



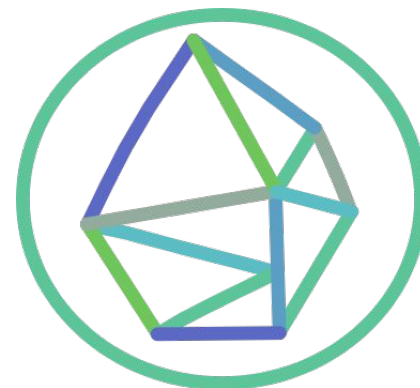
Le savoir FAIR d'OpenSILEX

Sarra Abidri, Isabelle Alic, Arnaud Charleroy, Anne Tireau

Séminaire interne MISTEA - 15 décembre 2025

Le savoir FAIR d'OpenSILEX

- OpenSILEX - c'est qui c'est quoi ?
- Pourquoi et comment faire FAIR
- OpenSILEX les grands principes
- OpenSILEX et MISTEA
 - Ontologies, graphes de connaissances et variables
 - Valorisation des données : modélisation et statistiques -> *SilexExplorer* !



OpenSILEX c'est quoi ?



Cellule de production logicielle... mais pas que !

Notre objectif : Permettre aux équipes de recherche de répondre aux enjeux climatiques, agroalimentaires et environnementaux en structurant les données !

Médiateur
scientifique

*Interpréter les besoins scientifiques, interface
entre les scientifiques et l'informatique*

Production Logicielle



Ambassadeur
scientifique

*Participer à la valorisation scientifique :
données exploitées, partagées et reconnues*

Structurant et
fédérateur

*Fournir la sémantique, les standards
et les vocabulaires contrôlés*

Orientations scientifiques

Expert garant et
validateur

*Participer à la modélisation
et à la qualité des données*

Facilitateur
analytique

- Assurer la compatibilité et la connexion avec les pipelines de traitement des données
- Préparer les données pour l'analyse et la modélisation

Nouveaux sujets de fort intérêt scientifique :

- ❖ Nouvelles thématiques mobilisant les concepts encore en structuration Agroécologie et Agroforesterie
- ❖ Intelligence artificielle et construction / communication avec des ontologies



OpenSILEX c'est qui ?

Présentation de l'équipe et organisation

Exploitation

- Arnaud CHARLEROY **Resp. Techno**
- Brigitte CHEVAL
- Nouha JEMAA
- **Farzaneh KAZEMIPOUR** **Animatrice**
- Pascal NEVEU (retraité chargé de mission)
- Sandrine VRIGNON

Resp. : Anne TIREAU
Resp. Adj. : Isabelle ALIC

Production et intégration

- Sarra ABIDRI
- Lydia ALIANE
- Renaud COLIN **Infra. Phenome**
- Sandrine COLLAVIZZA
- Maximilian HART
- Sébastien PRADO
- Yvan ROUX

Référent-es Communautés

Infra. Phenome/Emphasis
Isabelle ALIC

Viticulture et Oenologie
Arnaud CHARLEROY

UE INRAE végétal
Anne TIREAU

Bio-procédé et environnement
Virginie ROSSARD (LBE)

Ferments du Futur
Emilie FERNANDEZ (Migale)

Bio-agresseurs
Lydia ALIANE

INRAE

Pourquoi et comment faire FAIR ?

2017 - Mauguio - Romain Chapuis

Expérimentation sur la résistance à la sécheresse de 60 variétés de Maïs

- Collecte de données environnementales
- Collecte de traits, données phénotypiques



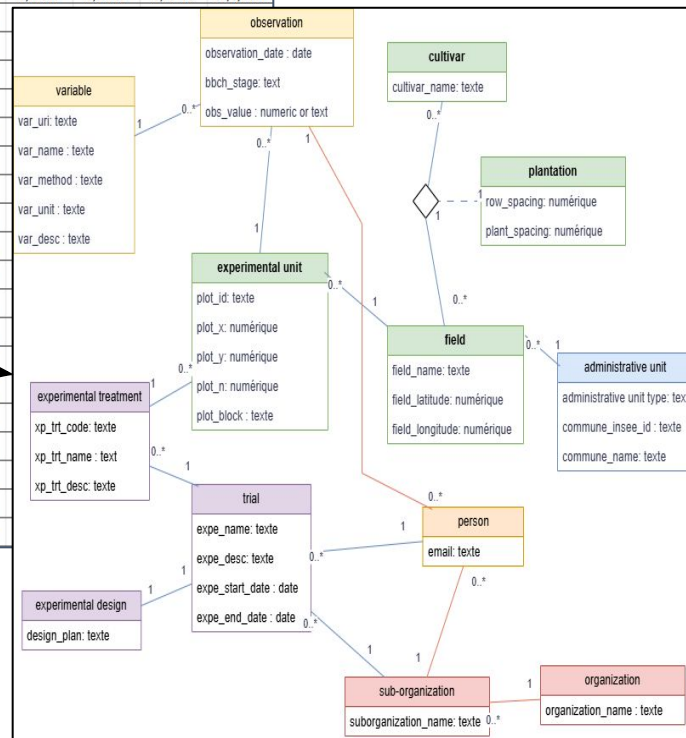
CodeB	Code_Trial	X	Y	Plot_ID	Treatment	Replicate	Block	Genotype_ID	Plot_Alias_Old	Length	Inter	Nb_Row	Nb_Seed	Nb_Ear_Spl	Sow	Emer	Visible_Leave	Date	Vig	Date	Lodging_Angle	Data
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_7_WW_2_19_5	MAU17-PG	19	5	2	WW																	
MAU17-PG_60_WW_3_19_6	MAU17-PG	19	6	3	WW																	
MAU17-PG_56_WW_4_19_7	MAU17-PG	19	7	4	WW																	
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	8	5	WW																	
MAU17-PG_FP_WW_6_19_9	MAU17-PG	19	9	6	WW																	
MAU17-PG_FP_WW_7_19_10	MAU17-PG	19	10	7	WW																	
MAU17-PG_46_WW_8_19_11	MAU17-PG	19	11	8	WW																	
MAU17-PG_41_WW_9_19_12	MAU17-PG	19	12	9	WW																	
MAU17-PG_11_WW_10_19_13	MAU17-PG	19	13	10	WW																	
MAU17-PG_44_WW_11_19_14	MAU17-PG	19	14	11	WW																	
MAU17-PG_8_WW_12_19_15	MAU17-PG	19	15	12	WW																	
MAU17-PG_58_WW_13_19_16	MAU17-PG	19	16	13	WW																	
MAU17-PG_52_WW_66_19_21	MAU17-PG	19	21	66	WW																	
MAU17-PG_26_WW_67_19_22	MAU17-PG	19	22	67	WW																	
MAU17-PG_30_WW_68_19_23	MAU17-PG	19	23	68	WW																	
MAU17-PG_32_WW_69_19_24	MAU17-PG	19	24	69	WW																	
MAU17-PG_5_WW_70_19_25	MAU17-PG	19	25	70	WW																	
MAU17-PG_14_WW_71_19_26	MAU17-PG	19	26	71	WW																	
MAU17-PG_16_WW_72_19_27	MAU17-PG	19	27	72	WW																	
MAU17-PG_47_WW_73_19_28	MAU17-PG	19	28	73	WW																	
MAU17-PG_51_WW_74_19_29	MAU17-PG	19	29	74	WW																	
MAU17-PG_36_WW_75_19_30	MAU17-PG	19	30	75	WW																	
MAU17-PG_13_WW_76_19_31	MAU17-PG	19	31	76	WW																	
MAU17-PG_20_WW_77_19_32	MAU17-PG	19	32	77	WW																	
MAU17-PG_6_WW_78_19_33	MAU17-PG	19	33	78	WW																	
MAU17-PG_14_WW_26_18_4	MAU17-PG	18	4	14	WW																	
MAU17-PG_10_WW_25_18_5	MAU17-PG	18	5	15	WW																	
MAU17-PG_13_WW_24_18_6	MAU17-PG	18	6	13	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW																	
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW			</														

Pourquoi et comment faire FAIR ?

CodeB	Code_Trial	X	Y	Plot_id	Treatment	Replicate	Block	Hybrid_id	Plot	Row_Sample	Row	Time.ms	Count	NDVI	VI_2	REDrefc	NIRrefc	Date
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW	1	1	DZ_PG_67	1	1	2	166310	47	0,16077	0,72338	0,58733	0,81248	31/5/2017
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW	1	1	DZ_PG_67	1	1	3	268410	44	0,16259	0,72073	0,58908	0,81765	31/5/2017
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW	1	1	DZ_PG_67	1	1	2	172720	42	0,8981	0,05388	0,01923	0,34789	28/6/2017
MAU17-PG_38_WW_1_19_4	MAU17-PG	19	4	1	WW	1	1	DZ_PG_67	1	1	3	181420	44	0,88859	0,05927	0,02534	0,42049	28/6/2017
MAU17-PG_7_WW_2_19_5	MAU17-PG	19	5	2	WW	1	1	DZ_PG_30	2	1	2	278110	40	0,16345	0,71943	0,57192	0,79495	31/5/2017
MAU17-PG_7_WW_2_19_5	MAU17-PG	19	5	2	WW	1	1	DZ_PG_30	2	1	3	285610	40	0,15948				
MAU17-PG_7_WW_2_19_5	MAU17-PG	19	5	2	WW	1	1	DZ_PG_30	2	1	2	203420	42	0,87667				
MAU17-PG_7_WW_2_19_5	MAU17-PG	19	5	2	WW	1	1	DZ_PG_30	2	1	3	211520	43	0,86788				
MAU17-PG_60_WW_3_19_6	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	294210	39	0,15974				
MAU17-PG_60_WW_3_19_6	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	301810	41	0,16129				
MAU17-PG_60_WW_3_19_6	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	227120	41	0,89088				
MAU17-PG_60_WW_3_19_6	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	234820	43	0,871				
MAU17-PG_56_WW_4_19_7	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	311710	40	0,15513				
MAU17-PG_56_WW_4_19_7	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	320310	41	0,15471				
MAU17-PG_56_WW_4_19_7	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	259720	37	0,89373				
MAU17-PG_56_WW_4_19_7	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	268220	40	0,88648				
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	329410	40	0,1655				
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	342410	39	0,15733				
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	280630	46	0,8948				
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	288830	41	0,88005				
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	2	356810	43	0,15488				
MAU17-PG_51_WW_5_19_8	MAU17-PG	19	6	3	WW	1	1	DZ_PG_49	3	1	3	365510	39	0,14395				
MAU17-PG_FP_WW_6_19_9	MAU17-PG	19	9	6	WW	1	1	DKC4590	6	1	2	301850	45	0,90404				
MAU17-PG_FP_WW_6_19_9	MAU17-PG	19	9	6	WW	1	1	DKC4590	6	1	3	308560	34	0,86535				

Beaucoup de fichiers .csv de données
=> manque de documentation, impossible à aligner avec un fichier d'une autre plateforme !

Pour documenter : faire un schéma de données
=> Restructurer le fichier .csv, plusieurs tableaux avec contrainte intégrité
=> Ajouter l'identification des entités

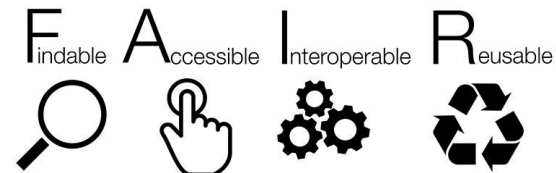


INRAE

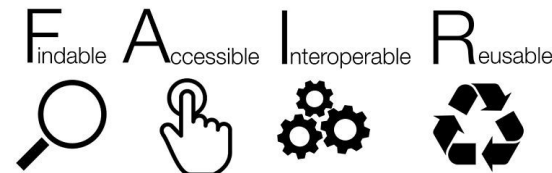
OpenSILEX
Séminaire interne MISTEA - 15 décembre 2025

Pourquoi et comment faire FAIR ?

2025 - Équipe de recherche qui veut entraîner son modèle -> besoin de jeux de données Mais



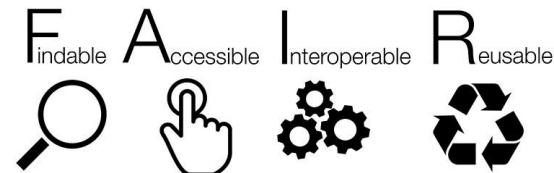
Pourquoi et comment faire FAIR ?



2025 - Équipe de recherche qui veut entraîner son modèle -> besoin de jeux de données Mais

- Comment **savoir que ce jeu de données existe** ?
- Comment **accéder** à ce jeu de données ? Qui peut y accéder ?
- Comment **interpréter** les données ? Comment les **lier** au jeu de données récupérées au jeu de données de l'Unité Expérimentale de Clermont-Ferrand ?
- Comment **ré-utiliser** ces données dans un autre contexte, interdisciplinaire ? Comment ont été générées ces données ?

Pourquoi et comment faire FAIR ?



2025 - Équipe de recherche qui veut entraîner son modèle -> besoin de jeux de données Mais

- Comment **savoir** que ce jeu de données existe ?
 - **F : Findable / Facilement trouvable**
Faciliter la découverte des données (et de leurs métadonnées)
tant pour les humains que pour les machines
 - Les données ont un **PID** (Persistent Identifier, identifiant unique et pérenne)
 - Les données sont **décrites par des métadonnées**
 - Ces **métadonnées incluent le PID** des données qu'elles décrivent
 - Les données sont déposées dans un **entrepôt de données**



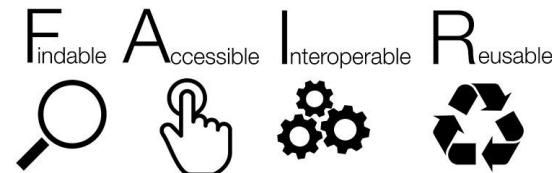
Métadonnée : donnée servant à définir ou décrire une autre donnée,
quel qu'en soit le support (papier, électronique ou autre).

- DOI : <https://doi.org/10.74335/MyDOI>
- Espèce : Maïs
- Localisation : Mauguio (France)
- Année : 2017
- Modalités expérimentation : irrigation
- Contact scientifique : Romain Chapuis



INRAE

Pourquoi et comment faire FAIR ?

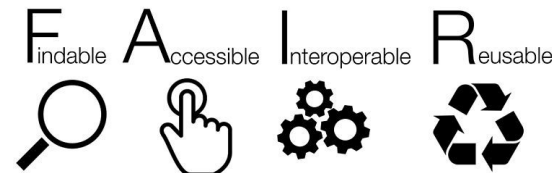


2025 - Équipe de recherche qui veut entraîner son modèle -> besoin de jeux de données Mais

- Comment **accéder** à ce jeu de données ? Qui peut y accéder ?
 - **A : Accessible**
Permettre l'accès aux données et leur téléchargement, ce qui peut inclure l'authentification et l'autorisation
 - Les données sont **accessibles à travers un protocole** de communication standard
 - Ce protocole est **libre et ouvert**
 - Ce protocole permet un accès par **authentification si besoin**
 - Les métadonnées restent accessibles même si les données ne le sont pas (disparues ou inaccessibles)

Donnée ouverte ne veut pas dire accessible n'importe qui !
“Données ouvertes autant que possible et fermées autant que nécessaire”

Pourquoi et comment faire FAIR ?



2025 - Équipe de recherche qui veut entraîner son modèle -> besoin de jeux de données Mais

- Comment **interpréter** les données ? Comment les **lier** à d'autres jeux de données ?

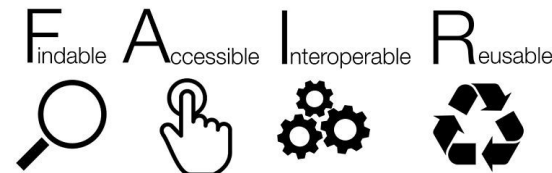
- **I : Interopérable**

Permettre l'exploitation et l'intégration des données quel que soit l'environnement informatique utilisé

- Les données sont décrites avec un **vocabulaire contrôlé**
- Le vocabulaire utilisé respecte les principes FAIR
- Les **métadonnées sont contextualisées** avec des liens vers d'autres données

Importance de la **qualité des métadonnées**, d'utiliser des **métadonnées de référence** !

Pourquoi et comment faire FAIR ?



2025 - Équipe de recherche qui veut entraîner son modèle -> besoin de jeux de données Mais

- Comment **ré-utiliser** ces données dans un autre contexte, interdisciplinaire ?

- **R : Réutilisable**

Permettre la réutilisation des données pour de futures recherches

- Les **métadonnées** contiennent toutes les informations qui peuvent être utiles
- Une **licence de réutilisation** est attribuée aux données
- La description des données indique leur **provenance**
- Le partage des données suit les standards de la communauté scientifique

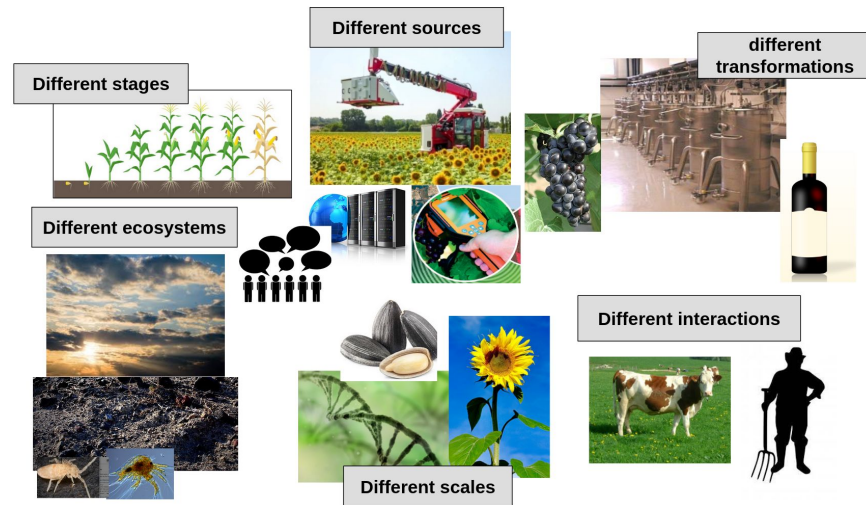
Donnée ouverte ne veut pas dire accessible n'importe qui !

“Données ouvertes autant que possible et fermées autant que nécessaire”

Défi des données dans l'agriculture et les sciences de l'environnement

Les Expérimentations ou Observations

- Sont complexes et ne peuvent pas être reproduites
- Nombreux besoins en ressources diverses
- Ensembles de données énormes et très **complexes**
- **Contexte interdisciplinaire**



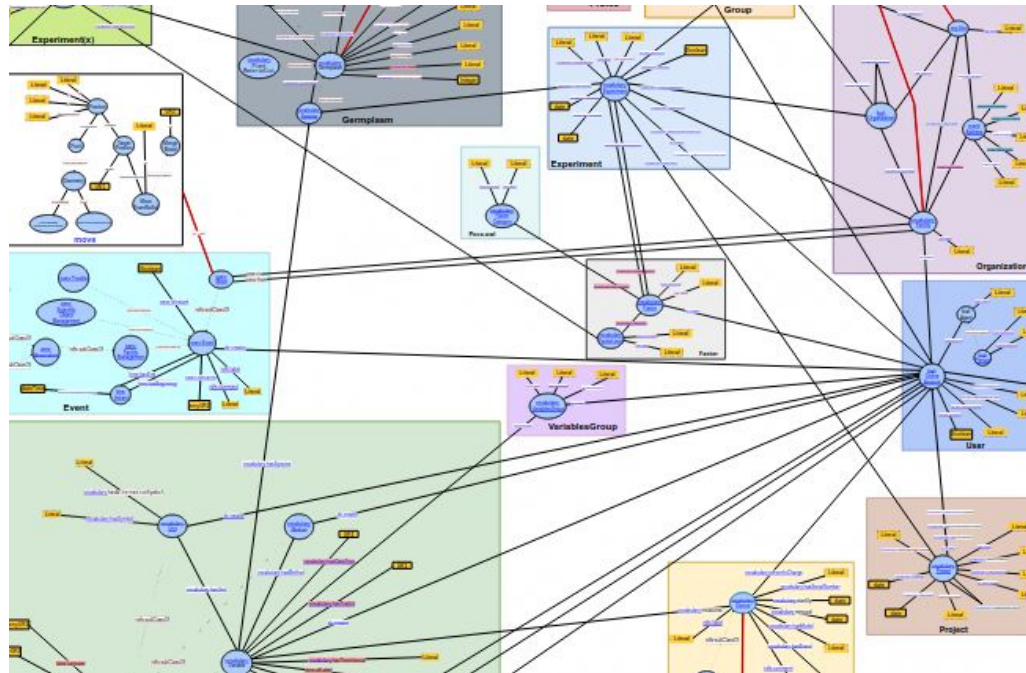
-> Besoins importants de **transparence** et de **reproductibilité** du traitement des données

Valoriser les données : ré-analyses, méta-analyses et nouvelles analyses
→ impossible sans **une gestion avancée des données**



Une approche de gestion des données

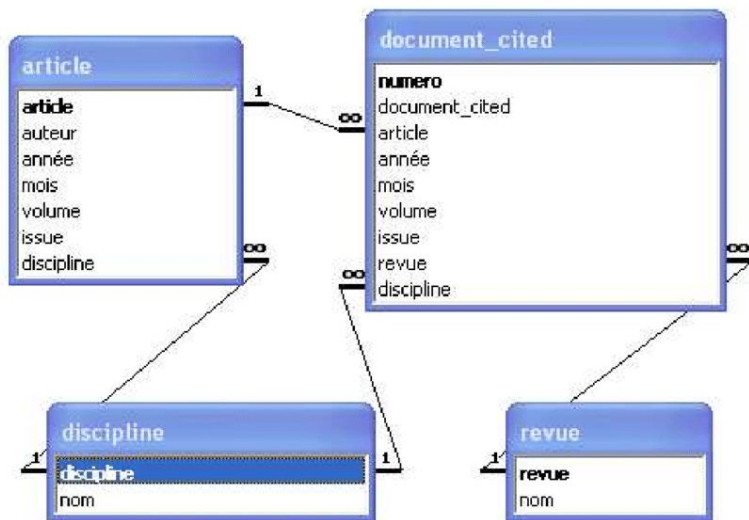
Linked Data (lisibles par machine) : utilisent des **ontologies** et permettent de construire des ensembles de données pour la découverte de connaissances, l'offre de services, la validation de résultats, l'analyse prescriptive, etc.



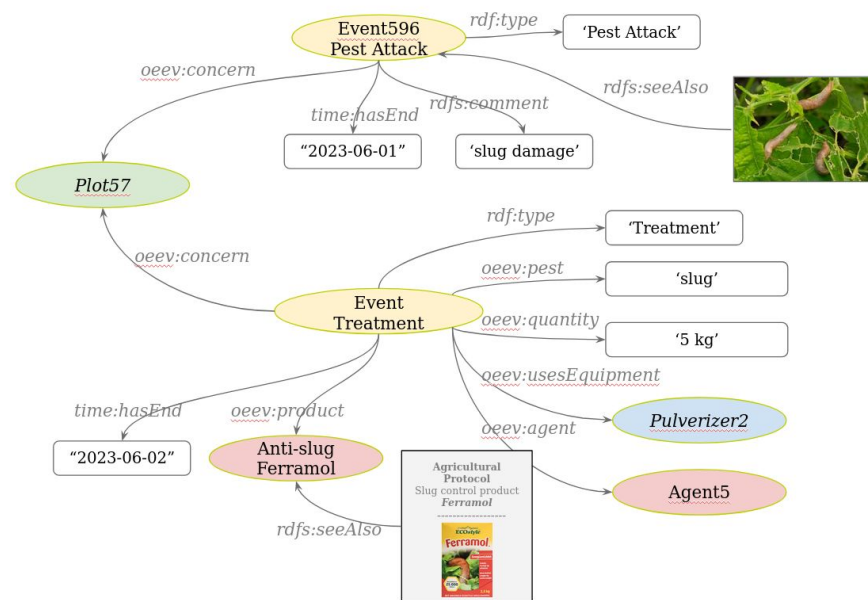
Données FAIR : quelques définitions

- **Métadonnée** : est une **donnée servant à définir ou décrire une autre donnée**, quel qu'en soit le support (papier, électronique ou autre).
- **Ontologie** : Une ontologie **formalise les concepts et les termes** utilisés pour **décrire un domaine**, et **formalise les relations sémantiques** entre les différents concepts (“A au moins un”, “Utilise”, “Est une partie de”...)
- **Sémantique** : La sémantique est l’étude des **différents sens des mots et du langage**. Dans le cas du **Web Sémantique**, il s’agit notamment de **donner du sens aux liens** qui sont créés entre les différentes entités.

Base de données relationnelles vs Web Sémantique



Exemple BDD Relationnelle



Exemple Web Sémantique

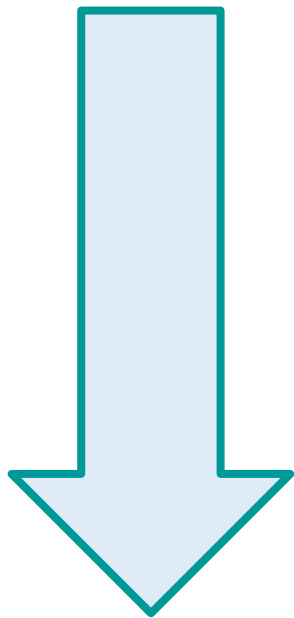
Web des données

Les données sont :

- Interconnectées
- Lisibles par machine
- Structurées par des ontologies

Cela permet :

- Une découverte plus facile
- La validation de résultats
- L'analyse prescriptive
- etc.



Structurer les données...

en ajoutant de la sémantique...

en les liant...

augmente leur valeur.

Nouveaux Systèmes d'Information



Nos objectifs

- Fournir des méthodes et composants pour la mise en œuvre de systèmes d'information
- Fournir des outils et services pour :
 - la collecte
 - l'organisation
 - le stockage
 - l'échange
 - l'exploration

de l'information

- S'adapter à diverses communautés


Créez des données FAIR, structurez vos données

Identification

- Tout peut être identifié : plantes, expériences, capteurs, événements, etc.
- Persistant, sans ambiguïté, résolvable

Sémantique

- Conventions de nommage
- Vocabulaire contrôlé
- Relations formalisées entre les entités
- Annotation et enrichissement des données



URI of plant:
<<http://phenome.fr/arch/2017/c17000118>>

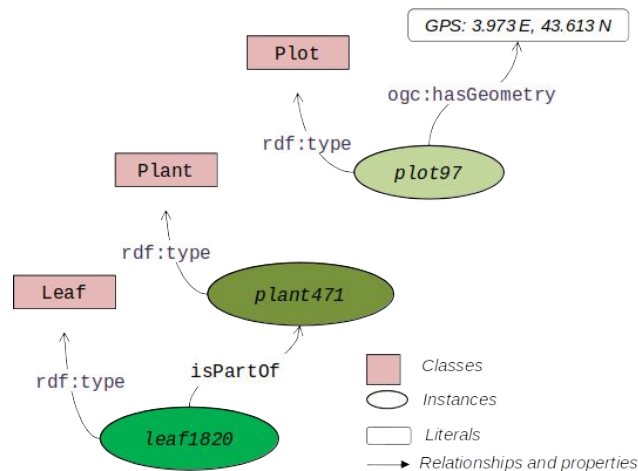
URI of pot:
<<http://phenome.fr/arch/2013/pc13001542>>

URI of cart:
<<http://phenome.fr/arch/2013/ct1300123>>

URI of cabin:
<<http://phenome.fr/arch/2018/ac180015>>

URI of camera:
<<http://phenome.fr/arch/2018/ac180019>>

URI of image: <<http://phenome.fr/arch/2017/ic17002295855>>



Créez des données FAIR, structurez vos données

Identification

- Tout peut être identifié : plantes, expériences, capteurs, événements, etc.
- Persistant, sans ambiguïté, résolvable

Sémantique

- Conventions de nommage
- Vocabulaire contrôlé
- Relations formalisées entre les entités
- Annotation et enrichissement des données

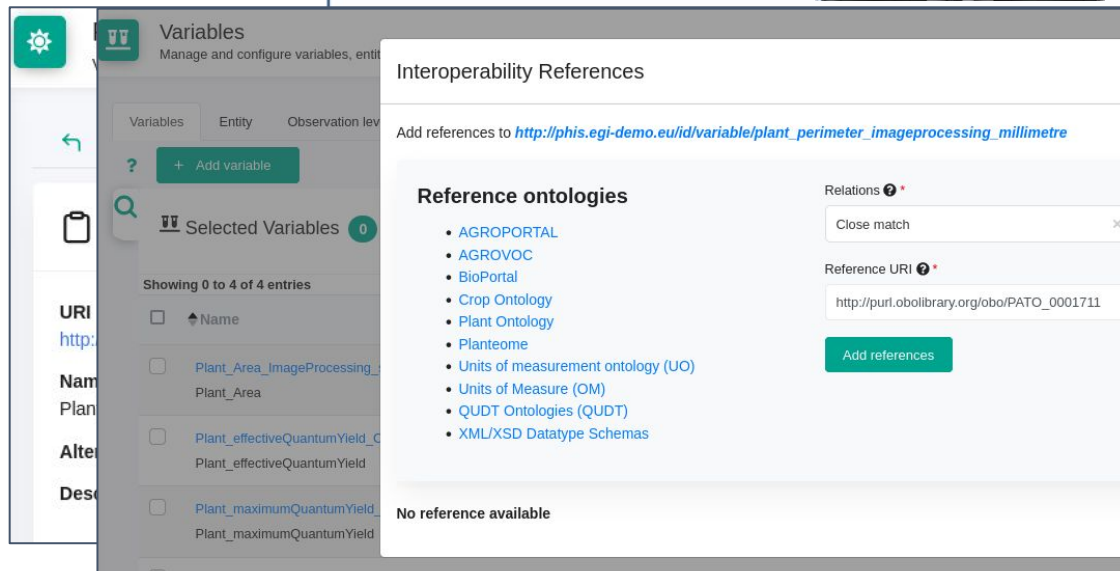
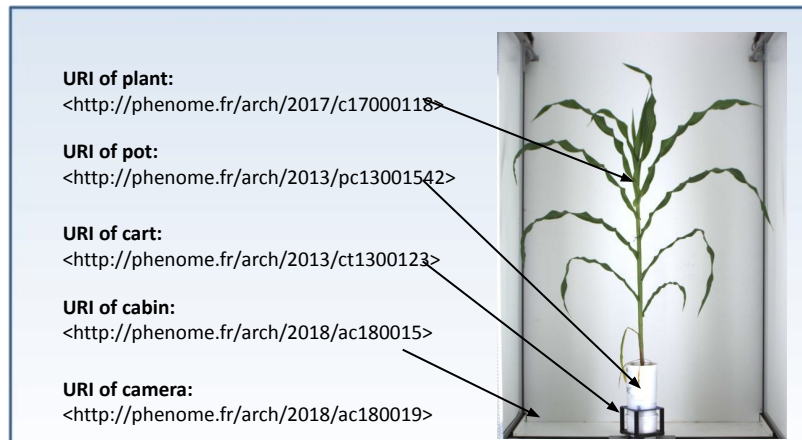
Comment ?

OpenSILEX, un système d'information piloté par des ontologies

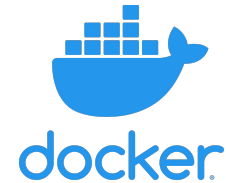
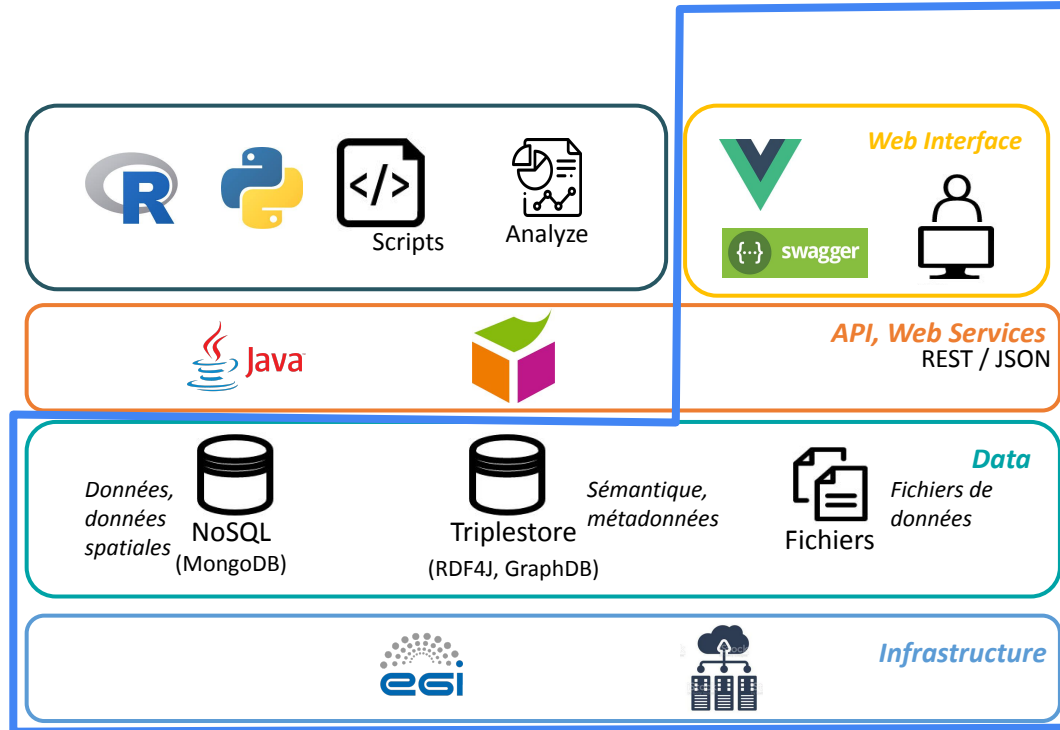


INRAE

OpenSILEX
Séminaire interne MISTEA - 15 décembre 2025



OpenSILEX - Architecture simplifiée



Galaxie OpenSILEX - les communautés bénéficiaires



Phénotypage des plantes et suivi de cultures



Sixtine

Grandes cultures
Essais Systèmes
Arboriculture - Vigne

Agro-Écologie



OpenSILEX

Viticulture - Oenologie

Vitis
Explorer



VitiOeno

Instance Ressources
Numériques

Environnement et procédés



Ferments du futur



INRAE

OpenSILEX
Séminaire interne MISTEA - 15 décembre 2025

Insectes



Bio-agresseurs



Galaxie OpenSILEX - les communautés bénéficiaires



Phénotypage des plantes et suivi de cultures



Grandes cultures
Essais Systèmes
Arboriculture - Vigne

Vigne
LaCage GCIE
Arbo
U2E

Agro-Écologie



~30 systèmes d'information
~200 utilisateurs
Plusieurs Peta de données

OpenSILEX

Viticulture - Oenologie

Vitis
Explorer



Environnement et procédés



Ferments du futur



INRAE

OpenSILEX
Séminaire interne MISTEA - 15 décembre 2025

Insectes



Bio-agresseurs



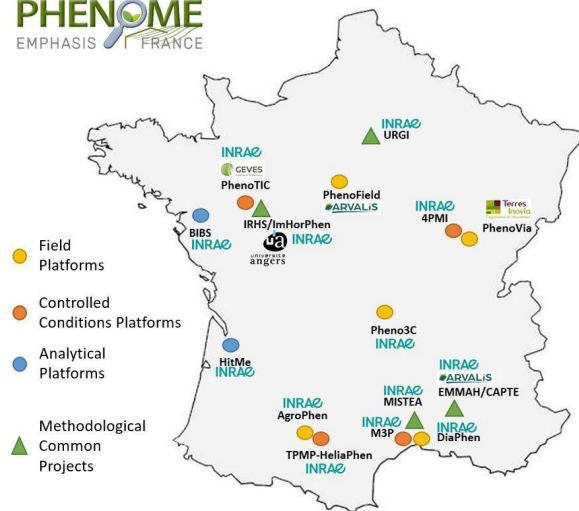


OpenSILEX pour le phénotypage végétal

PHIS : Phenotyping Hybrid Information System

- **Phénotype** : ensemble des caractères apparents d'un individu, correspondant à une réalisation du **génotype**.
- Le phénotype est l'expression du génotype dans un **environnement particulier**
=> Importance de connaître l'environnement
- **Phénotypage Haut-Débit** : techniques permettant d'acquérir énormément de données très rapidement, et de mesurer en continu la croissance et le développement des plantes.

PHENOME
EMPHASIS FRANCE

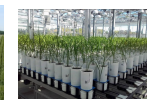


INRAE

ARVALIS

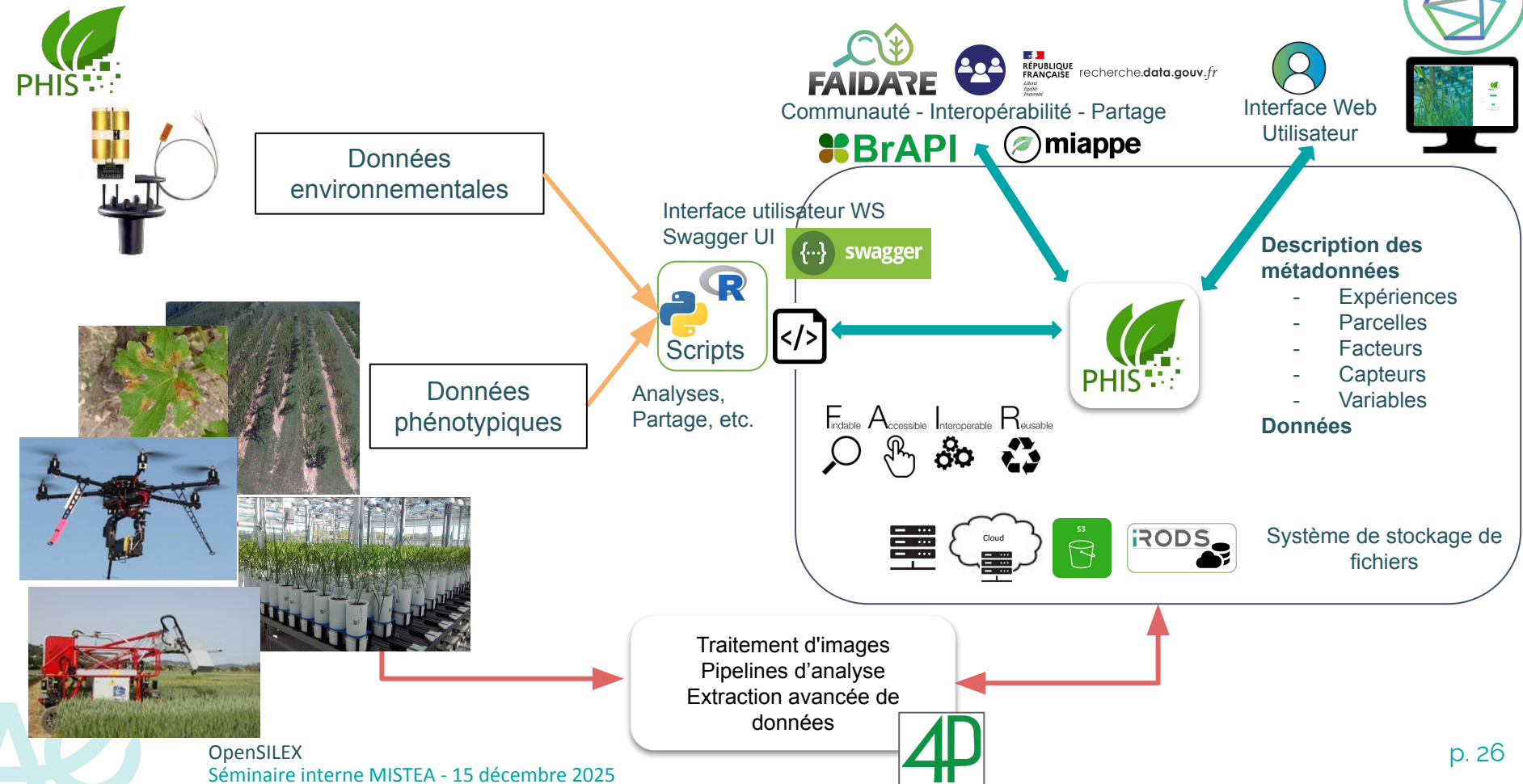
Terres Inovia

GEVES



INRAE

PHIS : un système d'information piloté par des ontologies dédié au phénotypage végétal



- 6 systèmes d'informations déployés
- Projets menés sur plusieurs plateformes

=> Nécessité d'avoir des vocabulaires partagés



PHIS - Diaphen

Plateforme Mauguio

<https://phenome.inrae.fr/diaphen/app/>

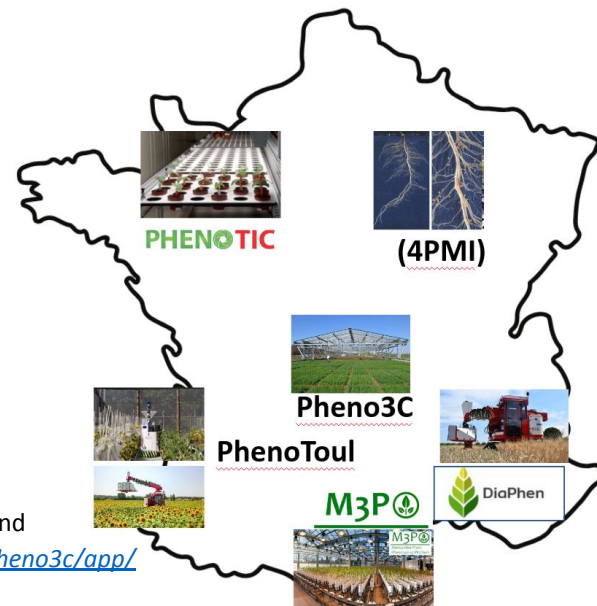
?



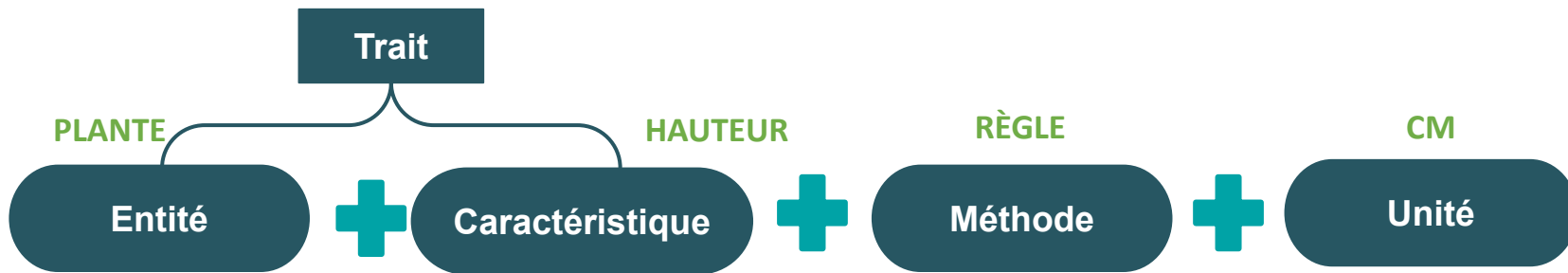
PHIS - Pheno3c

Plateforme Clermont-Ferrand

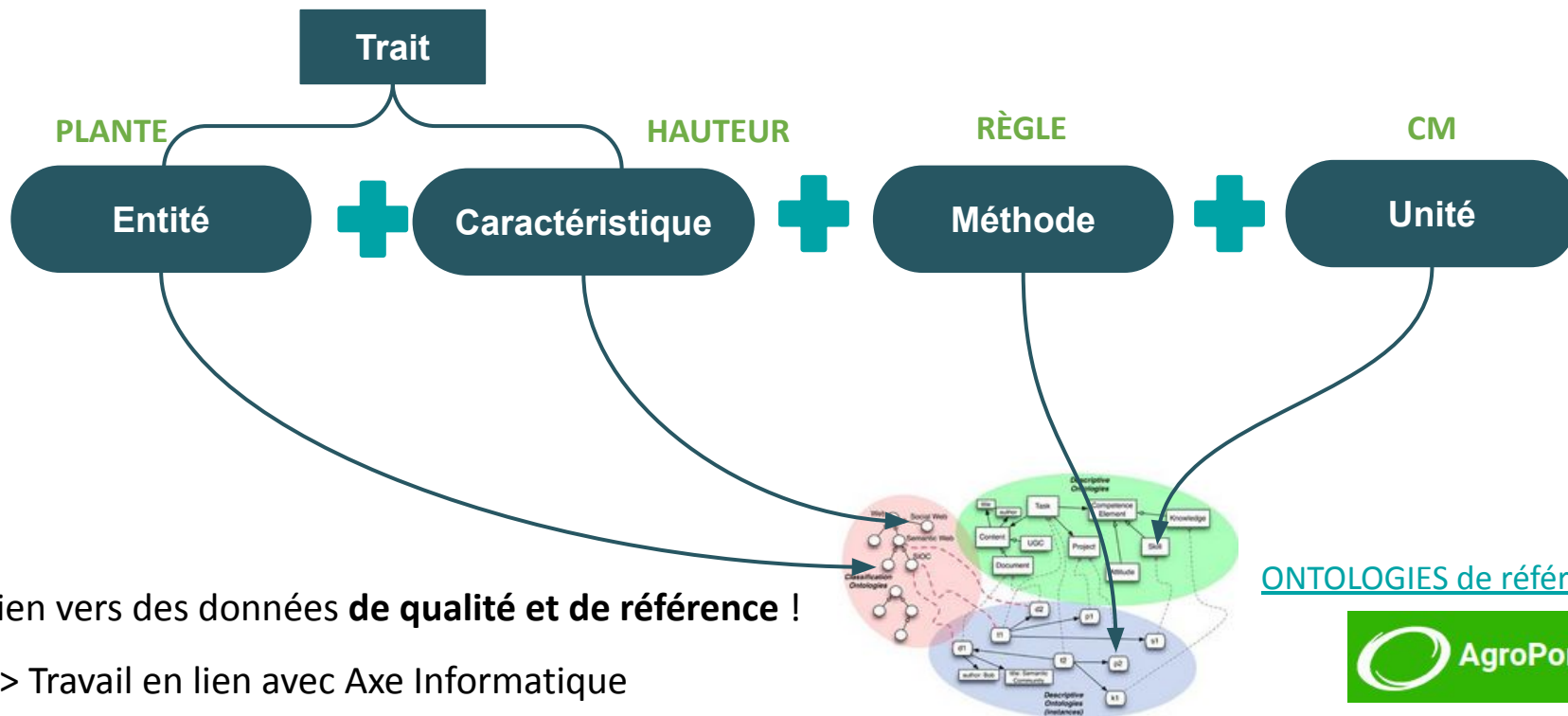
<https://phenome.inrae.fr/pheno3c/app/>



Un modèle de Variables partagé



Un modèle de Variables partagé



ONTOLOGIES de référence



INRAE

Démonstration



You can connect as guest

Connect as guest

 Email or URI

 Password

Forgot your password ?

Login

PHIS - Phenotyping Hybrid Information System
Version 1.4.9-rdg
Based on OpenSILEX version 1.4.9-rdg
Copyright ©2021 INRAE



SilexExplorer

**Rendre les données OpenSILEX accessibles, intégrées et
prêtes pour l'analyse**

Un service GraphQL + packages R / Python au service de la recherche



Pourquoi SilexExplorer ?

❖ Des données riches mais dispersées

- Données génétiques, environnementales, phénotypiques...
- **Hétérogènes** et stockées dans des bases différentes (RDF / JSON)
- Pourtant **complémentaires** pour l'analyse

➡ Les chercheurs ont besoin de les intégrer **sans gérer la complexité technique**.

❖ Trop de temps perdu à assembler les données

- Les services REST d'OpenSILEX sont nombreux : expériences, objets scientifiques, capteurs, variables, facteurs...
- Les chercheurs doivent **recupérer, nettoyer, fusionner** les résultats eux-mêmes

➡ Long, fastidieux, non reproductible.



C'est quoi SilexExplorer

Un service pour explorer facilement les données OpenSILEX

❖ Service GraphQL :

un point d'accès unique pour accéder à l'ensemble des données.

❖ Packages R & Python (SilexExplorerR, SilexExplorerPy)

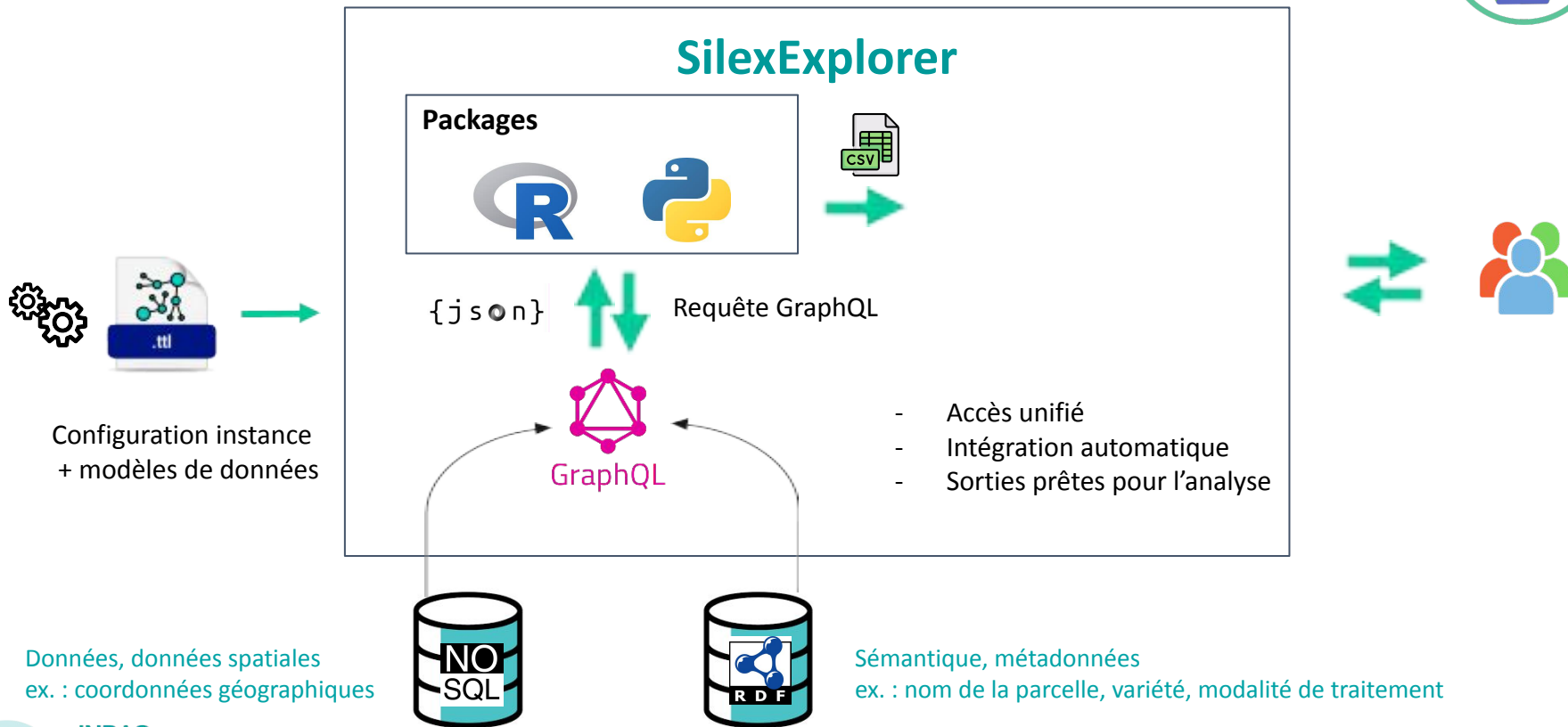
→ filtrage, extraction, structuration, production des tables.

→ **Orienté utilisateur :**

- Tableaux prêts à être exploités pour l'analyse
- Possibilité d'écrire vos propres requêtes GraphQL

```
query MyQuery {  
  Plant {  
    _id  
    _type  
    label  
    geometry {  
      graph  
      type  
      coordinates  
    }  
  }  
}
```

Comment ça marche ?



INRAE



Illustration — Détection des groupes de répétition

Objectif

- Récupérer les objets scientifiques d'une expérience et leurs caractéristiques (ressource génétique, modalités...)
- Regrouper automatiquement les objets scientifiques présentant les mêmes caractéristiques
- Récupérer les données mesurées pour chaque groupe
- Visualiser ensuite la variabilité intra- et inter-groupe



Illustration — Détection des groupes de répétition

Étape 1 — Récupérer les objets scientifiques de l'expérience

Fonction utilisée : `get_os_by_exp`

Ce qu'elle retourne

- liste des objets scientifiques
- leurs caractéristiques :
 - type (Plante, Parcelle...)
 - modalités (niveau de facteur)
 - ressources génétiques

uri	name	type	Irrigation	Germplasm_type_1	Germplasm_1	Species_1
http://www.phenome-fppn.fr/m3p/arch/2017/c17000001	0001/DZ_PG_47/ ZM4362/test:id/ factor/za17- bis.irrigation.wd/ Veg_1/01_01/ ARCH2017-03-30	Plant	WD	SeedLot	ZM4362	Zea mays
http://www.phenome-fppn.fr/m3p/arch/2017/c17000002	0002/DZ_PG_07/ ZM4351/test:id/ factor/za17- bis.irrigation.wd/ Veg_1/01_02/ ARCH2017-03-30	Plant	WD	SeedLot	ZM4351	Zea mays
http://www.phenome-fppn.fr/m3p/arch/2017/c17000003	0003/DZ_PG_26/ ZM4387/test:id/ factor/za17- bis.irrigation.wd/ Veg_1/01_03/ ARCH2017-03-30	Plant	WD	SeedLot	ZM4387	Zea mays
http://www.phenome-fppn.fr/m3p/arch/2017/c17000004	0004/DZ_PG_12/ ZM1304_lot1/ test:id/factor/	Plant	WD	SeedLot	ZM1304_lot1	Zea mays



Illustration — Détection des groupes de répétition

Étape 2 — Regrouper les objets scientifiques partageant les mêmes caractéristiques

Fonction utilisée : `group_dict`

Ce qu'elle retourne

Un tableau contenant :

- l'ID du groupe (caractéristiques communes)
- le nombre d'objets dans chaque groupe

Group	Number of Elements
NaN_group	4
WD_ScionLot_ZM1289_Zea mays	10
WD_ScionLot_ZM3234_Zea mays	2
WD_ScionLot_ZM4112_Zea mays	16
WD_ScionLot_ZM4373_Zea mays	10
WD_ScionLot_ZM4383_Zea mays	10
WD_ScionLot_ZM4453_Zea mays	8
WD_ScionLot_ZM1304_lot1_Zea mays	10
WD_ScionLot_ZM4342_Zea mays	10
WD_ScionLot_ZM4343_Zea mays	10
WD_ScionLot_ZM4344_Zea mays	10
WD_ScionLot_ZM4345_Zea mays	10
WD_ScionLot_ZM4346_Zea mays	10
WD_ScionLot_ZM4347_Zea mays	10
WD_ScionLot_ZM4348_Zea mays	10

Illustration — Détection des groupes de répétition



Étape 3 — Récupérer les données mesurées sur chaque groupe

Fonction utilisée : `get_data_os`

Ce qu'elle retourne

- Des données organisées par variable
- Pour chaque variable :
 - identifiant de l'objet scientifique
 - valeurs mesurées
 - date

URI	Leaf_Area_ImageAnalysis_squareMeter	Date
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.305309818	2017-05-15T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.263307888	2017-05-13T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.253655772	2017-05-12T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.230782124	2017-05-10T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.198242297	2017-05-08T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.179725674	2017-05-07T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.153705557	2017-05-05T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.138666844	2017-05-04T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.109103799	2017-05-02T00:00:00.000Z
http://www.phenome-fppn.fr/m3p/arch/2017/c17000039	0.08927577	2017-05-01T00:00:00.000Z



Illustration — Détection des groupes de répétition

Analyses possibles à partir des groupes :

❖ Analyse intra-groupe:

Étudier la variabilité au sein d'un même groupe

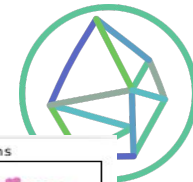
- cohérence des objets scientifiques partageant les mêmes caractéristiques
- détecter les écarts, anomalies, comportements atypiques

❖ Analyse inter-groupes

Comparer des groupes entre eux

- évaluer l'effet des modalités / traitements (ex. WD vs WW)
- détecter des différences significatives de réponse par variable

Illustration — Détection des groupes de répétition



Analyse intra-groupe

Variabilité au sein d'un groupe

But : Visualiser la dispersion des objets scientifiques appartenant à un même groupe

Sortie : Pour chaque variable, un graphique regroupant les courbes individuelles des objets scientifiques du groupe.

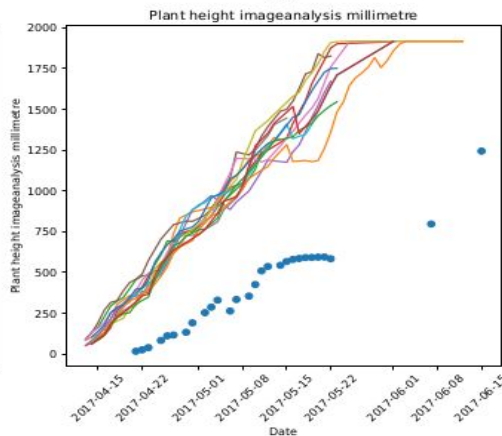
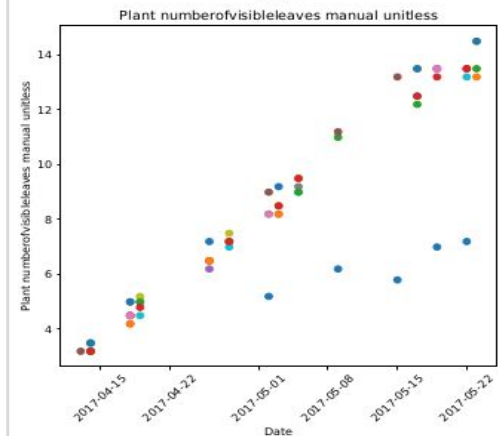
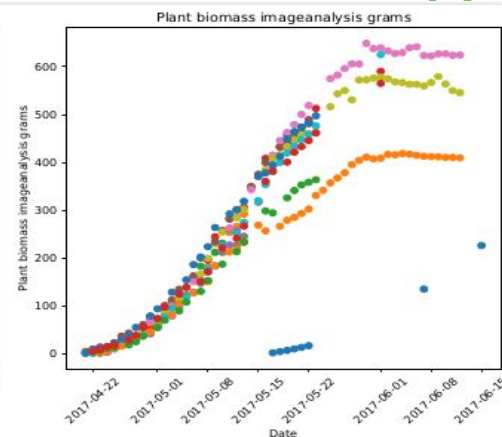
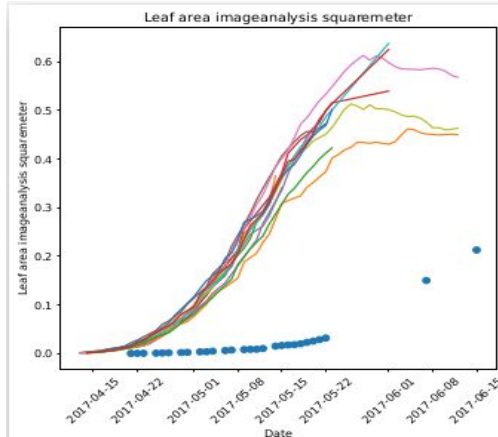


Illustration — Détection des groupes de répétition



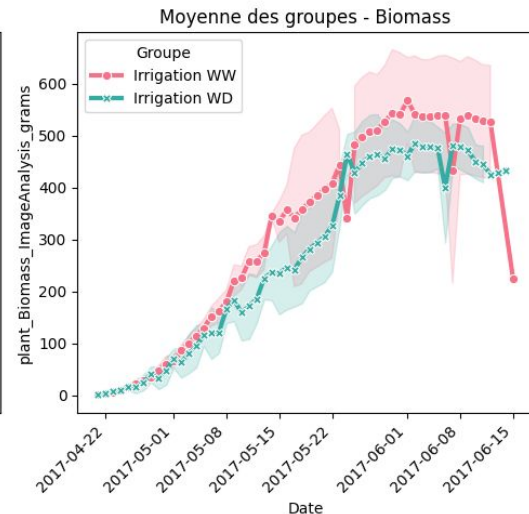
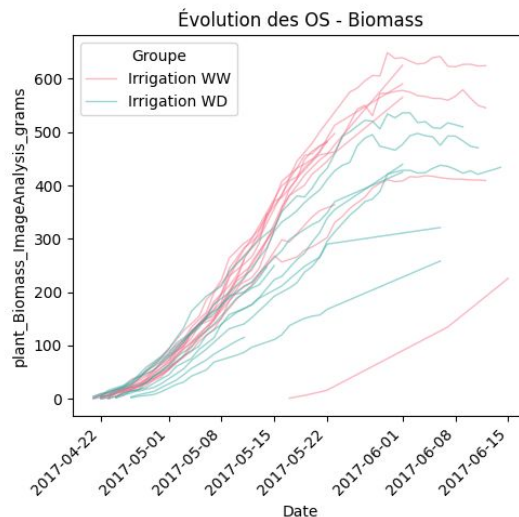
Analyse inter-groupes — Comparaison entre groupes

But : Comparer l'évolution des variables entre des groupes soumis à différents traitements (ex. modalités d'irrigation).

Sortie :

Pour chaque variable, un graphique affichant :

- **les courbes individuelles** des objets scientifiques de chaque groupe,
- **la moyenne** de chaque groupe,
- **l'écart-type**, illustrant la variabilité entre groupes.





Perspectives

- Collaborations bienvenues : vos cas d'étude enrichissent directement l'outil
- Développement conjoint de nouvelles analyses
- Intégration en cours avec StatGenHTP (analyses phénotypage à haut débit)

Conclusion



❖ OpenSILEX

- Une **production logicielle** mais pas que !
- Un accompagnement à l'utilisation du logiciel (**formation, support**)
- Un appui des communautés pour une **meilleure formalisation des concepts** (à l'aide d'ontologies) et **des données nécessaire dans la recherche interdisciplinaire**
- Des **formations à la gestion de données**

❖ Formalisation des concepts :

- **Les variables**, un sujet complexe
=> véritable **question de recherche**, menée en lien avec l'**axe Informatique**

❖ Exploration des données - analyse et modélisation - Collaborations à construire

- Faire des lien avec des modèles (sélection variétale, ...)
- Confronter les modèles à la réalité des données
- Utiliser des données OpenSILEX pour des modèles / statistiques

OpenSILEX : liens utiles



- ❖ Site web OpenSILEX : <http://opensilex.org/>
- ❖ Contact : opensilex@inrae.fr
- ❖ Instance Demo OpenSILEX : <http://opensilex.org/sandbox/app/>
- ❖ Documentation utilisateur : <http://opensilex.org/doc/GuideUtilisateurOpenSilex.pdf>
- ❖ Comment contribuer ?
 -  ➤ Github repository : <https://github.com/OpenSILEX/>
 - Documentation développeur : <https://opensilex.github.io/docs-community-dev/>
 -  ➤ OpenSILEX Docker : <https://github.com/OpenSILEX/opensilex-docker-compose>
- ❖ Vidéo de présentation de la Version 1.0.0-rc+5 : https://youtu.be/jB9Hqh0Wn_Y
- ❖ Tutoriel de déclaration des variables : <https://youtu.be/d1pGuh8PAKA>
- ❖ SilexExplorer
 -  ➤ SilexExplorerPy : <https://github.com/OpenSILEX/silex-explorer-py>
 -  ➤ SilexExplorerR : <https://github.com/OpenSILEX/silex-explorer-r>