



**HAL**  
open science

## Les outils de gestion de données de phénotypage soutenus par INRAE

Cyril Pommier, Anne Delaunay

► **To cite this version:**

Cyril Pommier, Anne Delaunay. Les outils de gestion de données de phénotypage soutenus par INRAE. AG RARE, INRAE, May 2023, Toulouse, France. hal-04314299

**HAL Id: hal-04314299**

**<https://hal.science/hal-04314299>**

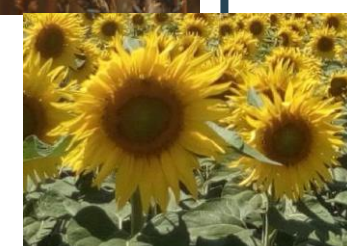
Submitted on 29 Nov 2023

**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 - NonCommercial 4.0 International License



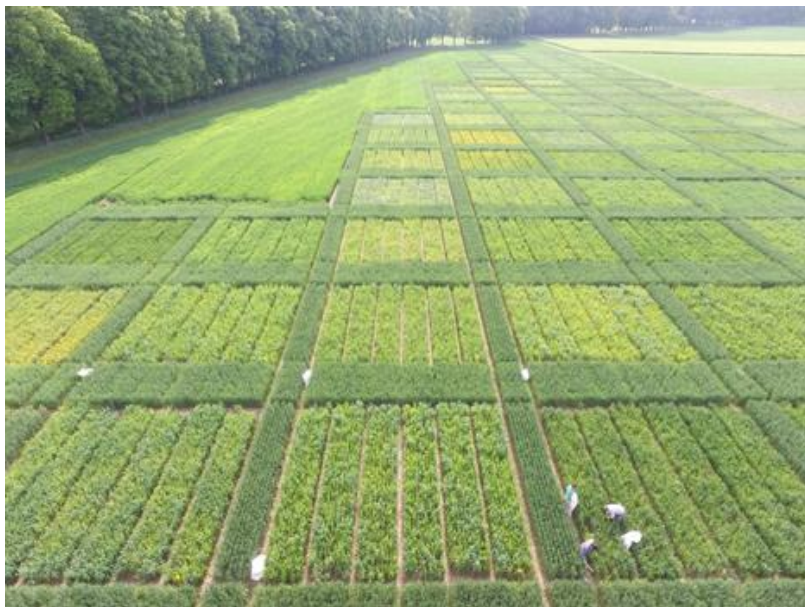
# Les outils de gestion de données de phénotypage soutenus par INRAE

# Plan de la session

- **Introduction Anne Delaunay et Cyril Pommier (15 minutes)**
- **Un outil dédié à l'acquisition des données à INRAE : Adonis WebApp et Mobile (Fabien SURAULT, Audrey PALAIS-SAINT-AGATHE INRAE) (20 minutes + 10 minutes de questions)**
- **PHIS et Sixtine pour la gestion de données de Phénotypage (Anne Tireau, Isabelle Alic INRAE) (20 minutes + 10 minutes de questions)**
- **Thalia DB/OLGA (Franciane Nuissier, Jacques Lagnel INRAE) (20 minutes + 10 minutes de questions)**

# Les UE AgroEcoSystem BAP SPE d'INRAE

- › 41 UE INRAE → 21 Unités Expérimentales dans la coordination des UE des 3 départements dont 12 labellisées ISC
- › + de 1700 ha de surface agricole en grandes cultures, arboriculture, vigne, cultures maraîchères
- › Conditions agropédoclimatiques contrastées



- › Conduites culturales différenciées
- › Un environnement caractérisé / en cours de caractérisation (climat, sols)
- › Une caractérisation du matériel végétal du bas débit (notations "à la main") au haut-débit, avec vecteur instrumenté (drône, phéno-mobile, ...)
- › Des expérimentations en petites parcelles, en grandes parcelles,  
↳ Des données nombreuses et diversifiées sont produites sur les UE  
annuelles ou pluriannuelles, des approches analytiques et systémiques.

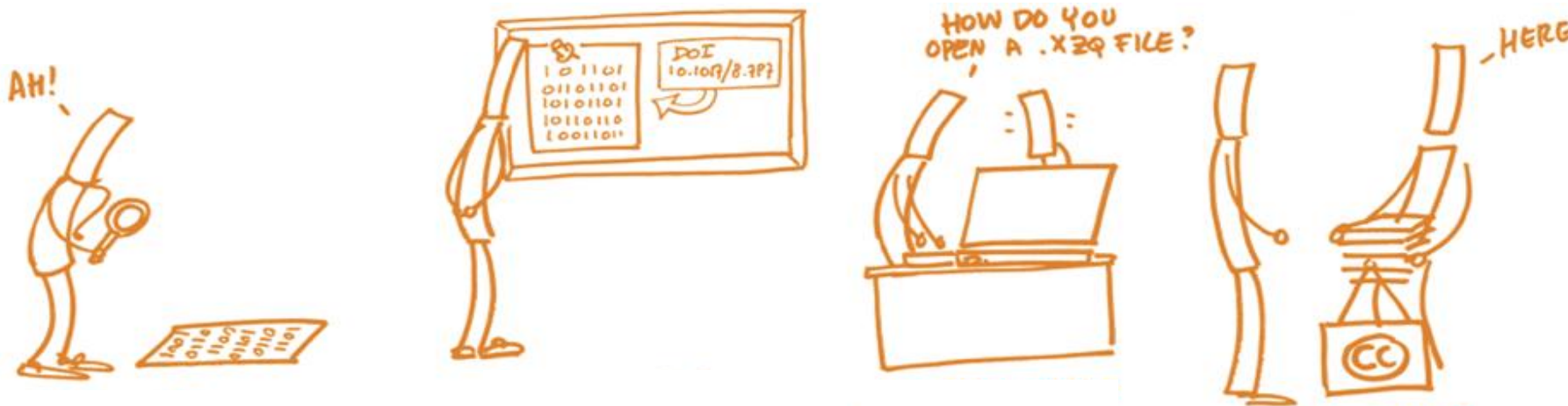
- ⇒ **Des données nombreuses et diversifiées sont produites sur les UE**
  
- ⇒ **Accompagnement des UE par les départements depuis ~2018 pour les aider à organiser la gestion de leurs données**
  
- ⇒ **Intérêt de partager cette expérience avec les CRB plantes :**
  - ❑ **Proximité de métier : par exemple entretien, caractérisation de matériel végétal**
  - ❑ **Certains CRB sont intégrés dans des UE, pour une partie ou la totalité de leurs activités**

- **FAIR data**





# FAIR Data Principles



**FINDABLE**

Unique identifiers and metadata are used to allow data to be located quickly and efficiently

**Ids  
Metadata  
Index**

**ACCESSIBLE**

Data is open, free and universally available for research discovery efforts

**Open Protocole  
Perennial  
Metadata**

**INTER-OPERABLE**

A common programming language is used to allow use in a broad range of applications

**Technical &  
Data  
(Semantics  
Linked Data  
Vocabularies)**

**REUSABLE**

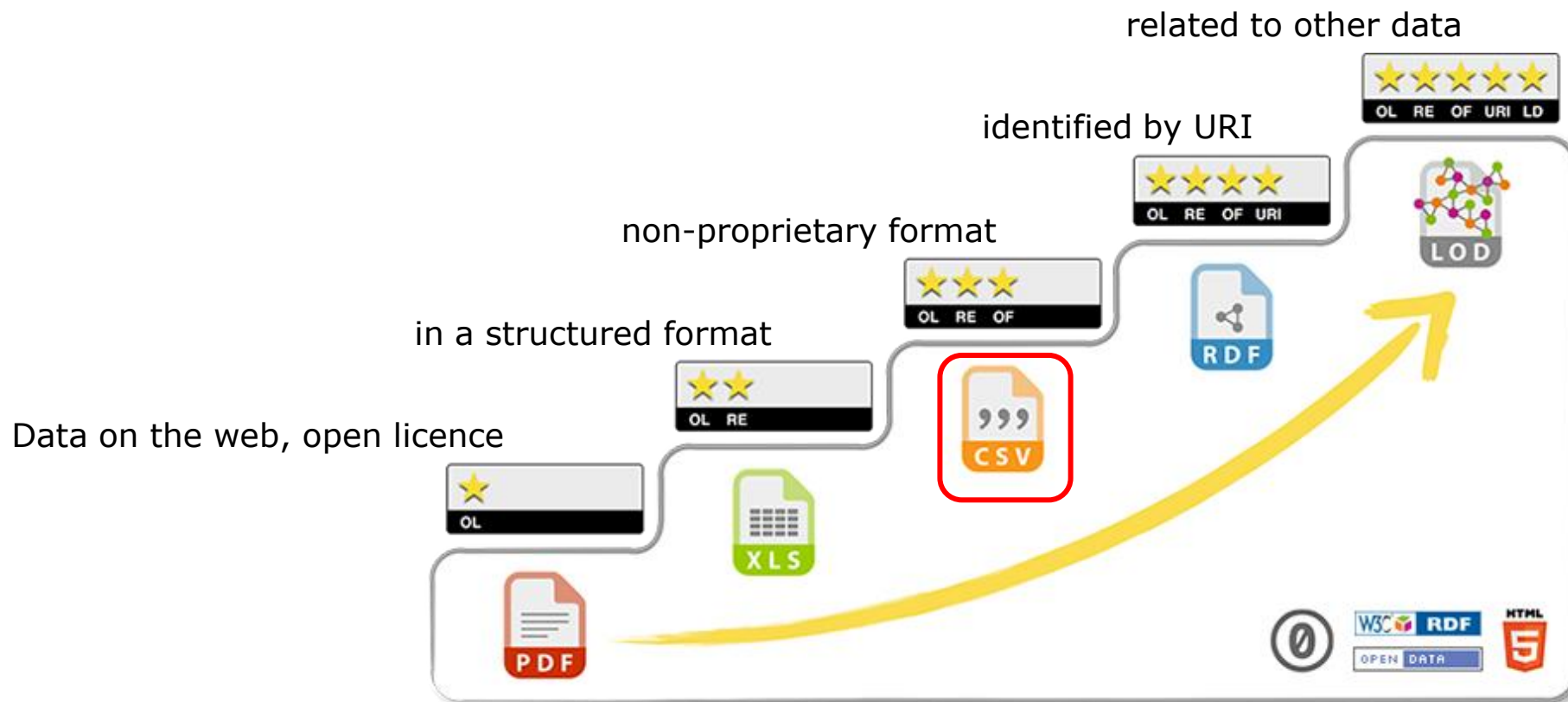
All data is clearly described and outlines associated data-use standards

**License  
Well described  
Provenance  
Standards**

Wilkinson et al., **The FAIR Guiding Principles for scientific data management and stewardship.** *Scientific Data* 3 (2016)



# Objectif FAIR: 5 stars Open Data

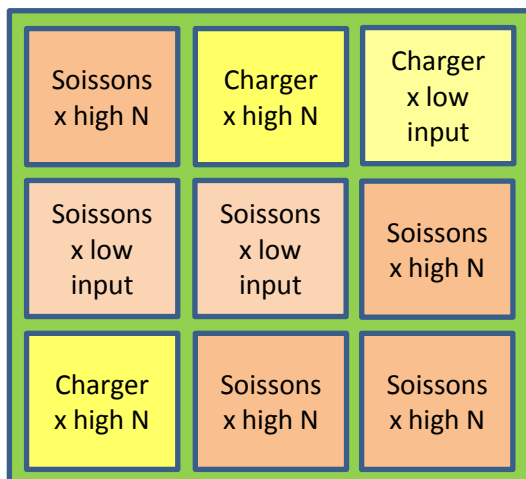


Mise en place de la science ouverte et du FAIR implique une coopération multidisciplinaire:

- Biologistes
- Bioinformaticiens
- Spécialistes de la gestion de données (ontologies/semantics, standards, ...)

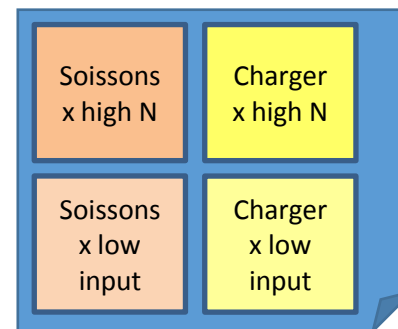


AG BRC4Plants, 31 mai au 1<sup>er</sup> juin 2023

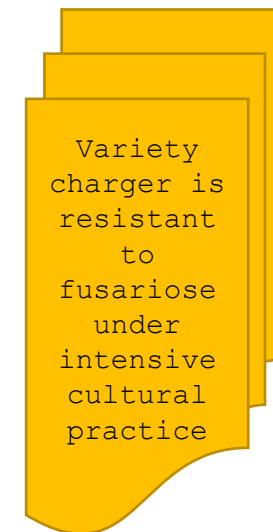


Derivation, Reduction

données dérivées, calculées, réduites, indicateurs



Publication



Genotype	Treatment	N input	Date	Rep	Fusariose
<b>Soissons</b>	<b>low input</b>	15,3225129	15/11/2011	1	5
<b>Soissons</b>	<b>low input</b>	15,3143556	16/11/2011	2	7

## VARIABLES

Tracabilité  
Mesures brutes  
Pheno & Env  
Nettoyage de Données  
Reproductibilité & Provenance

Genotype	Treatment	Fusariose
<b>Soissons</b>	<b>low input</b>	6

## INDICATEURS

Nouveau calcul pour chaque question scientifique  
Un dataset experimental → plusieurs dataset dérivés

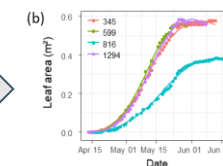
## DESCRIPTEURS

**PRIMAIRES**  
Characterisation  
Phenodes  
accessions

**Raw data long term conservation**

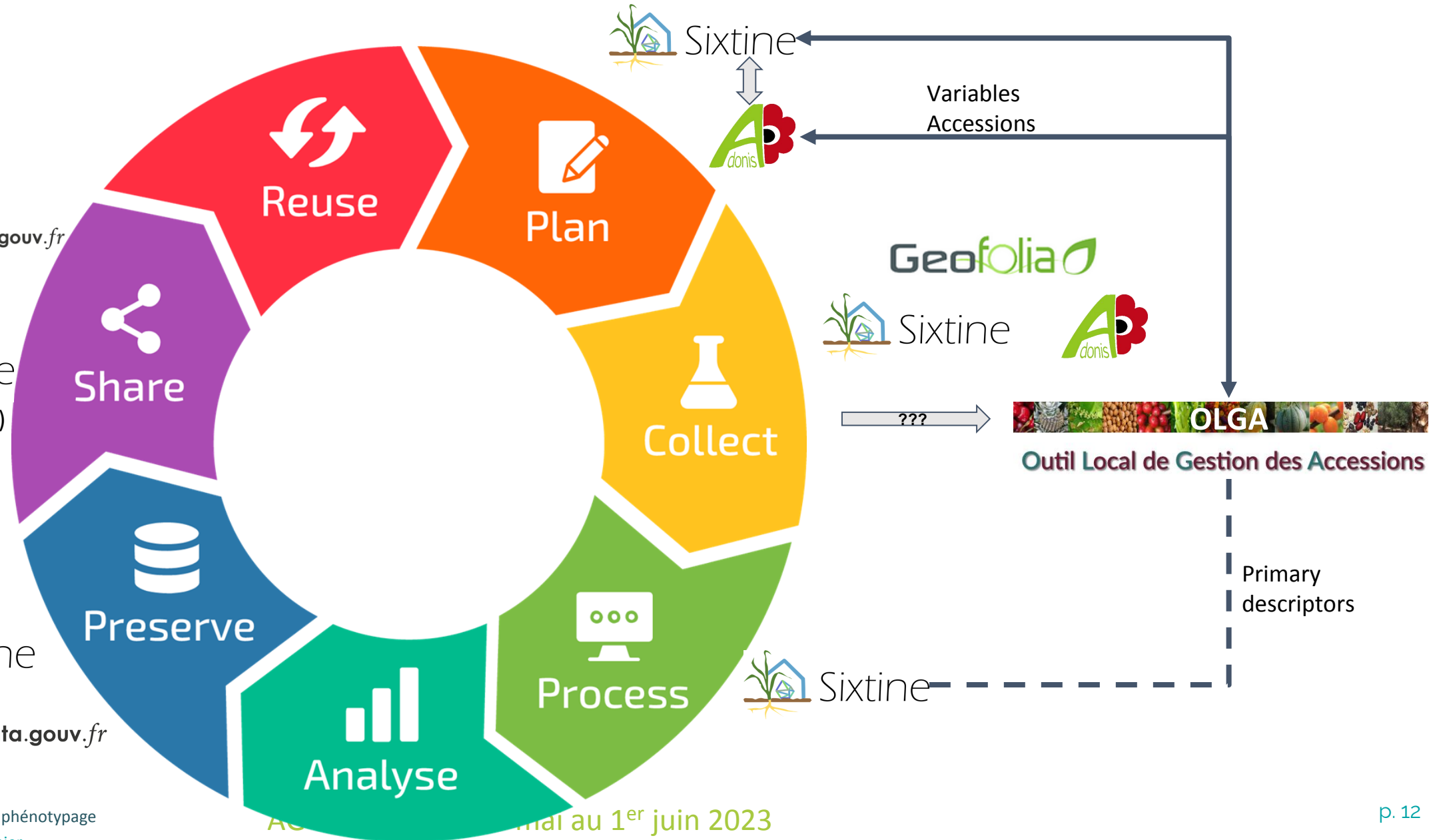


year	Site	Env	Date	D20.air	D20cum.air
2012	Gaillac	Gai12	2012-05-21	0.38382247	0.38382247
2012	Gaillac	Gai12	2012-05-22	0.5249961365	0.9088186066
2012	Gaillac	Gai12	2012-05-23	0.5882208771	1.4970394837
2012	Gaillac	Gai12	2012-05-24	0.9890314415	2.4860709251
2012	Gaillac	Gai12	2012-05-25	1.3484068027	3.8344777278
2012	Gaillac	Gai12	2012-05-26	0.9767565078	4.8112342356
2012	Gaillac	Gai12	2012-05-27	1.0450063392	5.8562405749



Accession_ID	genetic.group	genotypic.main	beta.Tnight.Flo	beta.Ri
11430_H	lodent	0	2.2	-0.58
A3_H	Other	-1454.2	48.2	1.63
A310_H	Other	-1238.7	105.5	-5.01
A347_H	Other	-3763.3	181.6	-2.88
A374_H	Other	1030.9	-10.6	-1.09
A375_H	Other	-1120.9	71.9	-1.37
A554_H	Other	-1732.2	76.7	-1.64
A5707_H	Other	2794.9	100.9	2.05

# Cycle de vie de la donnée de Phénotypage





- Planification
- Saisie au champs et en serre



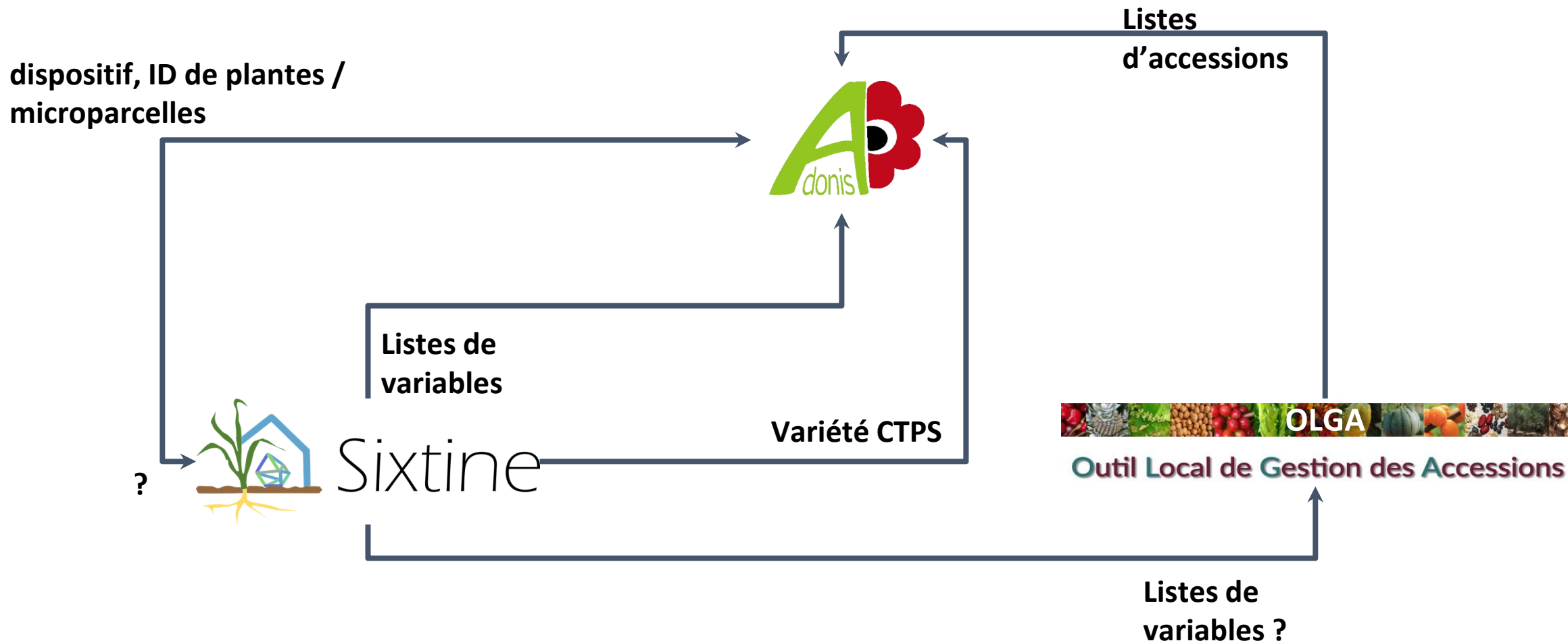
## Sixtine

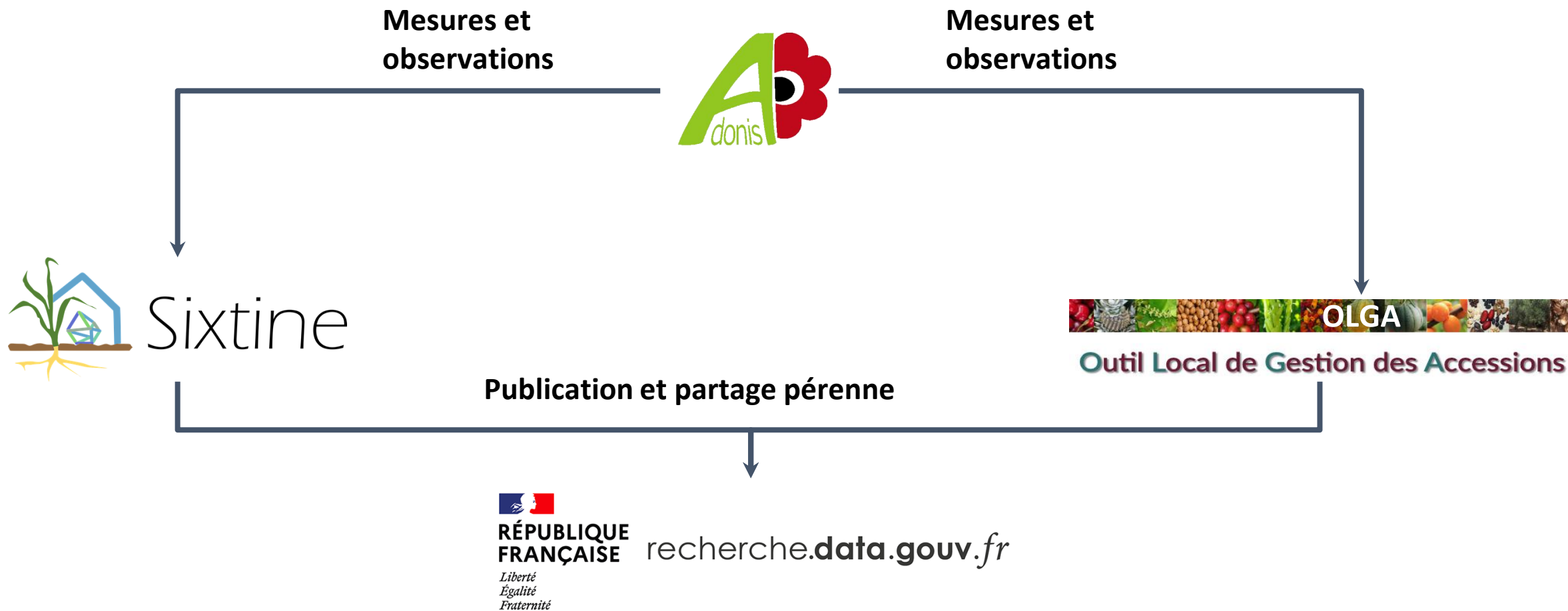
- Planification/design expérimental
- Documentation des parcelles
- Organisation des UE
- Partage/création de listes de variables
- Collecte quotidienne des données
- Consolidation, nettoyage avec des outils (R)
- Dérivation, Analyse, Analyse d'Image



## Outil Local de Gestion des Accessions

- Gestion des CRB
- liste d'identifiant d'accessions
- Phenotypes primaires
- Autres phénotypes ?
- Gestion d'expériences de caractérisation simple ?








## Entrepot

- **Description minimale simple**
  - Abstract décrivant l'expérience
  - Auteurs et personnes impliquées
- **Fichiers de données**
- **Lien vers les données complexes/volumineuses**
  - PHIS / Sixtine
- **Connexion automatisée en cours**

### Phenological data of 26 varieties of cherry, apricot, peach and apple trees since 2016 in a French orchard of Toulence, Gironde, France

Version 3.0





Alletru, David; Gazel, François; Corradi, Michel; Grell, Marie-Laure; Delmas, Marine, 2020, "Phenological data of 26 varieties of cherry, apricot, peach and apple trees since 2016 in a French orchard of Toulence, Gironde, France", <https://doi.org/10.15454/TJTV6N>, Recherche Data Gouv, V3, UNF:6:HPzk9LBM04Yk5dx5fRsT6Q== [fileUNF]

**Description**

This dataset contain phenological observations and growth data of 7 cherry, apricot and apple trees varieties and 5 peach trees varieties evaluated since 2016.

These data come from observations performed since 2016 on "Les Jarres" estate belonging to the Arbicultural

1 à 9 de 9 Fichiers

	<p><b>Apple phenology.tab</b></p> <p>Données tabulaires - 21.1</p> <p>Publié 26 juil. 2022</p> <p>7 téléchargements</p> <p>1 Variables, 147 Observations</p> <p>Phenological data of 7 apple trees</p>
	<p><b>Apricot phenology.tab</b></p> <p>Données tabulaires - 20.7</p> <p>Publié 26 juil. 2022</p> <p>5 téléchargements</p>

31 mai au 1<sup>er</sup> juin 2023

## Entrepôt

- Description minimale simple
  - Abstract décrivant l'expérience
  - Auteurs et personnes impliquées
- Fichiers de données
- Lien vers les données complexes/volumineuses
  - PHIS / Sixtine
- Connexion automatisée en cours

## Standard de données



- Minimal Information About Plant Phenotyping Experiment
  - Inclus MCPD, BrAPI
  - Implémenté dans Sixtine et Olga
  - Trois fichiers de descriptions
    - Biological Material
    - Study (experiment)
    - Observed variables

### Phenological data of 26 varieties of cherry, apricot, peach and apple trees since 2016 in a French orchard of Toulence, Gironde, France

Version 3.0



Alletru, David; Gazel, François; Corradi, Michel; Grell, Marie-Laure; Delmas, Marine, 2020, "Phenological data of 26 varieties of cherry, apricot, peach and apple trees since 2016 in a French orchard of Toulence, Gironde, France", <https://doi.org/10.15454/TJTV6> N, Recherche Data Gov, V3, UNF:6:HPzk9LBM04Yk5dx5fRsT6Q== [fileUNF]

#### Description

This dataset contains phenological observations and growth data of 7 cherry, apricot and apple tree varieties and 5 peach tree varieties evaluated since 2016.


These data come from observations performed since 2016 on "Les Jarres" estate belonging to the Arbicultural

1 à 9 de 9 Fichiers

	<p><b>Apple phenology.tab</b> Données tabulaires - 21.1 Ko Publié 26 juil. 2022 7 téléchargements 1 Variables, 147 Observations Phenological data of 7 apple varieties</p>
	<p><b>Apricot phenology.tab</b> Données tabulaires - 20.7 Ko Publié 26 juil. 2022 5 téléchargements</p>

 **8-Info-BiologicalMaterial.tab**  
Données tabulaires - 49.9 Ko  
Publié 3 nov. 2022  
104 téléchargements  
Accession identifier, "G...  
Taxonomy identifier, "G...  
**Biological Material**

 **10-Info-ObservedVariable.tab**  
Données tabulaires - 21.8 Ko  
Publié 3 nov. 2022  
132 téléchargements  
17 Variables, 91 Observations  
List of phenotypic and environmental variables for each observation in the dataset, following the MIAPPE data standard (https://www.miappe.org/)  
**Observed Variable**

 **11-Info-Study.tab**  
Données tabulaires - 6.4 Ko  
Publié 3 nov. 2022  
110 téléchargements  
11 Variables, 19 Observations  
List of studies, including location and date, following the MIAPPE data standard (https://www.miappe.org/)  
**Study**

31 mai au 1<sup>er</sup> juin 2023



# MERCI