

INRAE

> FAIR-DS

**Un logiciel de gestion de mesures expérimentales
ouvert et flexible**

Réunion des DU de PHASE

08/02/2024

François Moreews

Gwendal Virlet

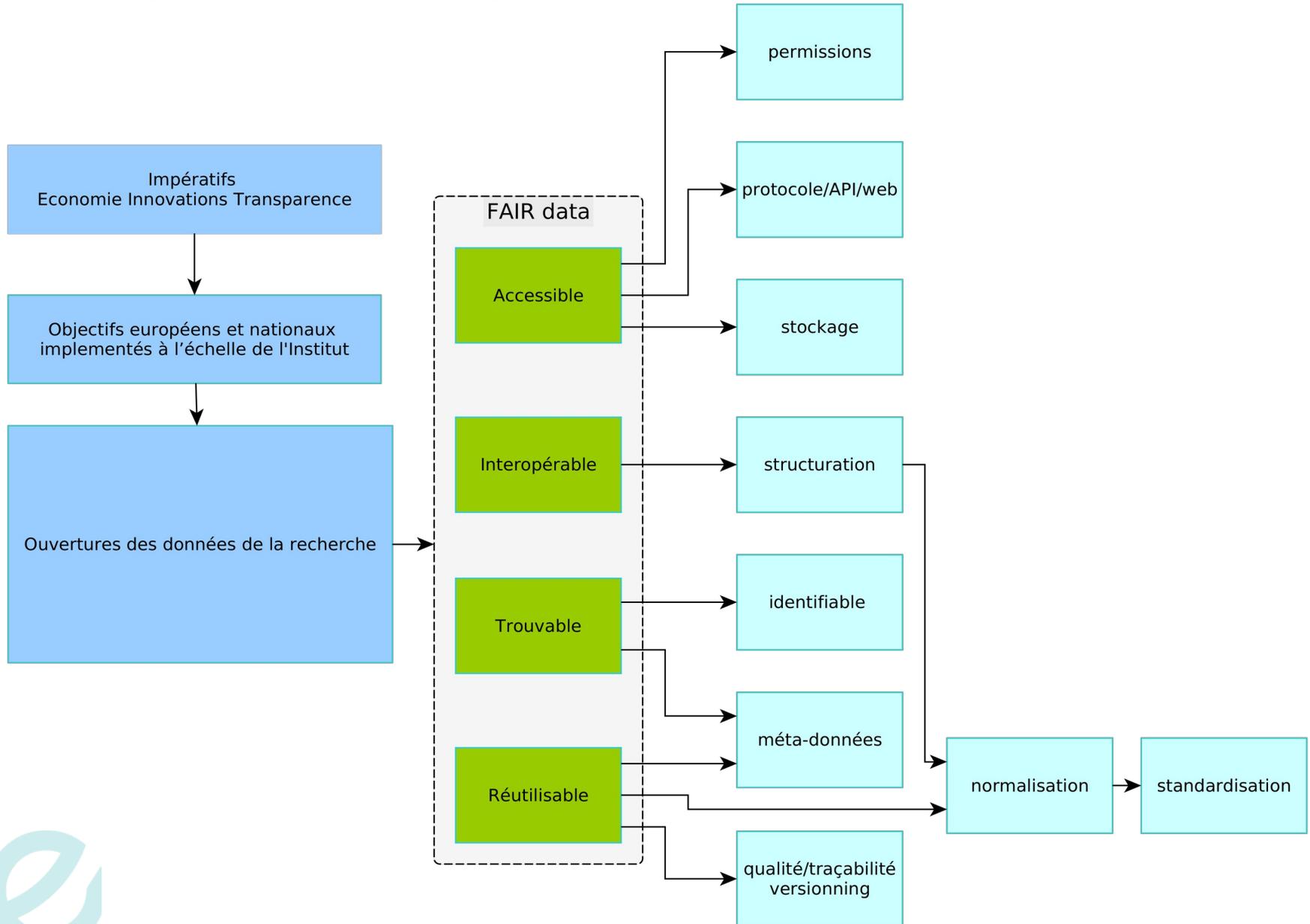
Sommaire

FAIR-DS – un logiciel de gestion de mesures expérimentales ouvert et flexible

- ◆ Enjeux et objectifs
- ◆ L'organisation du projet
- ◆ Matériel et méthodes
- ◆ Résultats et présentation de l'outil
- ◆ Analyse des freins et solutions



Enjeux et objectifs



Enjeux et objectifs

FAIR-DS – un logiciel de gestion de mesures expérimentales ouvert et flexible

◆ Objectifs du logiciel

- ◆ Ouverture des données produites à l'échelle du laboratoire conformément aux directives
 - ◆ Publication (DataVerse)
- ◆ Réutilisation des données de la recherche
 - ◆ En interne
 - ◆ Avec les partenaires externes
 - ◆ Avec la communauté scientifique
- ◆ Garantir la qualité et la traçabilité des données de la recherche
- ◆ Garantir la confidentialité et la propriété des données
 - ◆ “aussi ouvert que possible, aussi fermé que nécessaire”



Organisation du projet

FAIR-DS – un logiciel de gestion de mesures expérimentales ouvert et flexible

- ◆ Les parties prenantes

- ◆ Comité de pilotage :

- ◆ DU et scientifiques, groupe data

- ◆ Groupe Données:

- ◆ F. Herault, C. Mustiere, E. Bernard, F. Moreews , G. Virlet

- ◆ Equipe de développement :

- ◆ G. Virlet , F. Moreews + stagiaires



Organisation du projet

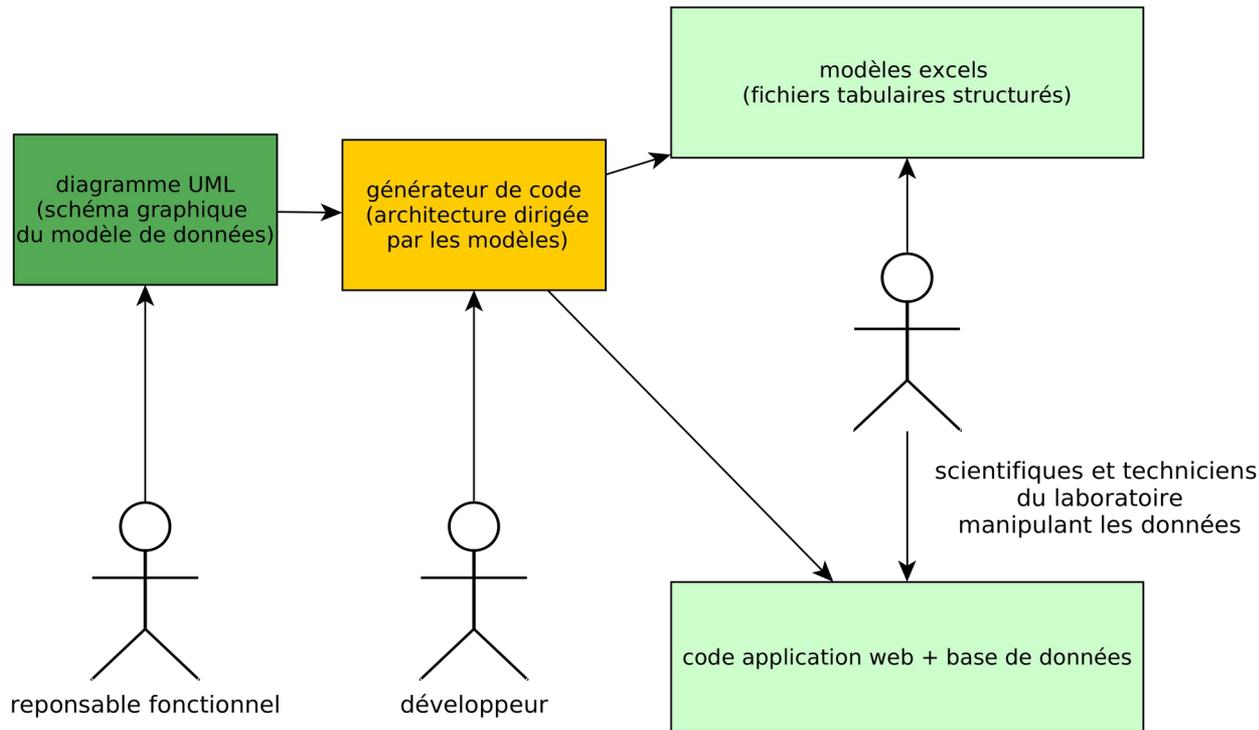
Historique : normalisation des données et qualité

- ◆ 2019 :
 - ◆ Phase initiale, état des lieux,
 - ◆ Inventaire des données
 - ◆ Création du diagramme des types de données expérimentales (UML)
 - ◆ Création des templates (modèles Excel) : début de structuration
- ◆ 2022 :
 - ◆ Début du développement de FAIR-DS , application web et base de données
- ◆ 2024 :
 - ◆ Mise en production d'une V1 et retours utilisateurs



Matériel et méthode

Des applications reconfigurables grâce à l'approche orientée modèle

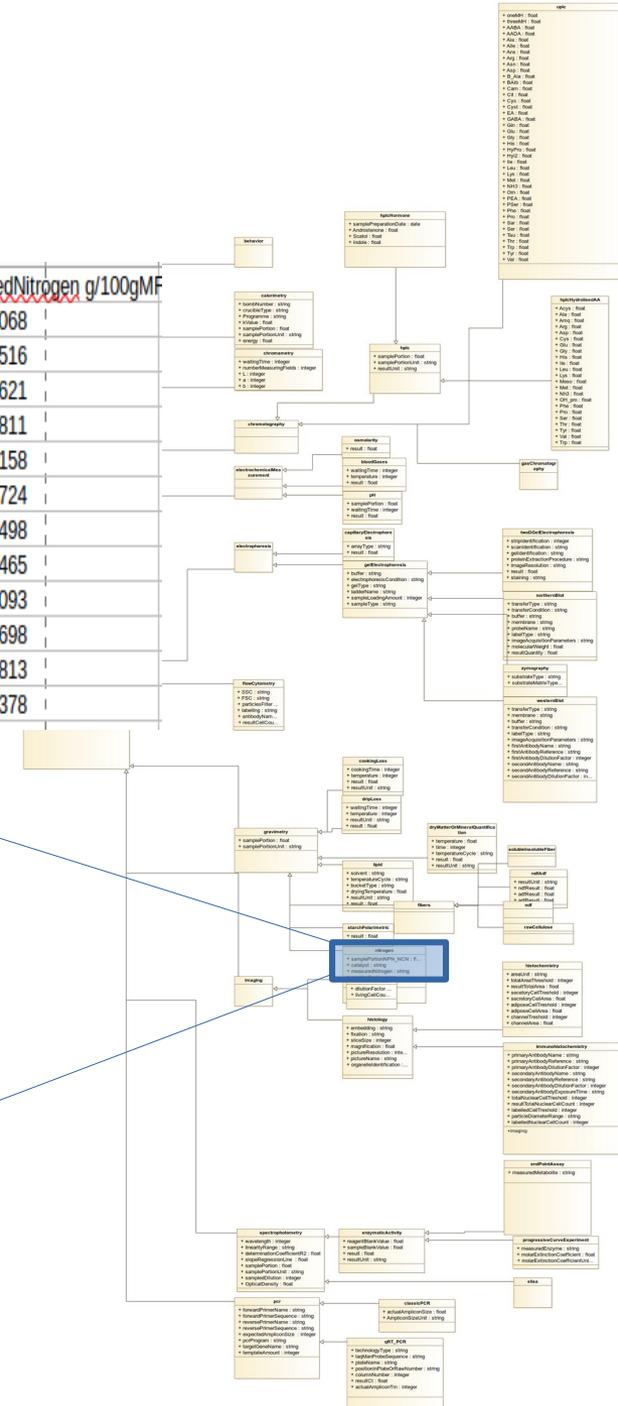
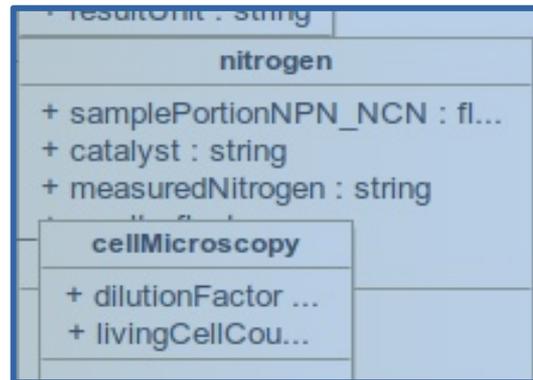


- + Code de qualité et évolutif
- + Modifications rapide du modèle de données et régénération de l'application a posteriori sans code
- Lenteur du développement initiale
- Prise en main complexe pour le développeur



Matériel et méthode

Laboratory	OperatingMode	CriticalAppar	CriticalProdu	CriticalProdu	RawDataPath	SampleID	SamplePortio	Result	Catalyst	MeasuredNitrogen g/100gMF
Dumas	FP-828	Balance DQ			PC Léco FP-	L2303 A68252	0.1126 g	3.02531068		
Dumas	FP-828				PC Léco FP-	L2303 A68252	0.1246 g	3.03674516		
Dumas	FP-828				PC Léco FP-	L2303 A68257	0.1019 g	1.08376621		
Dumas	FP-828				PC Léco FP-	L2303 A68257	0.1200 g	1.09372811		
Dumas	FP-828				PC Léco FP-	L2303 A68258	0.1153 g	1.02197158		
Dumas	FP-828				PC Léco FP-	L2303 A68258	0.1160 g	1.02538724		
Dumas	FP-828				PC Léco FP-	L2303 A68259	0.1000 g	1.31088498		
Dumas	FP-828				PC Léco FP-	L2303 A68259	0.1161 g	1.32879465		
Dumas	FP-828				PC Léco FP-	L2303 A68260	0.1047 g	1.43991093		
Dumas	FP-828				PC Léco FP-	L2303 A68260	0.1119 g	1.44892698		
Dumas	FP-828				PC Léco FP-	L2303 A68261	0.1736 g	2.76066813		
Dumas	FP-828				PC Léco FP-	L2303 A68261	0.2203 g	2.84771378		



Résultat et présentation de l'outil

Accessible : Sécurité – Gérer l'accès aux données finement tout au long de leur cycle de vie

- ◆ Base sécurisée par le LDAP INRAE

- ◆ Données sécurisées par type de mesure
 - ◆ lecture / écriture

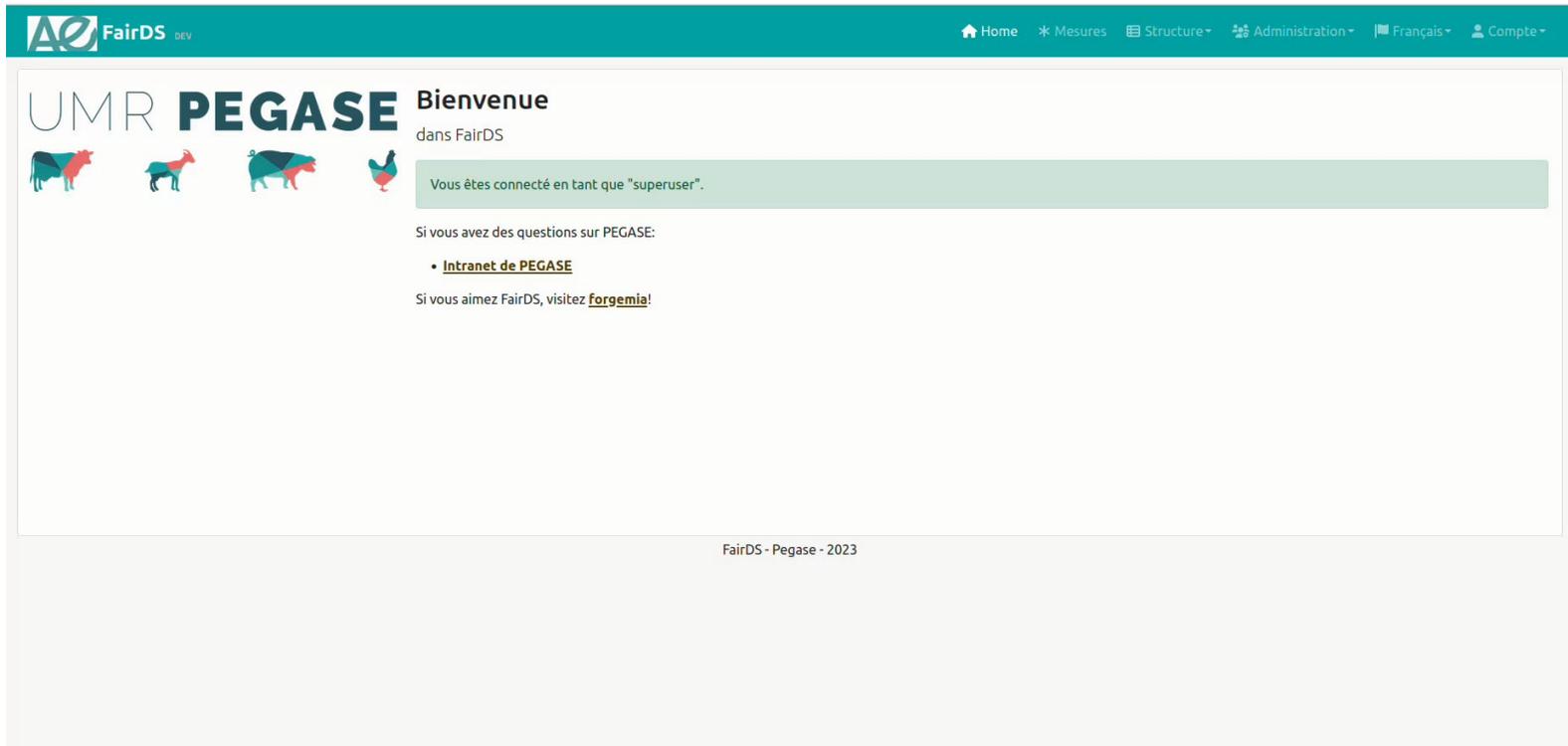
- ◆ Données sécurisées par la propriété
 - ◆ lecture / écriture

- ◆ Données sécurisées par le partage
 - ◆ lecture / écriture



Résultat et présentation de l'outil

Accessible : Sécurité – Gérer l'accès aux données finement tout au long de leur cycle de vie



The screenshot shows the user interface of the FairDS PEGASE application. At the top, there is a teal navigation bar with the FairDS logo and the text 'FairDS dev' on the left, and a menu on the right containing 'Home', 'Mesures', 'Structure', 'Administration', 'Français', and 'Compte'. Below the navigation bar, the main content area features the text 'UMR PEGASE Bienvenue dans FairDS' and four small icons representing different farm animals: a cow, a horse, a pig, and a chicken. A green notification box states 'Vous êtes connecté en tant que "superuser".'. Below this, there is a section for questions about PEGASE, with a bullet point for 'Intranet de PEGASE'. At the bottom of the main content area, it says 'Si vous aimez FairDS, visitez [forgemia!](#)'. The footer of the page contains the text 'FairDS - Pegase - 2023'.



Résultat et présentation de l'outil

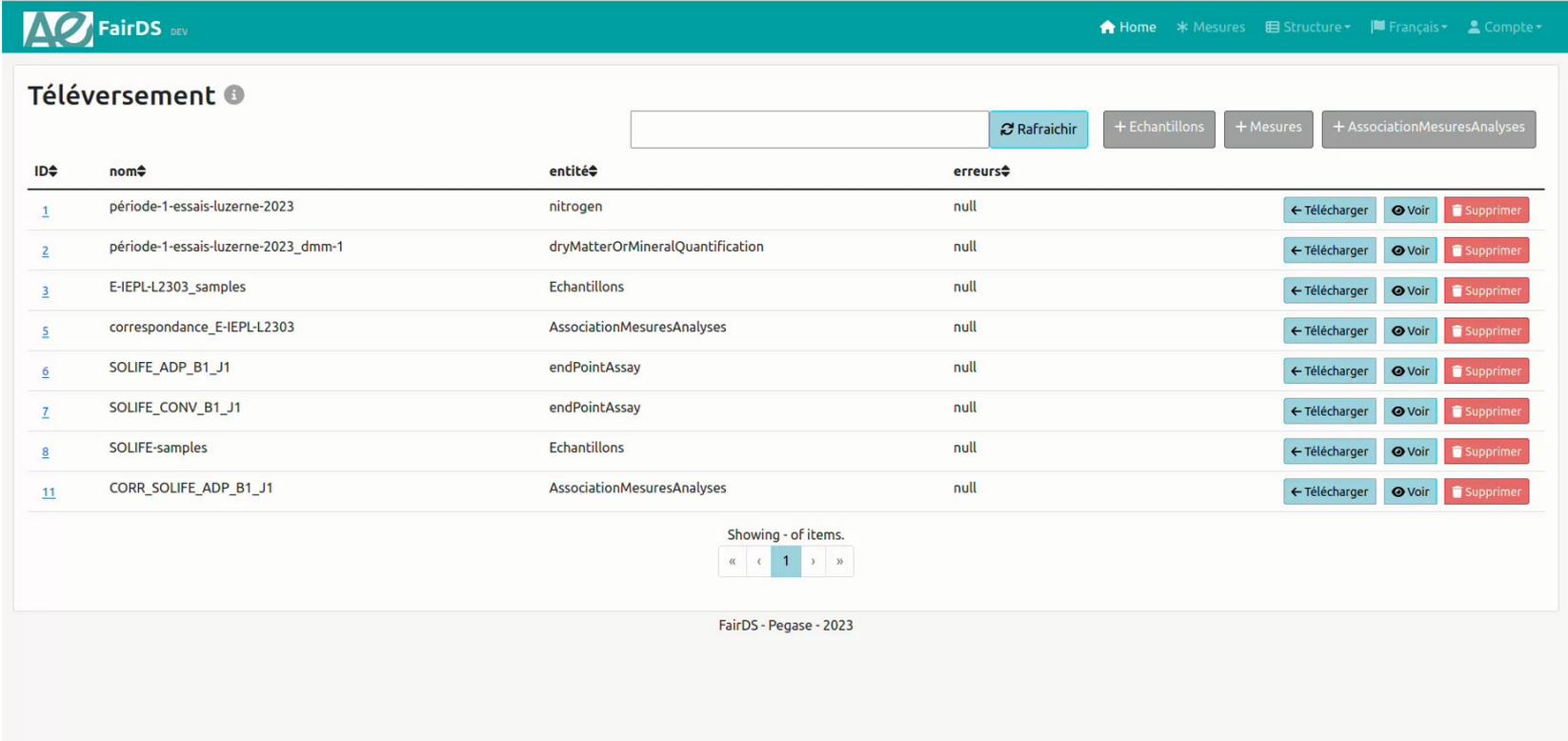
Interopérable : Des formats interopérables – Du dépôt des données à l'ouverture

- ◆ Téléversement de fichiers types complétés (Excel, puis csv, libre office)
- ◆ Téléchargement des fichiers téléversés (csv, puis Excel, libre office)
- ◆ Téléchargement de mesures (filtrées)



Résultat et présentation de l'outil

Interopérable : Des formats interopérables – Du dépôt des données à l'ouverture



The screenshot displays the FairDS web application interface. At the top, there is a navigation bar with the FairDS logo and the text "FairDS DEV". On the right side of the navigation bar, there are links for "Home", "Mesures", "Structure", "Français", and "Compte".

The main content area is titled "Téléversement" and features a search bar and several buttons: "Rafrâchir", "+ Echantillons", "+ Mesures", and "+ AssociationMesuresAnalyses".

Below the navigation bar is a table with the following columns: "ID", "nom", "entité", and "erreurs". The table contains 8 rows of data, each with a set of action buttons: "Télécharger", "Voir", and "Supprimer".

ID	nom	entité	erreurs			
1	période-1-essais-luzerne-2023	nitrogen	null	Télécharger	Voir	Supprimer
2	période-1-essais-luzerne-2023_dmm-1	dryMatterOrMineralQuantification	null	Télécharger	Voir	Supprimer
3	E-IEPL-L2303_samples	Echantillons	null	Télécharger	Voir	Supprimer
5	correspondance_E-IEPL-L2303	AssociationMesuresAnalyses	null	Télécharger	Voir	Supprimer
6	SOLIFE_ADP_B1_J1	endPointAssay	null	Télécharger	Voir	Supprimer
7	SOLIFE_CONV_B1_J1	endPointAssay	null	Télécharger	Voir	Supprimer
8	SOLIFE-samples	Echantillons	null	Télécharger	Voir	Supprimer
11	CORR_SOLIFE_ADP_B1_J1	AssociationMesuresAnalyses	null	Télécharger	Voir	Supprimer

Below the table, there is a pagination control showing "Showing - of items." and a set of navigation buttons: "«", "<", "1", ">", "»".

At the bottom of the page, there is a footer that reads "FairDS - Pegase - 2023".



Résultat et présentation de l'outil

Findable : Des données identifiables et 'trouvables'

- ◆ Une vue par type de mesure
- ◆ Un lien URI par enregistrement
- ◆ Une recherche simple possible par vue
- ◆ Par la suite, un moteur de recherche avancé avec indexation du contenu



Résultat et présentation de l'outil

Findable : Des données identifiables et 'trouvables'

AE FairDS DEV Home Mesures Structure Français Compte

EndPointAssay i

Rafraichir Télécharger + EndPointAssay

ID	Jeu de données	MeasuredMetabolite	ReagentBlankValue	SampleBlankValue	Result	ResultUnit	Wavelength	LinearityRange	DeterminationCoefficientR2	SlopeRegressionLine	SamplePortion	SamplePortion
14	SOLIFE_ADP_B1_J1	creatinine			458.704	mg/l		0-2000	0.9998	656.5	18	µl
15	SOLIFE_ADP_B1_J1	creatinine			456.707	mg/l		0-2000	0.9998	656.5	18	µl
16	SOLIFE_ADP_B1_J1	creatinine			678.218	mg/l		0-2000	0.9998	656.5	18	µl
17	SOLIFE_ADP_B1_J1	creatinine			680.007	mg/l		0-2000	0.9998	656.5	18	µl
18	SOLIFE_ADP_B1_J1	creatinine			1657.49	mg/l		0-2000	0.9998	656.5	18	µl



Résultat et présentation de l'outil

Reusable : Recherche, réplication et interprétation des données

- ◆ Des métadonnées de projet de recherche
 - ◆ Projet → Investigation → Etude → Essai
- ◆ Des métadonnées de mesure
 - ◆ Longueur d'onde utilisée (biochimie)
- ◆ Des métadonnées d'échantillon
 - ◆ Aliment, Tissu prélevé - animal
 - ◆ Liens potentiels avec barcode
 - ◆ Liens vers des fichiers lourds
- ◆ Sémantisation:
 - ◆ Introduction de vocabulaire contrôlé



Résultat et présentation de l'outil

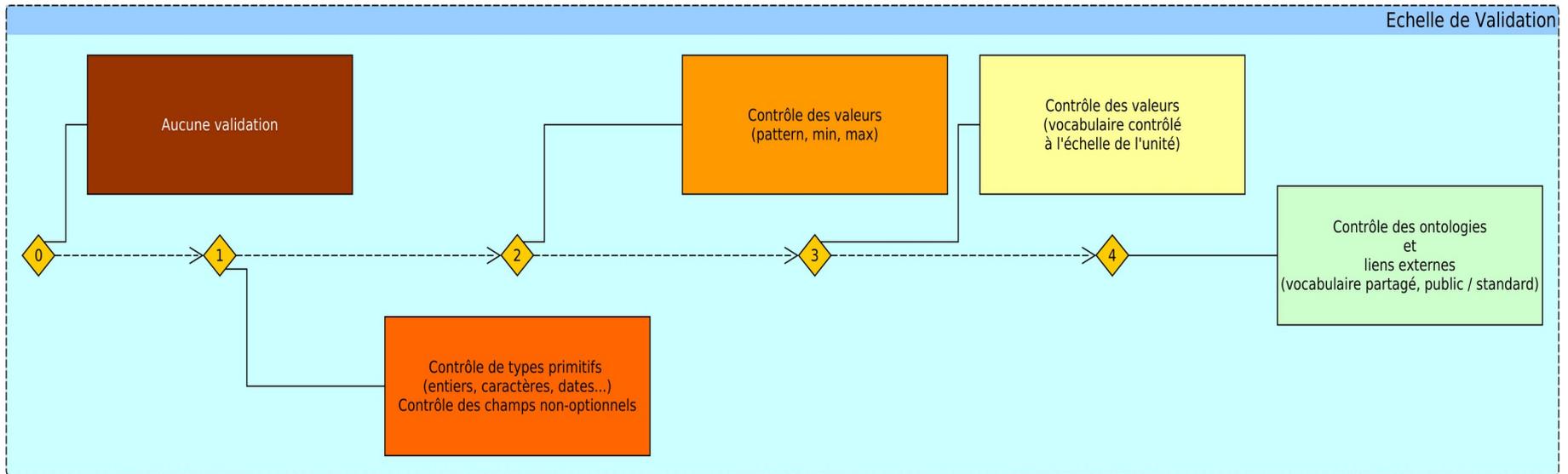
Reusable : Cohérence des données

- ◆ **Objectif**
 - ◆ Versionner les modifications des données
- ◆ **Comment ?**
 - ◆ Permettre des retours en arrière, à une date donnée
 - ◆ Globalement sur la base de données
 - ◆ Sur une recherche
- ◆ **Pourquoi ?**
 - ◆ Gestion des erreurs
 - ◆ L'utilisateur peut connaître la version des données analysées
 - ◆ Traçabilité



Résultat et présentation de l'outil

Reusable : Qualité des données - Validation



Résultat et présentation de l'outil

Ouverture des données

- ◆ Partage dans l'unité / INRAE
 - ◆ Système de permissions
- ◆ Envoi à [data.gouv.fr /dataverse](https://data.gouv.fr/dataverse)
 - ◆ Par téléchargement des données
- ◆ Partage avec le CATI SICPA



Analyse des freins et solutions

- ◆ Scientifiques

- ◆ Ouverture et partage des données dont ils sont à l'origine

- Selon leurs choix, leurs besoins, les directives

- ◆ Techniciens

- ◆ Intérêt pour l'outil

- Traçage des mesures et des échantillons

- Traçage des logiciels, appareils, réactifs

- ◆ Prise en main coûteuse (temps, habitudes)

- Nécessité d'avoir un outil plus bénéfique que couteux, globalement

- Système de validation progressif

- Nous seront présents pour la prise en main



FAIR-DS

- ◆ Gestion des données de mesure
- ◆ Type de données re-définissable
- ◆ Outil de suivi qualité
- ◆ Outil pour la FAIRisation des données
 - ◆ Au long du cycle de vie de la donnée
- F ++++
- A ++++
- I ++++
- R ++++

