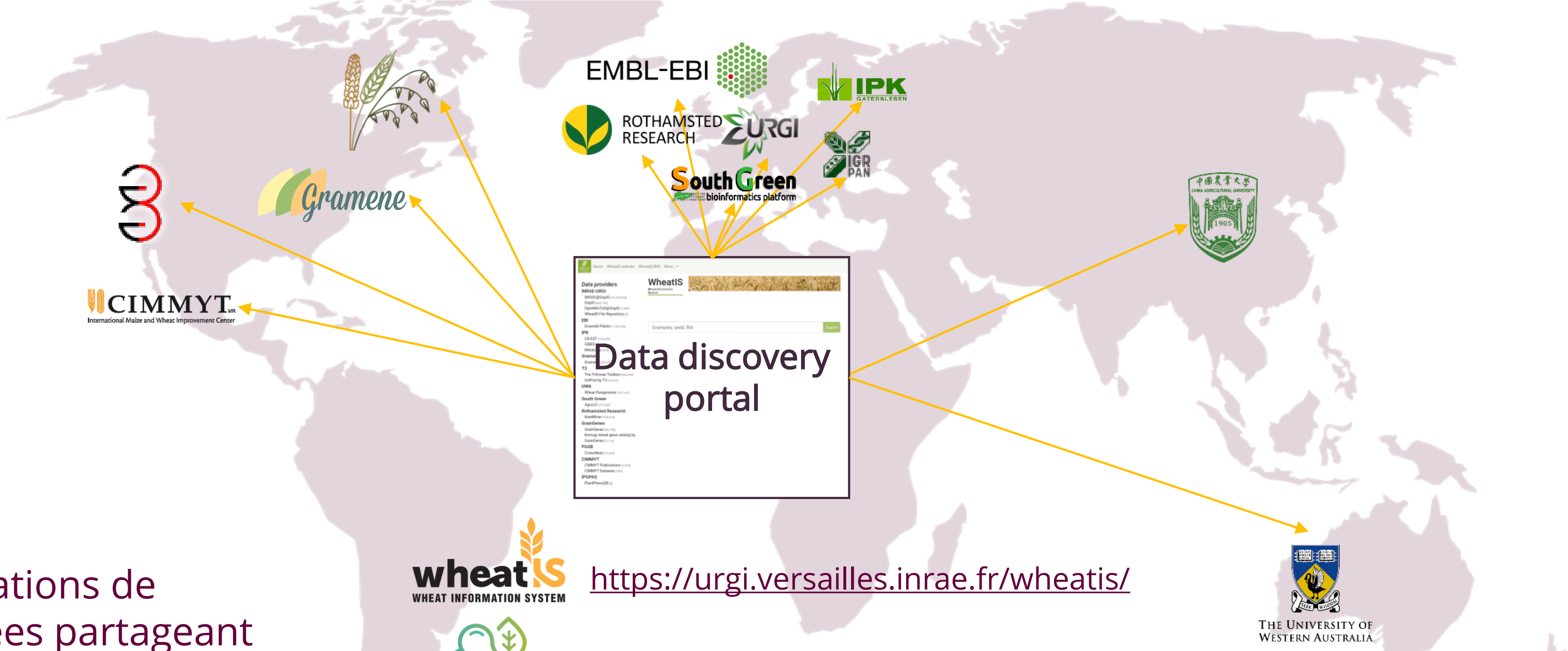
Des formations

- Visibles dans le catalogue de formation de l'IFB qui alimente automatiquement celui d'ELIXIR (TeSS)
- Du matériel de formation réutilisable accessible via une instance Moodle de l'IFB



crédit Jordane Ancelin

Portail de recherche des données scientifiques



Data discovery portal

Fédérations de données partageant les mêmes standards

wheatIS WHEAT INFORMATION SYSTEM https://urgi.versailles.inrae.fr/wheatis/

FAIDARE https://urgi.versailles.inrae.fr/faidare/

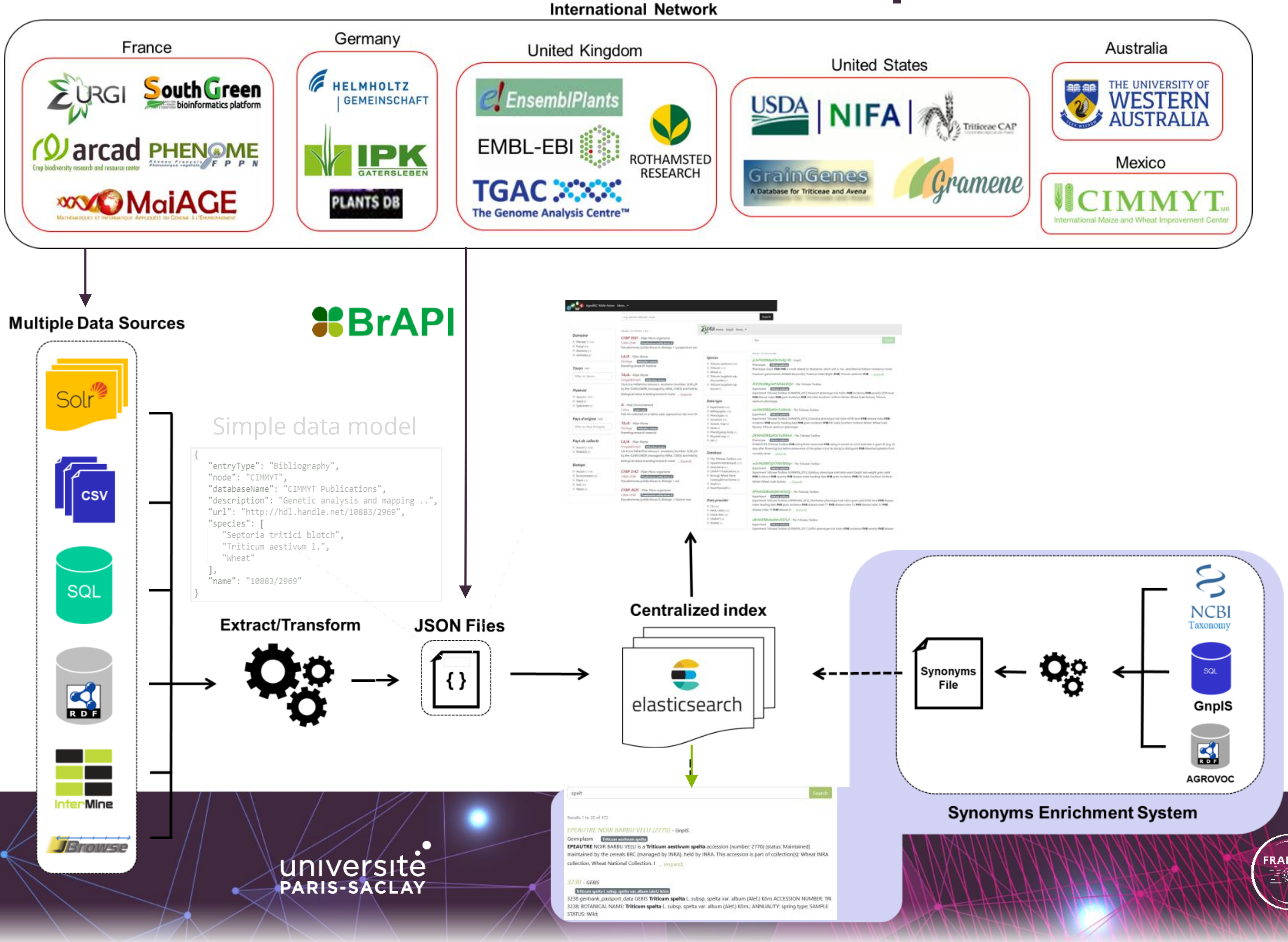
Portail de recherche des données scientifiques

wheatIS
WHEAT INFORMATION SYSTEM

28 bases de données indexées pointant vers 30 millions données

FAIDARE

40 bases de données indexées pointant vers 97 millions de données



```
{
   "entryType": "Bibliography",
   "node": "CIMMYT",
   "databaseName": "CIMMYT Publications",
   "description": "Genetic analysis and mapping ..",
   "url": "http://hdl.handle.net/10883/2969",
   "species": [
      "Septoria tritici blotch",
      "Triticum aestivum l.",
      "Wheat"
   ],
   "name": "10883/2969"
}
```

spelt
Search

Results 1 to 20 of 472

EPEAUTRE NOIR BARBU VELU (2776) - *GnpIS*

Germplasm Triticum aestivum spelta

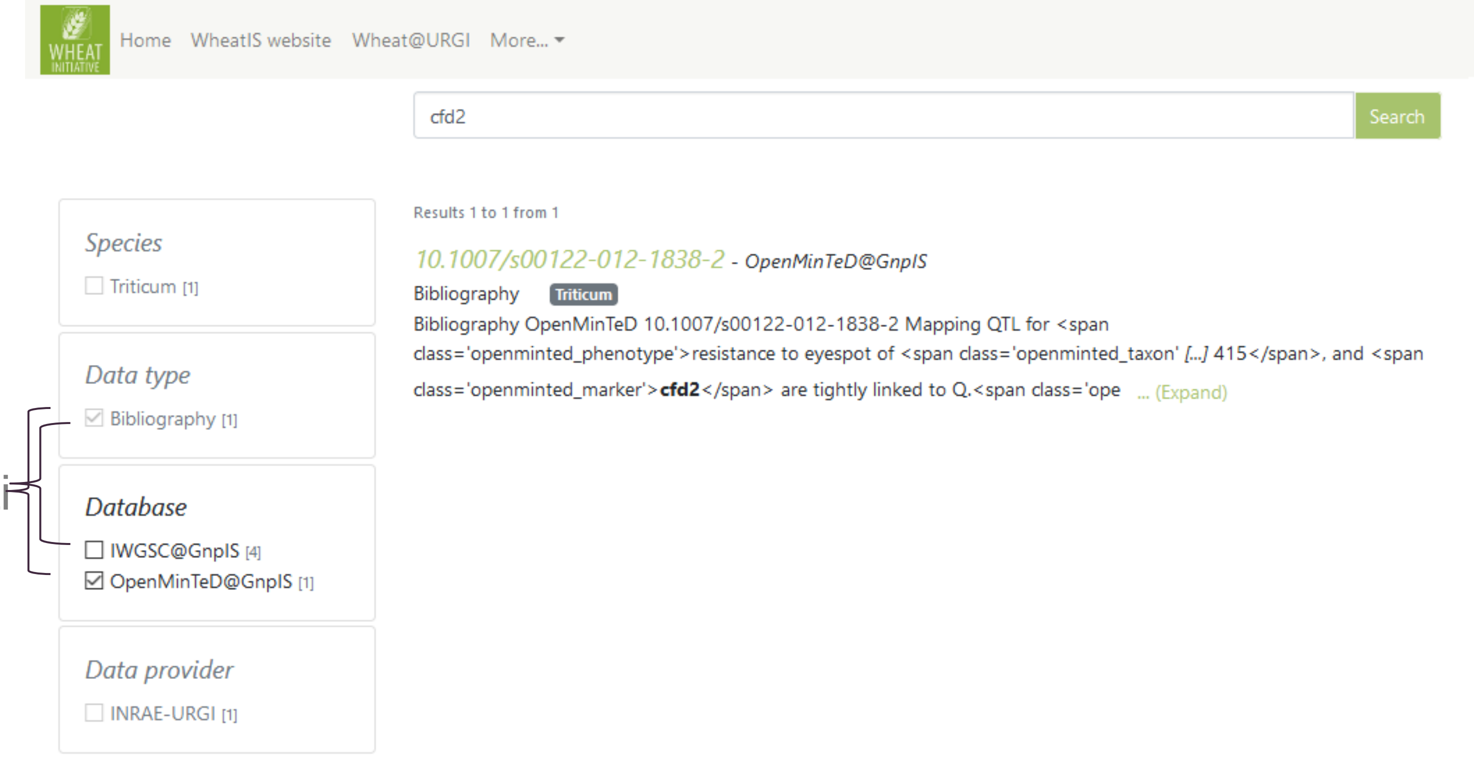
EPEAUTRE NOIR BARBU VELU is a **Triticum aestivum spelta** accession (number: 2776) (status: Maintained) maintained by the cereals BRC (managed by INRA), held by INRA. This accession is part of collection(s): Wheat INRA collection, Wheat National Collection. I ... (expand)

3238 - *GEBIS*

Triticum spelta L. subsp. spelta var. album (alef.) körn

3238 genbank_passport_data GEBIS **Triticum spelta** L. subsp. spelta var. album (Alef.) Körn ACCESSION NUMBER: TRI 3238; BOTANICAL NAME: **Triticum spelta** L. subsp. spelta var. album (Alef.) Körn.; ANNUALITY: spring type; SAMPLE STATUS: Wild;

Exemple de recherche



cfd2

Species

☐ Triticum [1]

Data type

☑ Bibliography [1]

Database

☐ IWGSC@GnpIS [4]

☑ OpenMinTeD@GnpIS [1]

Data provider

☐ INRAE-URGI [1]

Results 1 to 1 from 1

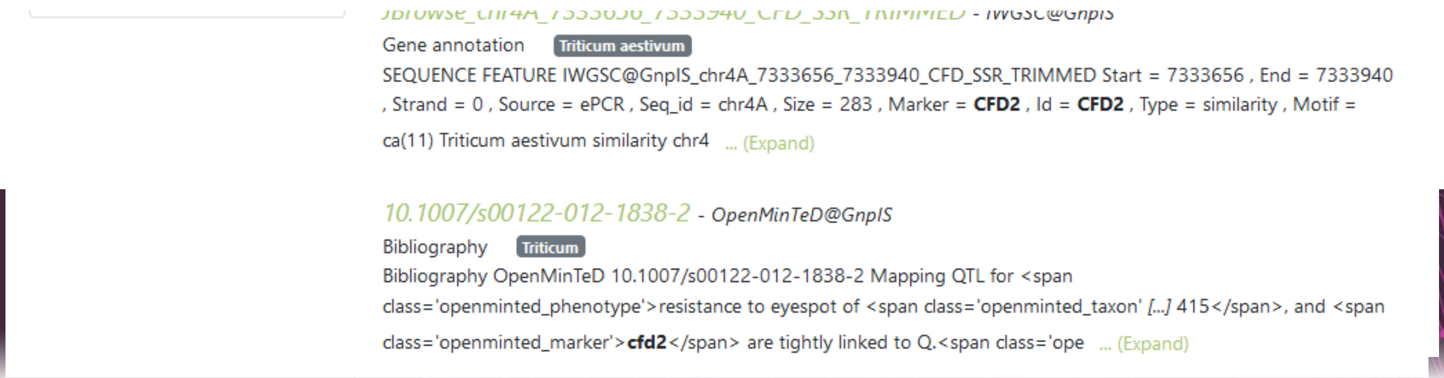
10.1007/s00122-012-1838-2 - *OpenMinTeD@GnpIS*

Bibliography **Triticum**

Bibliography OpenMinTeD 10.1007/s00122-012-1838-2 Mapping QTL for resistance to eyespot of <span class='openminted_taxon' *[...]* 415, and **cfd2** are tightly linked to Q.<span class='ope ... (Expand)

Union (OR query)

Intersection (AND query)



JBrowse_chr4A_7333656_7333940_CFD_SSR_TRIMMED - *IWGSC@GnpIS*

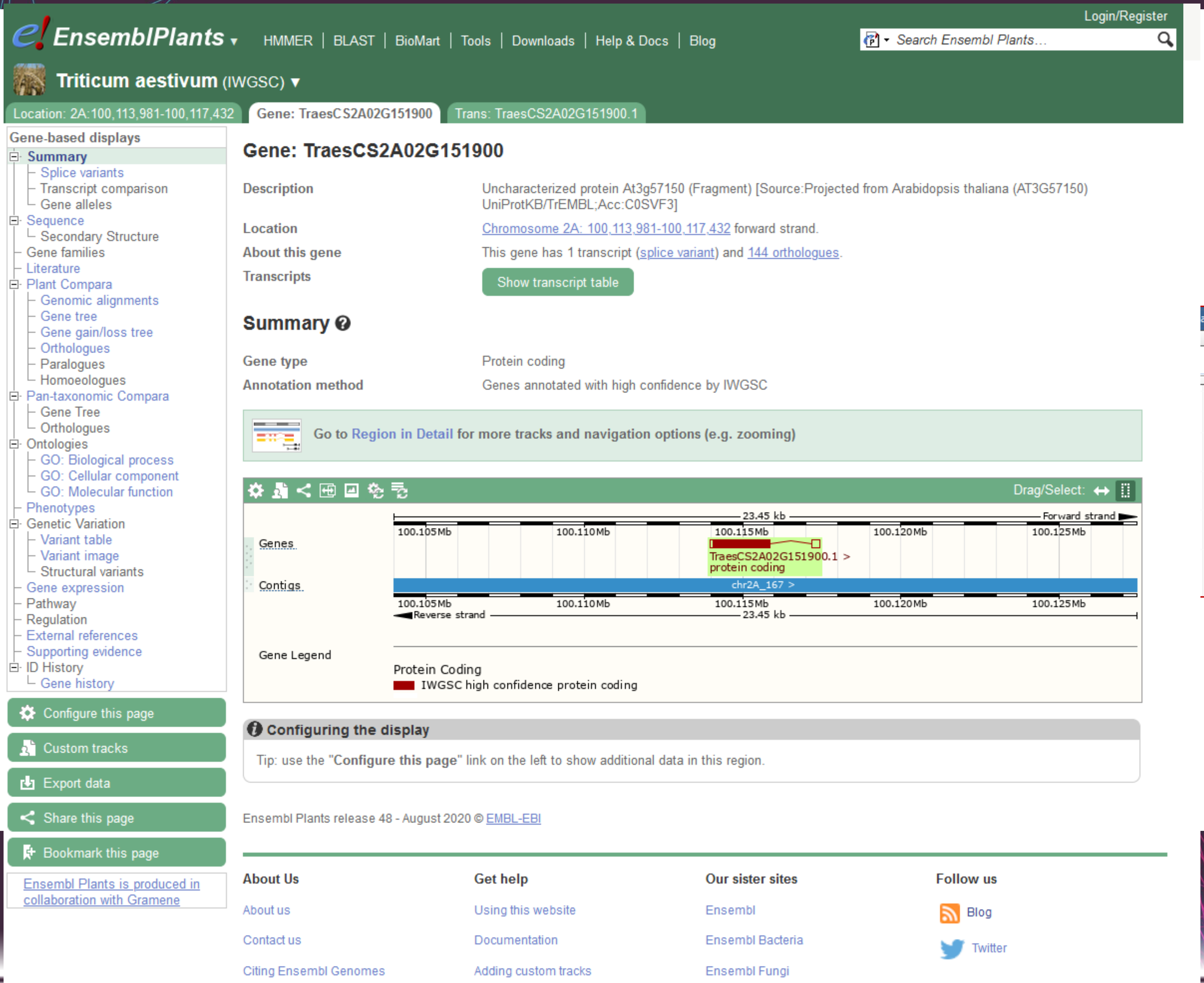
Gene annotation **Triticum aestivum**

SEQUENCE FEATURE IWGSC@GnpIS_chr4A_7333656_7333940_CFD_SSR_TRIMMED Start = 7333656 , End = 7333940 , Strand = 0 , Source = ePCR , Seq_id = chr4A , Size = 283 , Marker = **CFD2** , Id = **CFD2** , Type = similarity , Motif = ca(11) Triticum aestivum similarity chr4 ... (Expand)

10.1007/s00122-012-1838-2 - *OpenMinTeD@GnpIS*

Bibliography **Triticum**

Bibliography OpenMinTeD 10.1007/s00122-012-1838-2 Mapping QTL for resistance to eyespot of <span class='openminted_taxon' *[...]* 415, and **cfd2** are tightly linked to Q.<span class='ope ... (Expand)



Gene: TraesCS2A02G151900

Description	Uncharacterized protein At3g57150 (Fragment) [Source:Projected from Arabidopsis thaliana (AT3G57150) UniProtKB/TrEMBL;Acc:C0SVF3]
Location	Chromosome 2A: 100,113,981-100,117,432 forward strand.
About this gene	This gene has 1 transcript (splice variant) and 144 orthologues.
Transcripts	Show transcript table

Summary

Gene type	Protein coding
Annotation method	Genes annotated with high confidence by IWGSC

Go to Region in Detail for more tracks and navigation options (e.g. zooming)

Gene Legend

Protein Coding

🟥 IWGSC high confidence protein coding

Configuring the display

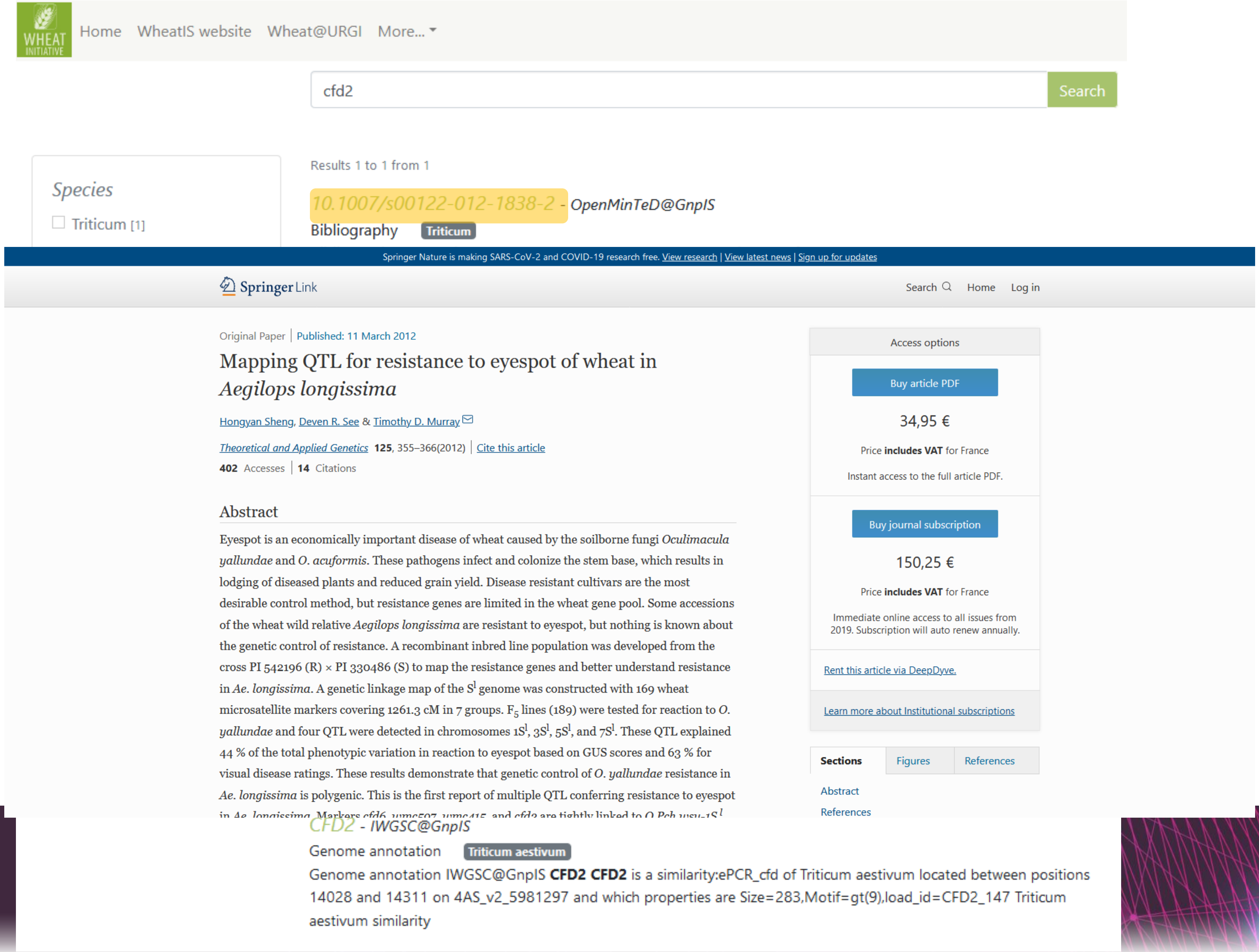
Tip: use the "**Configure this page**" link on the left to show additional data in this region.

Ensembl Plants release 48 - August 2020 © EMBL-EBI

About Us	Get help	Our sister sites	Follow us
About us	Using this website	Ensembl	Blog
Contact us	Documentation	Ensembl Bacteria	Twitter
Citing Ensembl Genomes	Adding custom tracks	Ensembl Fungi	

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR, DE LA RECHERCHE ET DE L'ESPACE

FRANCE 2030



Results 1 to 1 from 1

10.1007/s00122-012-1838-2 - OpenMinTeD@GnpIS

Bibliography **Triticum**

Springer Nature is making SARS-CoV-2 and COVID-19 research free. View research | View latest news | Sign up for updates

Original Paper | Published: 11 March 2012

Mapping QTL for resistance to eyespot of wheat in *Aegilops longissima*

Hongyan Sheng, Deven R. See & Timothy D. Murray

Theoretical and Applied Genetics **125**, 355–366(2012) | Cite this article

402 Accesses | **14** Citations

Abstract

Eyespot is an economically important disease of wheat caused by the soilborne fungi *Oculimacula yallundae* and *O. acuformis*. These pathogens infect and colonize the stem base, which results in lodging of diseased plants and reduced grain yield. Disease resistant cultivars are the most desirable control method, but resistance genes are limited in the wheat gene pool. Some accessions of the wheat wild relative *Aegilops longissima* are resistant to eyespot, but nothing is known about the genetic control of resistance. A recombinant inbred line population was developed from the cross PI 542196 (R) × PI 330486 (S) to map the resistance genes and better understand resistance in *Ae. longissima*. A genetic linkage map of the S^l genome was constructed with 169 wheat microsatellite markers covering 1261.3 cM in 7 groups. F_5 lines (189) were tested for reaction to *O. yallundae* and four QTL were detected in chromosomes $1S^l$, $3S^l$, $5S^l$, and $7S^l$. These QTL explained 44 % of the total phenotypic variation in reaction to eyespot based on GUS scores and 63 % for visual disease ratings. These results demonstrate that genetic control of *O. yallundae* resistance in *Ae. longissima* is polygenic. This is the first report of multiple QTL conferring resistance to eyespot in *Ae. longissima*. Markers *cfd6*, *wmc507*, *wmc415*, and *cfd2* are tightly linked to *O.Pch.wsu-1S*^{l}

34,95 €

Price **includes VAT** for France

Instant access to the full article PDF.

150,25 €

Price **includes VAT** for France

Immediate online access to all issues from 2019. Subscription will auto renew annually.

Rent this article via DeepDyve.

Learn more about Institutional subscriptions

CFD2 - IWGSC@GnpIS

Genome annotation **Triticum aestivum**

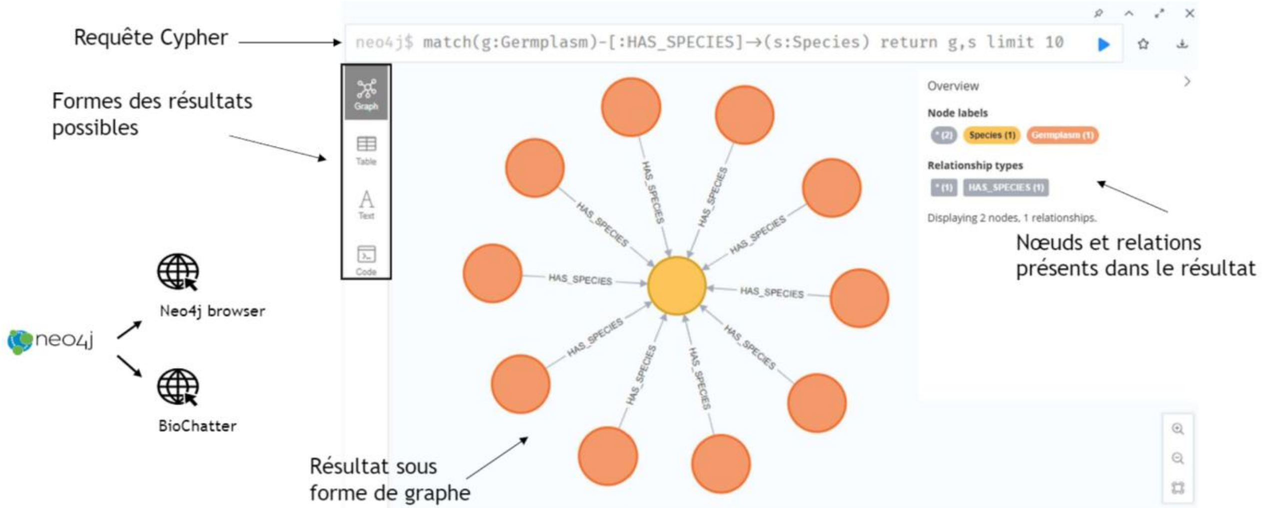
Genome annotation IWGSC@GnpIS **CFD2 CFD2** is a similarity:ePCR_cfd of Triticum aestivum located between positions 14028 and 14311 on 4AS_v2_5981297 and which properties are Size=283,Motif=gt(9),load_id=CFD2_147 Triticum aestivum similarity

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'ESPACE

FRANCE 2030

Notre R&D autour de la donnée

- Intégration de données dans des graphes de connaissances
- Utilisation de l’IA pour améliorer intégration, curation, interface homme-machine



```
neo4j$ match(g:Germplasm)-[:HAS_SPECIES]→(s:Species) return g,s limit 10
```

Overview

Node labels

Relationship types

Displaying 2 nodes, 1 relationships.

Requête Cypher

Formes des résultats possibles

Nœuds et relations présents dans le résultat

Résultat sous forme de graphe

Identification de germplasms avec leur espèce

Conclusions de notre retour d'expérience

- Les infrastructures de recherche, IFB et ELIXIR, ont permis de capitaliser sur le long terme les développement de ressources support
  - maintenance des ontologies, standards, recommendations
  - maintenance des entrepôts de données
  - personnels experts dans la standardisation et curation de données (*data stewards*)
- Les ressources support doivent être co-construites avec les communautés de recherche (ex: Wheat Initiative, RDA)
  - cadrage des besoins et expertise sur les données
  - personnes relai pour diffuser les bonnes pratiques

Identification des partenaires et des ressources:
- ➢ Ecosystème Recherche Data Gouv
- ➢ forum des ESFRI

- ➢ Impliquer l'infrastructure de recherche pertinente comme partenaire de projet
- ➢ Faire travailler les infrastructures ensemble

Conclusions de notre retour d’expérience

- Les activités de formation et de montée en capacité sont critiques
  - former les chercheurs
  - former des formateurs
  - développement de matériel de formation réutilisable
- Dans un contexte de grande dispersion des données, le développement d’une fédération de données facilite la mise en place de réseaux de collaborations entre bases de données

Identification de partenaires et de ressources:

- Ecosystème Recherche Data Gouv
- forum des ESFRI

Les systèmes de labellisation des bases de données partenaires devraient permettre d’identifier les forces et faiblesses pour consolider ces fédérations

Remerciements

URGI

Raphaël Flores : coordination opérationnelle des services

Cyril Pommier : co-pilote services données et fédération, responsable communauté plante d’ELIXIR, responsable données Phenome-Emphasis

Célia Michotey : co-pilote services données et fédération

Vanita Haurheeram : data steward BReIF et Pro-Wild

Erwan Le-Floch et Etienne Bardet : data steward AGENT

URGI team

IFB
INSTITUT FRANÇAIS DE BIOINFORMATIQUE

Morgane Thomas-Chollier : directrice

Jacques Van Helden : coordination scientifique

IFB-core team

Questions

NOUS CONTACTER

Michaël Alaux: michael.alaux@inrae.fr

Anne-Françoise Adam-Blondon: anne-francoise.adam-blondon@inrae.fr

https://urgi.versailles.inrae.fr/

https://www.ifb-elixir.fr/