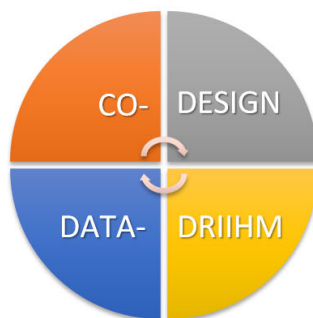


Co-design d'un portail "aiguilleur de la science ouverte" DATA-DRIIHM pour le partage des données interdisciplinaires des Observatoires Hommes-Milieus

Émilie Lerigoleur, Pascal Dayre, Mathieu Massaviol, Kristell Michel, Corinne Pardo, Jean-Claude Raynal, Marie-Laure Trémélo



Webinaire MATE-SHS Maté La Science Ouverte – 23 septembre 2021
La dynamique de l'ouverture des données : panorama d'outils



Le LabEx DRIIHM et ses 13 Observatoires Hommes-Milieus (OHM)



- Dispositif CNRS-INEE depuis 2007, financé par le LabEx depuis 2012
 - 13 observatoires (OHM)
 - ~ 1250 pers. / 200 labos / 133 EPST-univ.-grandes écoles / 27 partenaires territoriaux-associatifs
 - 866 projets annuels/thèses/post-doc
 - 73 champs disciplinaires
 - données hétérogènes "longue traîne"
 - Personnel : 1 Resp. ST + 1 IR gouv. + 1 IE (CDD) informaticien + groupe data-DRIIHM (~15 pers.)

- Une infrastructure de données construite brique par brique géocatalogues, web SIG, photothèque



Patagonia-Bahia exploradores (Chili) Pardo 2019
 E. Lerigoleur, P. Dayre, et al. - 23/09/21

Fonctionnement d'un observatoire OHM interdisciplinaire

Chaque OHM s'organise autour d'un **cadre socio-écologique** qu'un **événement fondateur** est venu bouleverser profondément

Exemples :

OHM	Cadre socio-écologique	Événement fondateur
Bassin minier de Provence	bassin minier	arrêt exploitation
Téssékéré (Sénégal)	désertification Sahel	Grande Muraille Verte
Littoral Caraïbe	contexte insulaire des petites antilles	accélération dynamiques urbanoportuaires
Littoral méditerranéen	pressions anthropiques	Gestion intégrée des zones côtières (GIZC)
Vallée du Rhône	fleuve Rhône	crues et Plan Rhône
Pyrénées Haut Vicdessos	vallée pyrénéenne	. fermeture usine . abandon pastoralisme

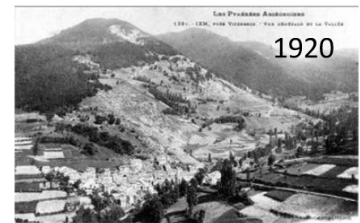


photo.driihm.fr

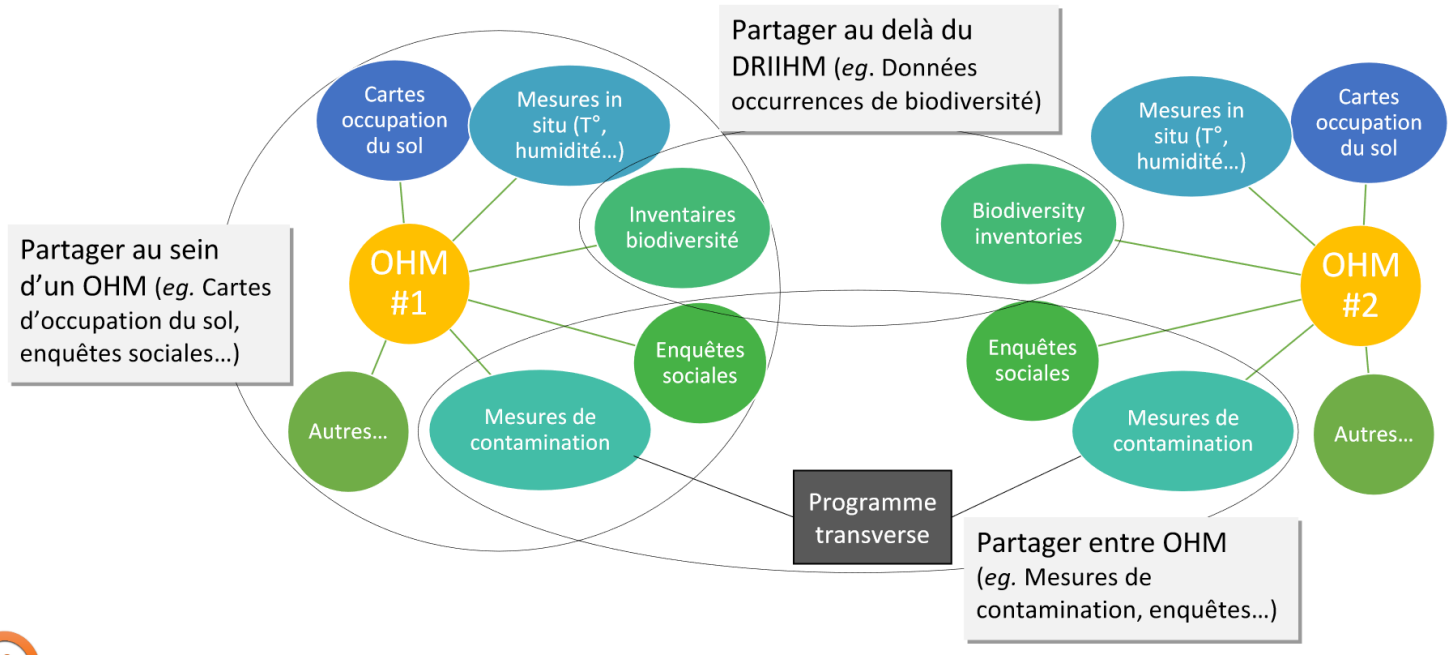


En savoir plus : driihm.fr

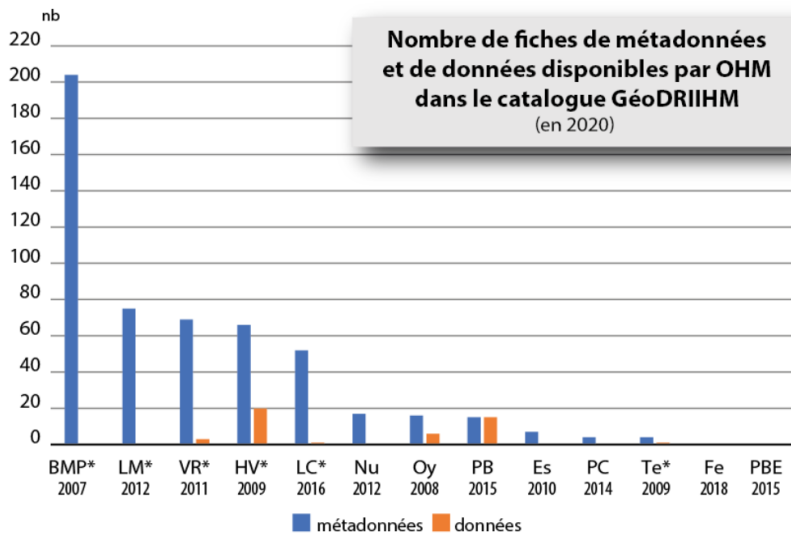


E. Lerigoleur, P. Dayre, et al. - 23/09/21

Un nécessaire partage des données au sein des OHM, entre les OHM et au delà



Des métadonnées disponibles, mais peu de données accessibles. Des outils de catalogage disponibles mais peu utilisés



BMP : Bassin minier de Provence; LM : Littoral méditerranéen; VR : Vallée du Rhône; HV : Haut-Videssos;
LC : Littoral Caraïbe; Nu : Nunavik; Oy : Oyapock; PB : Pays de Bitche; Es : Estarreja; PC : Pima County
Te : Tèssékéré; Fe : Fessenheim; PBE : Patagonia-Bahia Exploradores

* OHM doté d'au moins un ingénieur support depuis 3 ans et plus
20xx : année de création de l'OHM

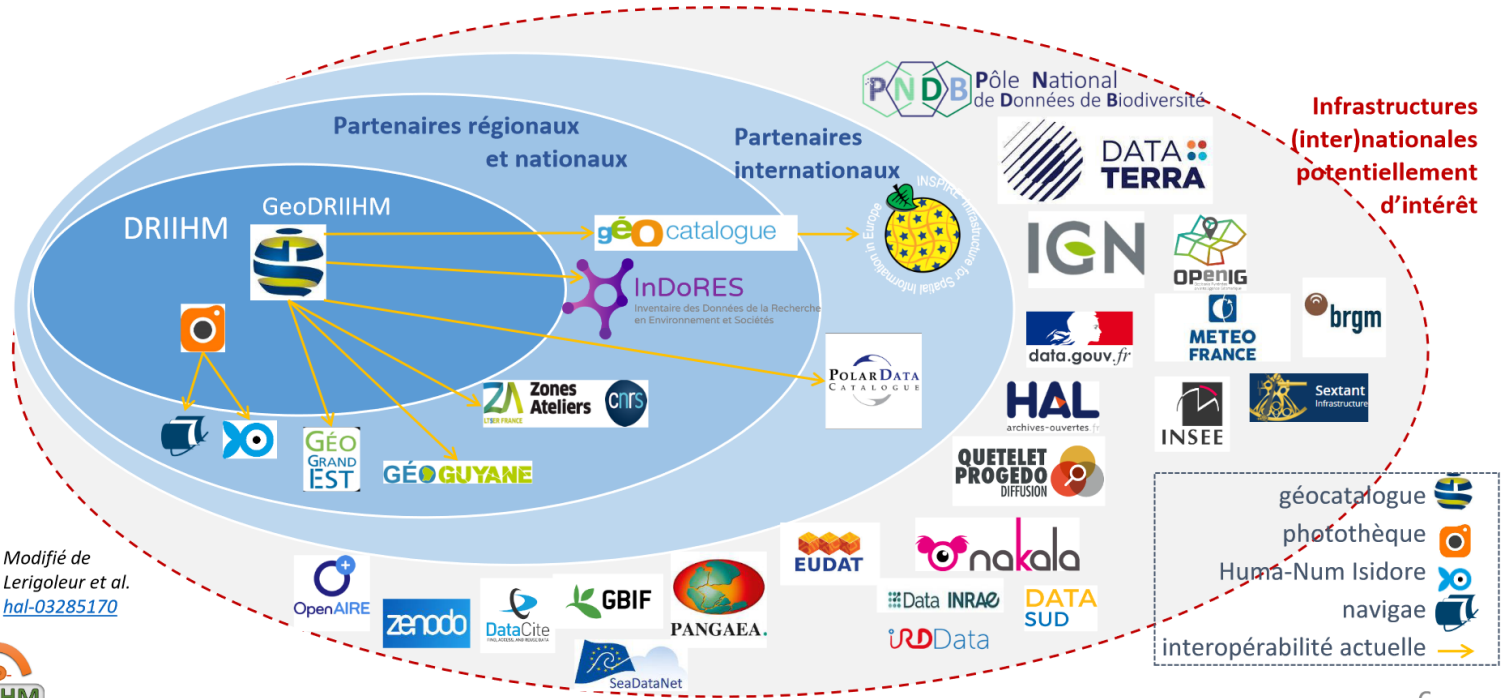
Outils de catalogage existants mais quelques limites

- Saisie des métadonnées coûteuse et peu ergonomique (logiciel Geonetwork)
- Pas de bénéfice direct dans la démarche
- Difficile d'évaluer l'interopérabilité (I) et la réutilisation (R) des données cataloguées avec cet outil



E. Lerigoleur, P. Dayre, et al. - 23/09/21

Interopérabilité complexe : comment connecter les données des OHM avec les infrastructures de données disciplinaires/institutionnelles ?

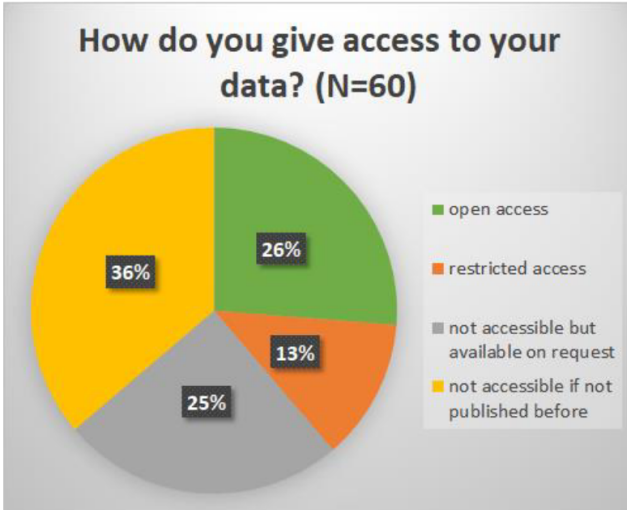


Modifié de Lerigoleur et al. [hal-03285170](https://hal.archives-ouvertes.fr/hal-03285170)

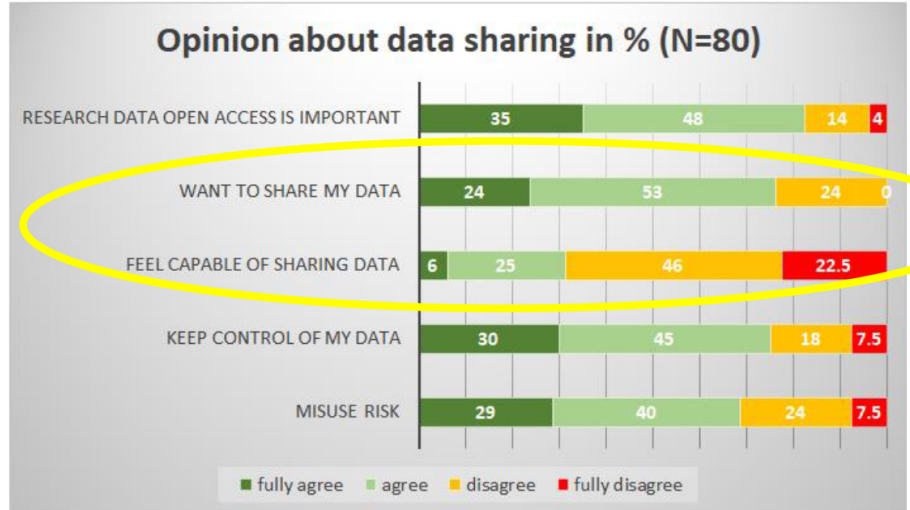


E. Lerigoleur, P. Dayre, et al. - 23/09/21

Un pas vers le partage et l'ouverture des données des OHM, mais nécessité d'accompagnement et prise en compte des potentiels risques associés



Jeu de données source en open access: data-DRIIHM 2018 survey
[DOI:10.34847/nkl.3f777nrt](https://doi.org/10.34847/nkl.3f777nrt)



Fanny Arnaud, Emilie Lerigoleur, Arnaud Jean-Charles, Iwan Le Berre, Corinne Pardo, et al. 2020, [10.18713/JIMIS-120620-6-3](https://doi.org/10.18713/JIMIS-120620-6-3) (hal-02895247)

En savoir plus : cf. communication orale lors des JA de Maté-SHS 2021
[\(support PDF, vidéo\)](#)



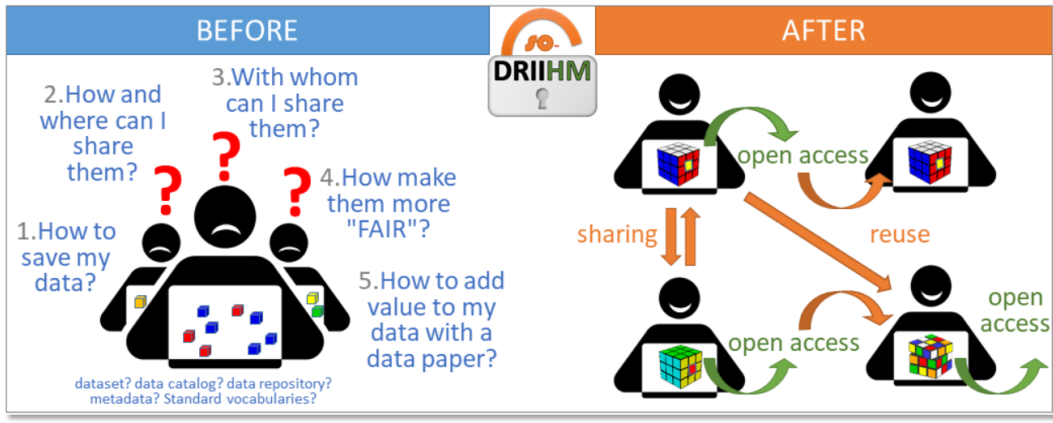
E. Lerigoleur, P. Dayre, et al. - 23/09/21

Le projet ANR SO-DRIIHM – objectifs

1. Promotion de la science ouverte : montrer les bénéfices du partage et de l'ouverture des données à la communauté scientifique DRIIHM

2. Co-construction d'une e-infrastructure ergonomique et interopérable *open source*, intégrant les outils existants et accompagnant les chercheurs dans la "FAIRisation" de leurs données

- ✓ ANR Open science Flash 2019
- ✓ 8 partenaires
- ✓ 5 OHM + 1 OHMi
- ✓ 98 k€
- ✓ Octobre 2020 - Mars 2023

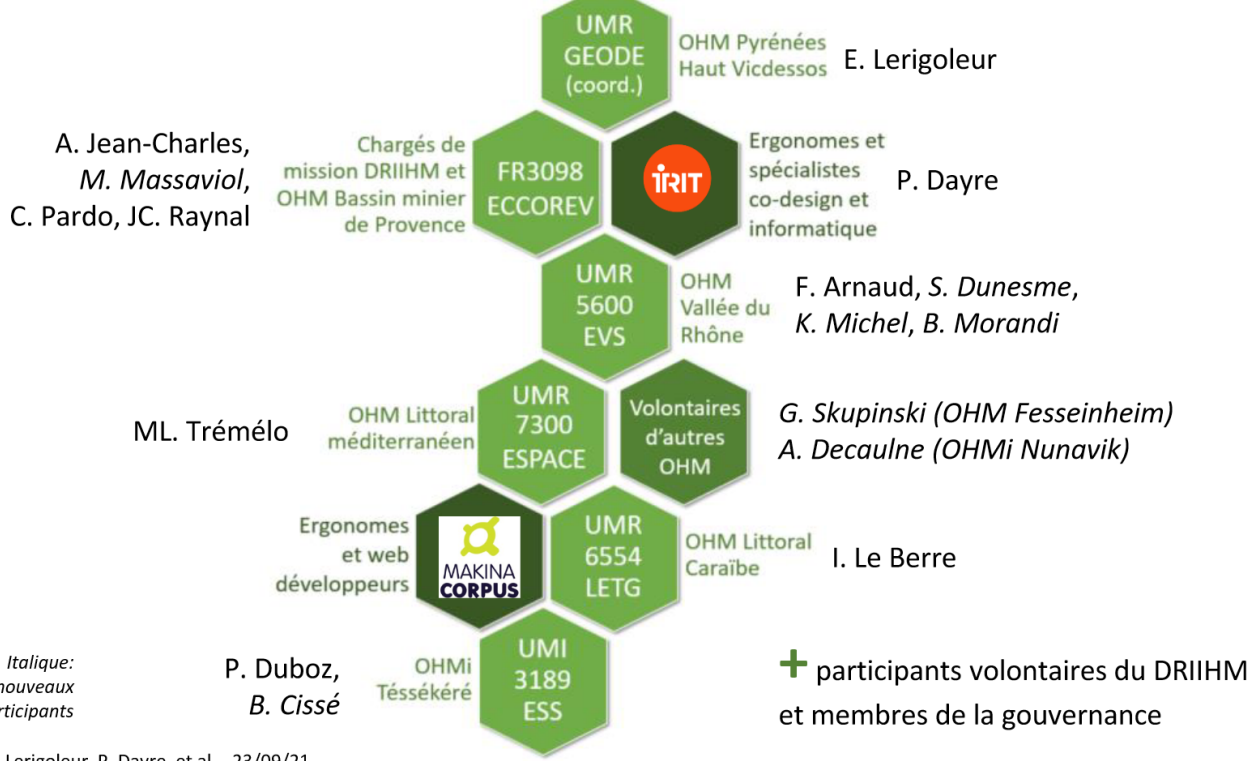


"As open as possible, as closed as necessary"
(H2020 EU Commission)



E. Lerigoleur, P. Dayre, et al. - 23/09/21

Le projet ANR SO-DRIIHM – 8 partenaires

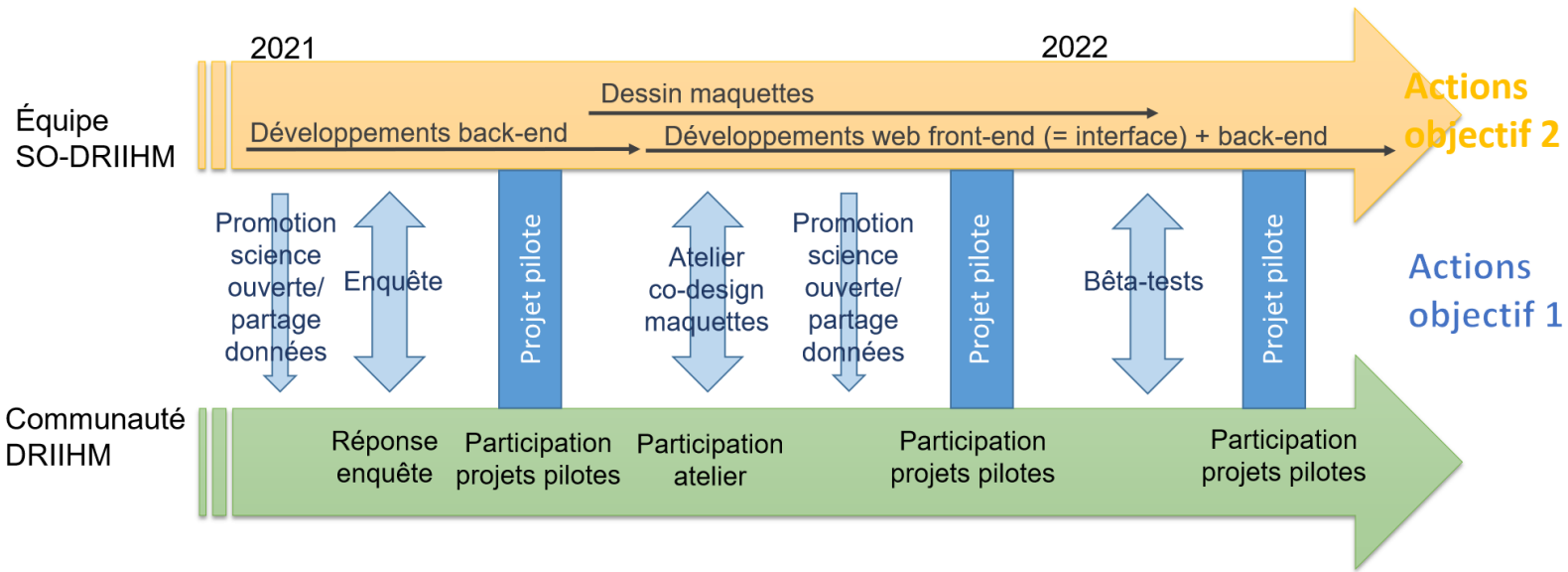


Italique: nouveaux participants

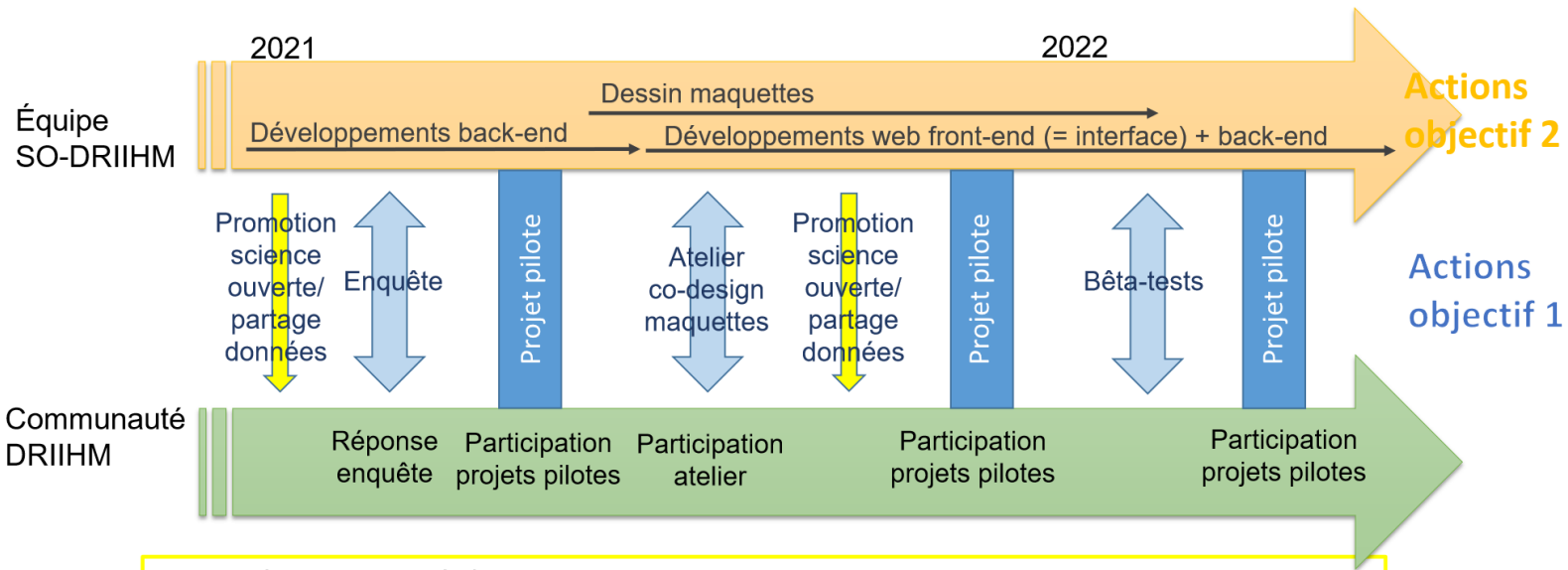


E. Lerigoleur, P. Dayre, et al. - 23/09/21

Co-construction de DATA-DRIIHM pour être le plus en phase avec les besoins de la communauté tout en suivant la feuille de route science ouverte



Co-construction de DATA-DRIIHM pour être le plus en phase avec les besoins de la communauté tout en suivant la feuille de route science ouverte

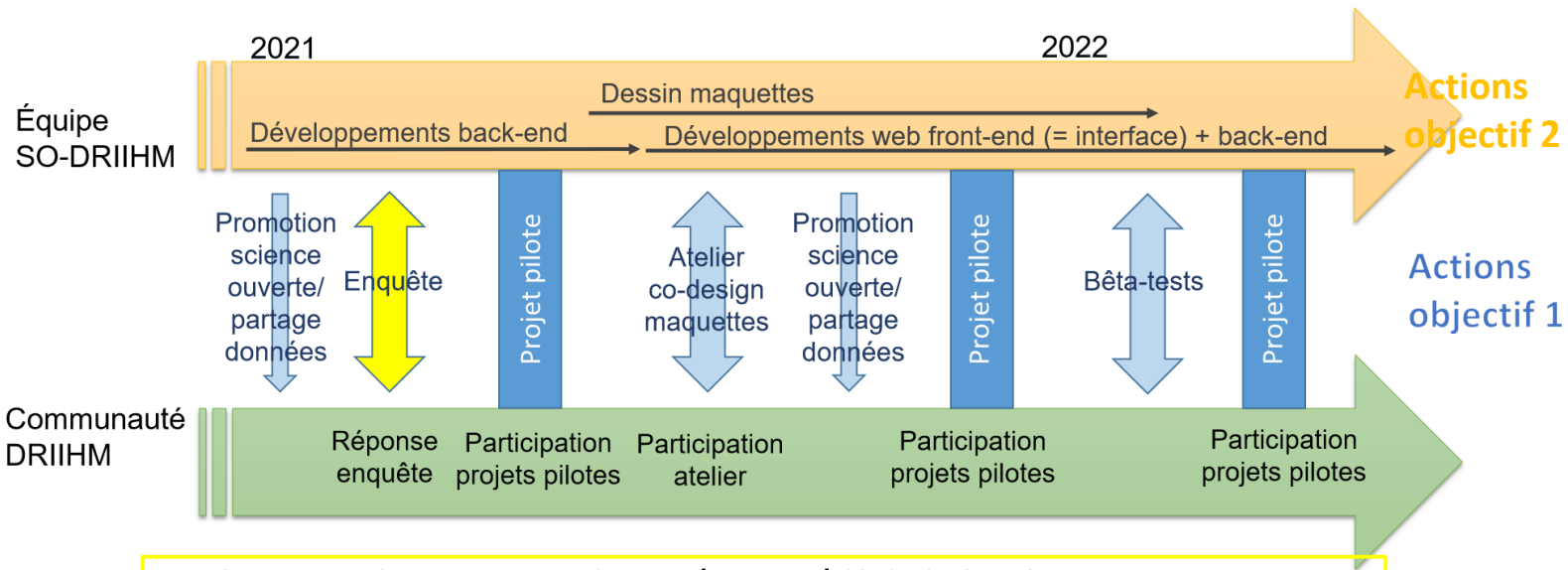


Actions de promotion de la science ouverte :

- ✓ nombreuses interventions dans webinaires/séminaires du LabEx et en dehors
 - ✓ animation d'un « café du DRIIHM » (24/03/21) sur les Plans de Gestion de Données et le RGPD
- Toutes les productions sont déposées dans la [collection HAL SO-DRIIHM](#)



Co-construction de DATA-DRIIHM pour être le plus en phase avec les besoins de la communauté tout en suivant la feuille de route science ouverte

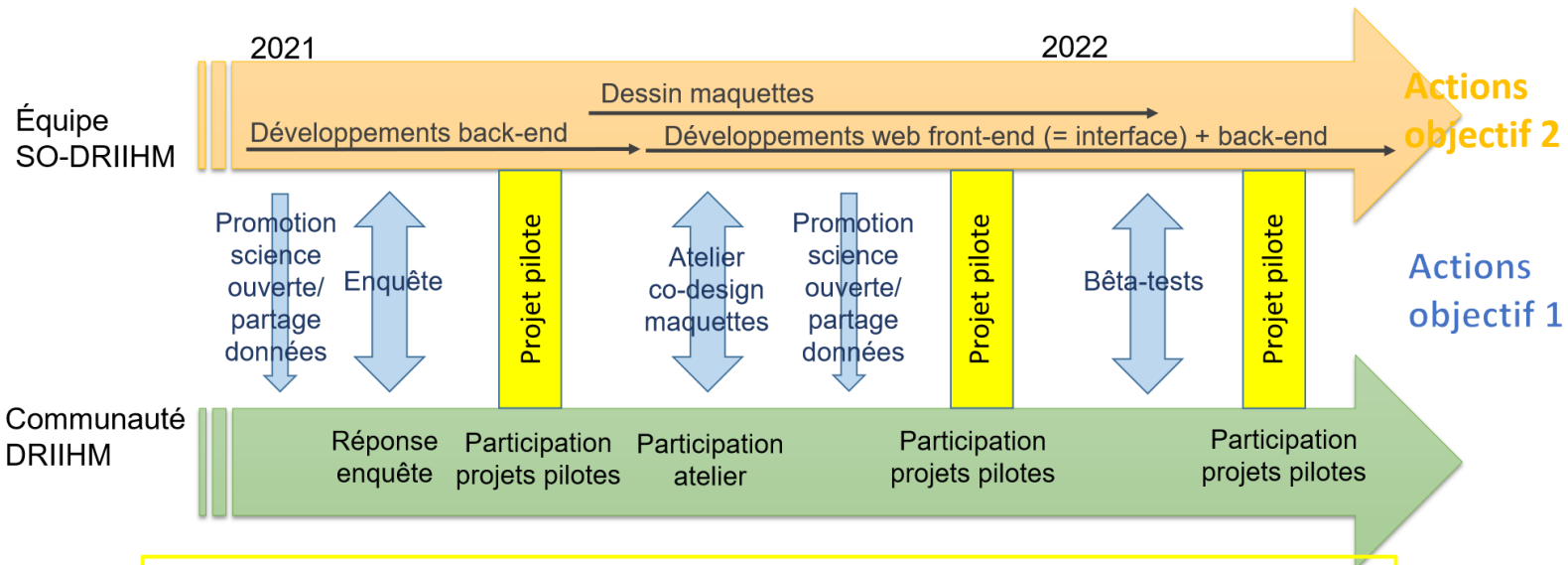


Enquête 2021 sur les pratiques : utilisation / partage / dépôt de données

- ✓ 107 répondants de 13 OHMs (~10%)
- ✓ Chercheurs et enseignants-chercheurs essentiellement (68%)
- ✓ Sciences naturelles (53%) et SHS (27%)



Co-construction de DATA-DRIIHM pour être le plus en phase avec les besoins de la communauté tout en suivant la feuille de route science ouverte

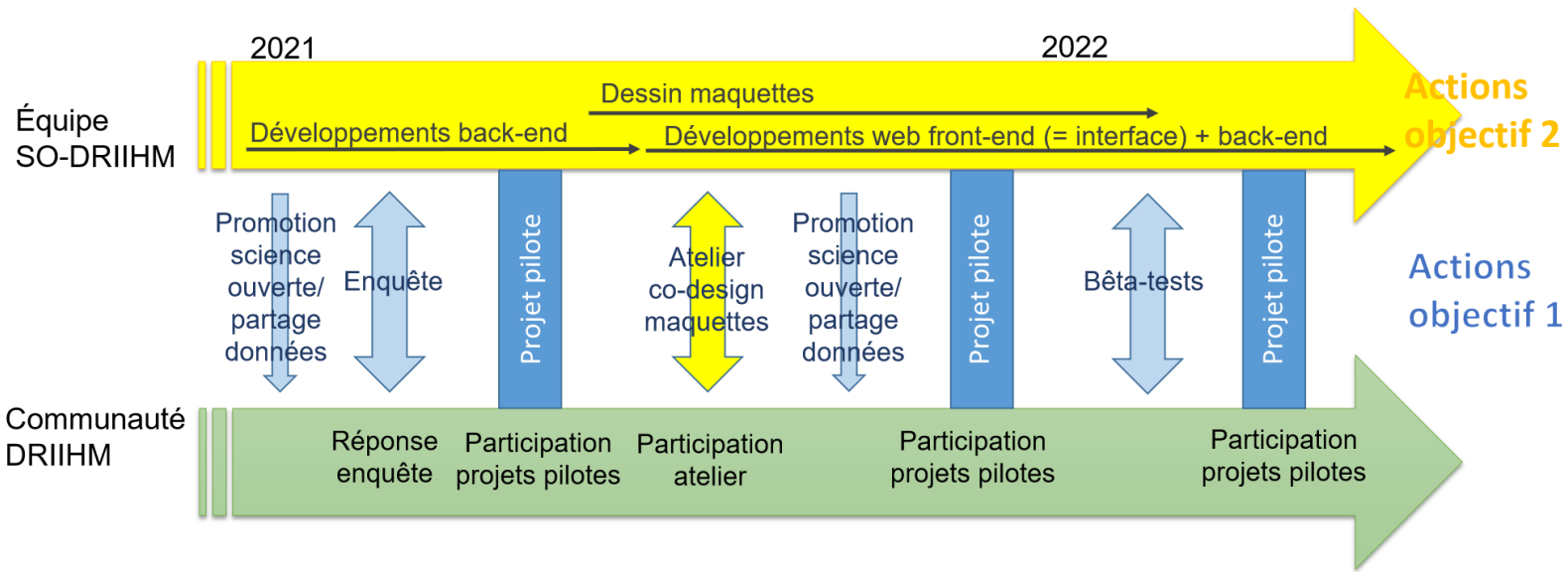


Accompagnement de 7 projets pilotes

- ✓ Projet Trafic : PGD, dépôt dans PNDB [DOI:10.48502/8bb5-pk85](https://doi.org/10.48502/8bb5-pk85), data paper [DOI:10.3897/BDJ.9.e69022](https://doi.org/10.3897/BDJ.9.e69022)
- ✓ Autres projets en cours de FAIRisation de données



Co-construction de DATA-DRIIHM pour être le plus en phase avec les besoins de la communauté tout en suivant la feuille de route science ouverte



→ Tous les échanges avec la communauté (projets pilotes, ateliers...) et les réponses à la dernière enquête nourrissent les réflexions et la conception des maquettes ...



Imaginer la future plateforme DATA-DRIIHM

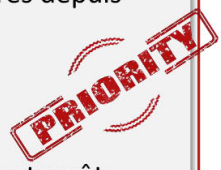
Carte d'identité



- **Nom :** DATA-DRIIHM
- **Langue :** anglais
- **Hébergeur web :** [CC-IN2P3](#) (CNRS)
- **Technologies :**
 - Django (python)
 - vue.js (javascript)
 - web sémantique (SparQL, RDF/RDFS, DCAT)
- **2 niveaux d'accès :**
 - Open access (site public)
 - Accès restreint à la communauté DRIIHM
- **Intégrée dans le site [driihm.fr](#) et couplée aux outils existants**

➔ **Développements en cours à l'IRIT –
Collaboration avec réseau alliance académique
[dataNooS](#)**

Un portail **aguilleur** de la science ouverte



- **Recherche de données** des OHM et autres depuis plusieurs sources/référentiels
- **Recherche de publications des OHM**
- **Aide au dépôt de données :**
 - Tableau de bord des dépôts multi-entrepôts
 - Liste d'entrepôts recommandés
 - Licences conseillées
 - Tags DRIIHM et OHM pour meilleure traçabilité
- **Bonnes pratiques pour rendre FAIR ses données**
- **Espace personnel avec tableau de bord d'activités**
- **Écrire un Plan de Gestion de Données**
- **Générer un brouillon de data paper**
- **Faciliter croisement données multidisciplinaires**



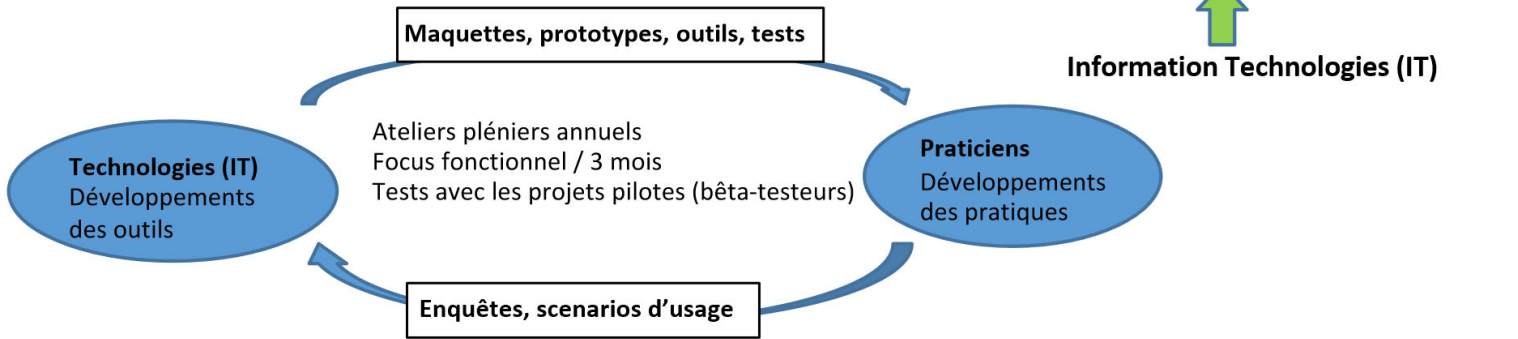
Méthodologie mise en oeuvre : le co-design

Co-design software-engineering process

« The aim of this process is to manage dynamic and integrative view on tools and practices to elaborate expectations, requirements and project management.

The design approach is the idea of co-evolution of tools, practices and actors. »

Allert et al., 2007. KP-LAB Knowledge Practices Laboratory... (hal-00593146)



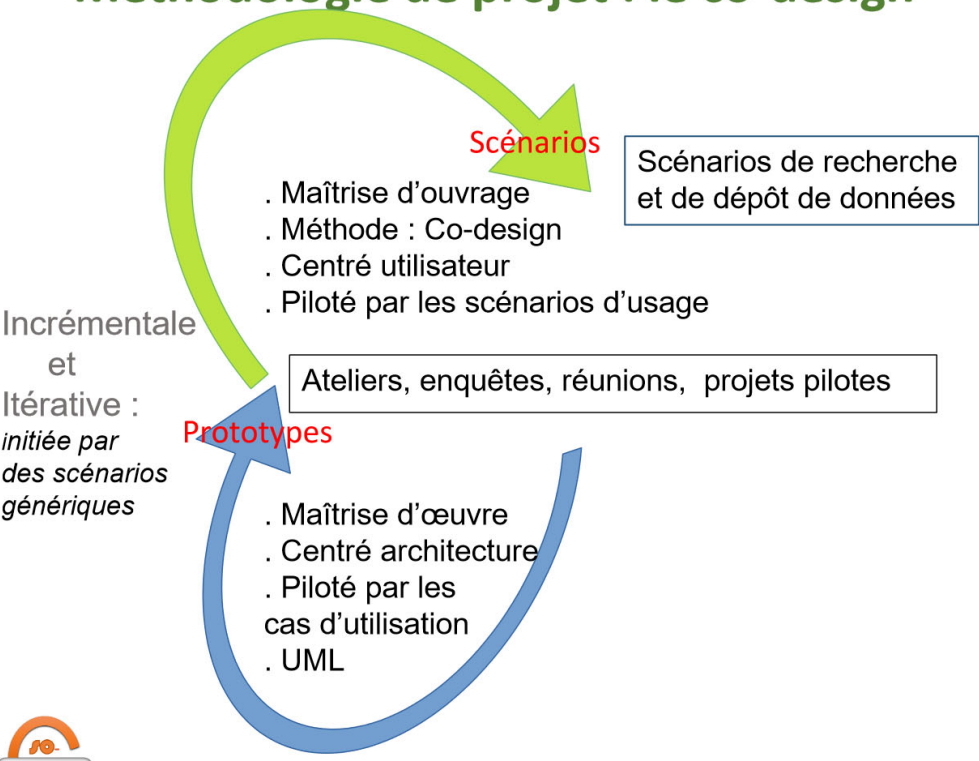
Mars-septembre 2021:

- Maquettage de l'interface utilisateur / espace personnel
- Atelier avec les utilisateurs pour affiner les scénarios d'usage de la plateforme et écrire de nouvelles spécifications
- Développement d'un prototype « Prove Of Concept » sur l'architecture technique de l'e-infrastructure
- Développement d'un démonstrateur du scénario de recherche et de construction de catalogues de jeux de données

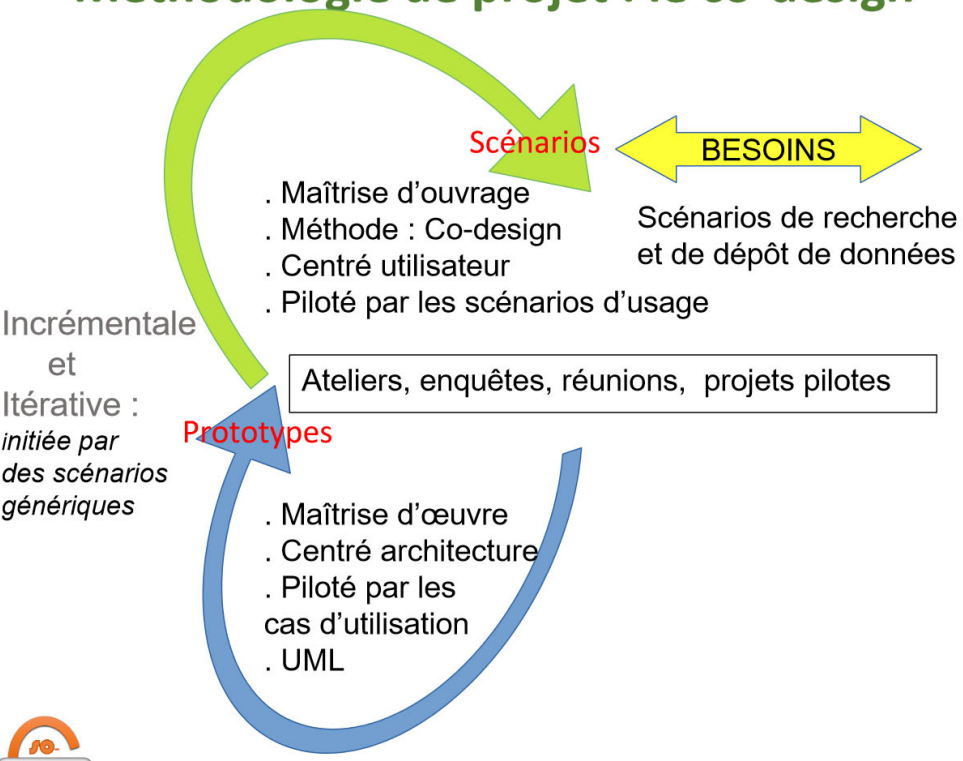


E. Lerigoleur, P. Dayre, et al. - 23/09/21

Méthodologie de projet : le co-design



Méthodologie de projet : le co-design



Communautés de producteurs/utilisateurs de données



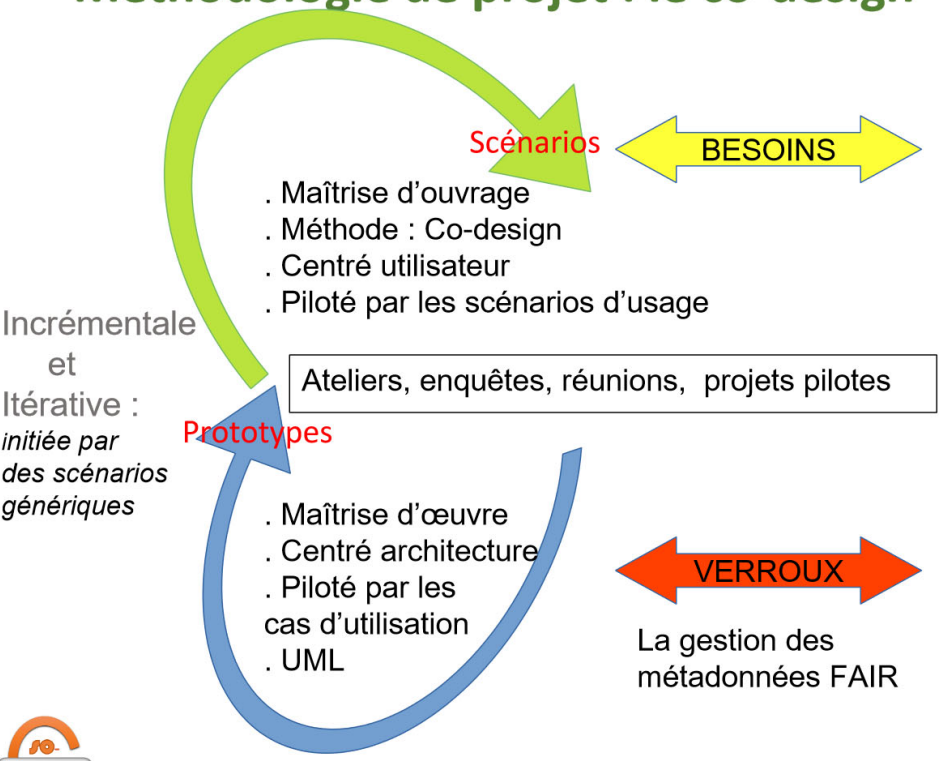
Communautés d'ingénieurs support



E. Lerigoleur, P. Dayre, et al. - 23/09/21



Méthodologie de projet : le co-design



Communautés de producteurs/utilisateurs de données



Communautés d'ingénieurs support

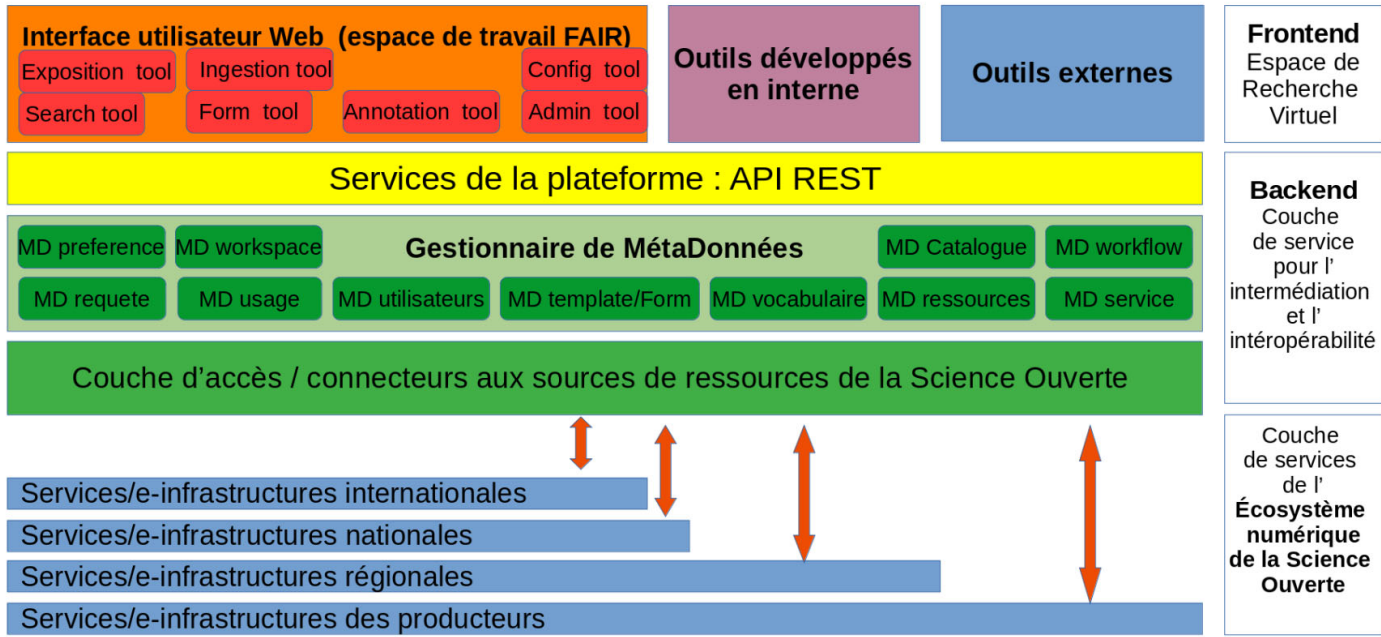
Recherche en informatique
Autres projets :

- Alliance académique sur les données DataNoos
- projet ANR Semantic4FAIR sur la FAIRisation de données météo.



E. Lerigoleur, P. Dayre, et al. - 23/09/21

Architecture de la plateforme DataNoos de médiation de la Science Ouverte

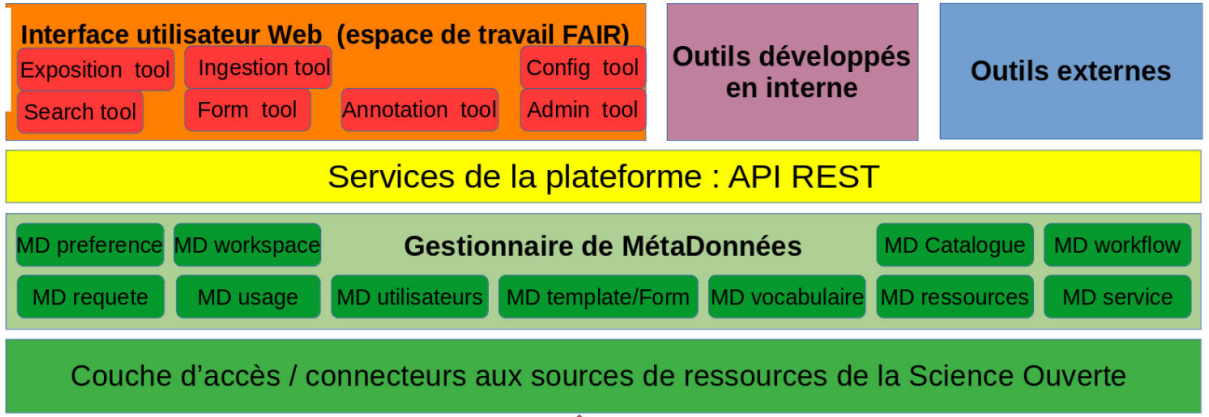


DataNoos.DataCore20210511

pascal.dayre@irit.fr

3

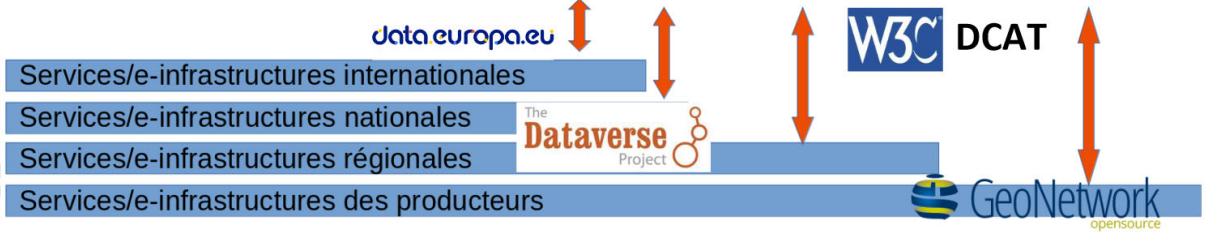
Architecture de la plateforme DataNoos de médiation de la Science Ouverte



Frontend
Espace de Recherche Virtuel

Backend
Couche de service pour l'intermédiation et l'interopérabilité

Couche de services de l'Écosystème numérique de la Science Ouverte

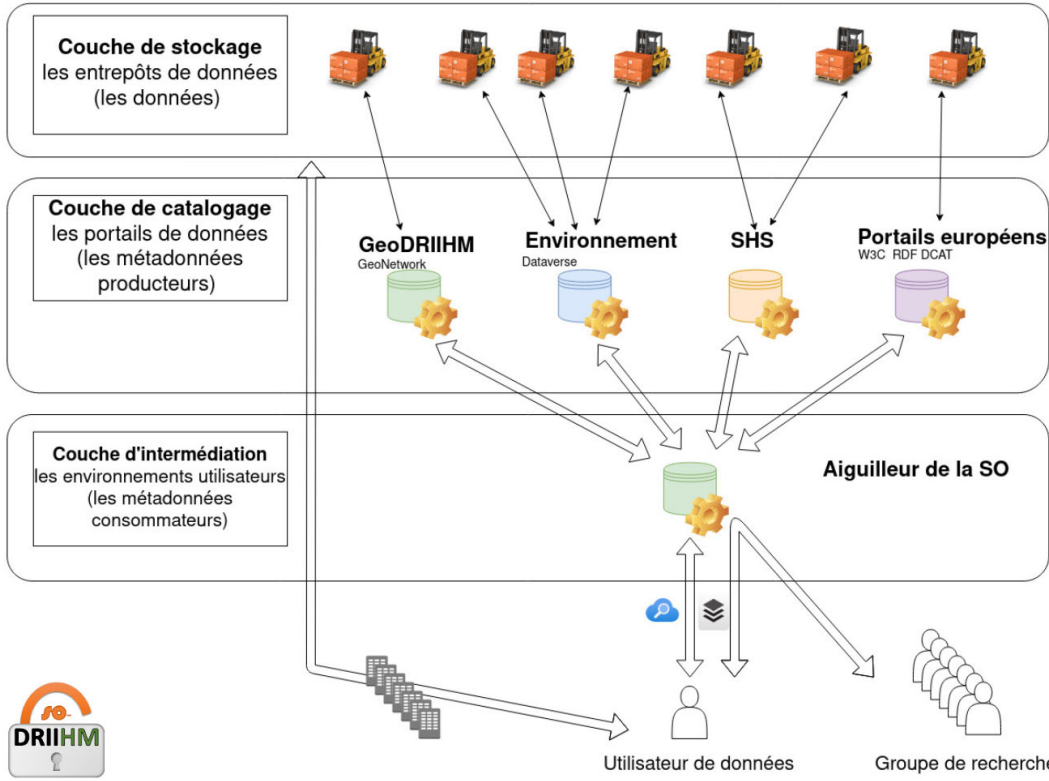


DataNoos.DataCore20210511

pascal.dayre@irit.fr

3

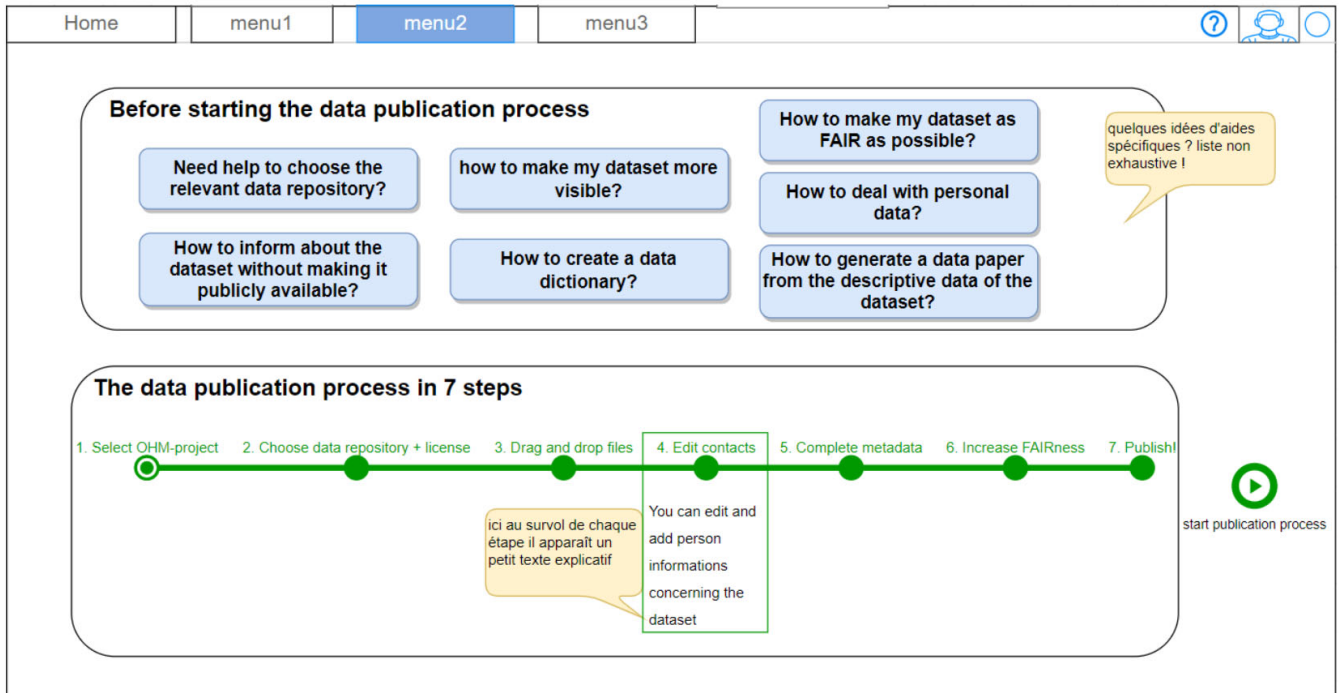
Scénario de constitution et de partage d'un catalogue de jeux de données par un ensemble de recherche sur internet via la plateforme Aiguilleur de la science ouverte



- Entrepôt de données (Data repository)**
- Jeu de données (dataset)**
- Catalogue de jeux de données (Catalog / Collection - métadonnées)**
- Portail de données**
- Service**
- Recherche de données (data search)**



Exemple : maquette du parcours utilisateur « Dépôt de données » proposée lors de l'atelier du 6 septembre 2021



Les résultats de l'atelier du 6 septembre 2021

- Validation et enrichissement des scénarios recherche et dépôt
- Ajustement, enrichissement et validation du recherche scénario de dépôt des données
- Validation du besoin « Espace de travail »
 - Tableau de bord de ses données
 - Notion de catalogue de données de travail
- Validation du besoin d'élargir la notion de données aux publications et aux autres produits numériques de la recherche (pré-prints, rapports, ...)
- Validation du besoin de rechercher des relations explicites voire implicites entre les données et les publications



Perspectives du projet

- **Prototype V2 :**
 - Spécification et développement de la nouvelle version de la recherche de jeux de données
 - Spécification et développement du scénario de dépôt et de son tableau de bord
 - Spécification de l'Espace de travail
 - Référencement de nouveaux services de catalogage
 - Démarrage de nouveaux développements avec un CDD et le partenaire Makina Corpus
 - Nouvelle mobilisation des ingénieurs support et élargissement du cercle des bêta-testeurs
- **Les autres actions de SO-DRIIHM se poursuivent :**
 - Accompagnement des projets pilotes dans la FAIRisation des données
 - Promotion de la science ouverte et transfert de « bonnes pratiques »
 - Veille sur le partage/ouverture des données de la recherche, en lien avec les initiatives nationales (recherche.data.gouv,...) et internationales (RDA, GOFAIR...)

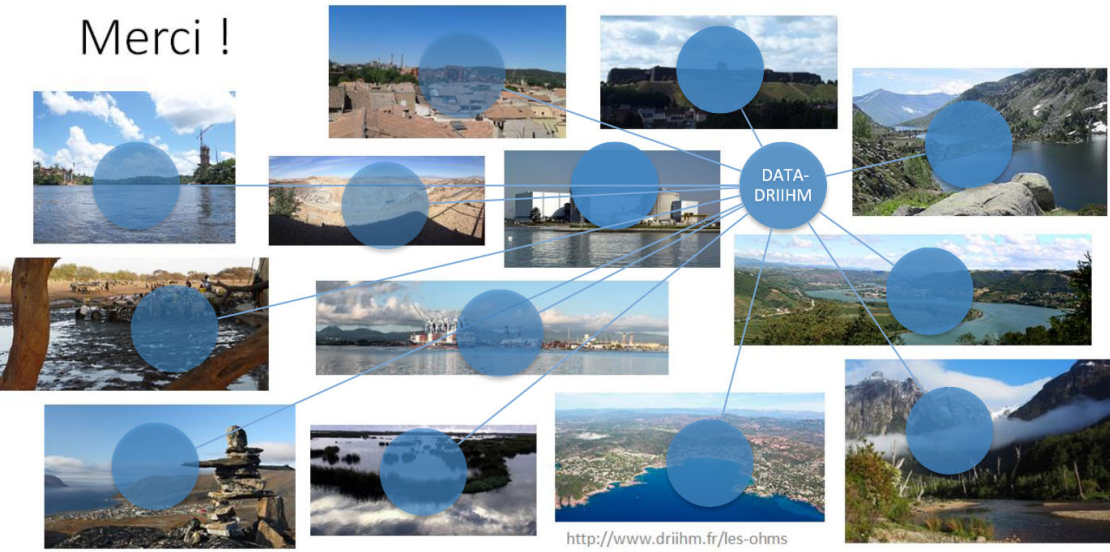


Conclusion

- La nouvelle plateforme de DATA-DRIIHM a pour but de fortifier les 2 communautés de pratiques et d'usages : les producteurs/utilisateurs des données et les ingénieurs support
- Le co-design permet d'impliquer ces 2 communautés dans l'évolution de ce futur outil de travail au cœur de leurs activités
- À la différence des méthodes agiles, ici c'est l'utilisateur final qui est impliqué
- Nous avons noté beaucoup d'interactions lors du dernier atelier
- L'implication des praticiens est un bon moyen pour sensibiliser à la science ouverte et pour faire une bonne conduite du changement
- Par cette co-construction nous complétons bien les enquêtes
- L'identification des verrous permet d'alimenter les travaux de recherche en informatique en amont : ainsi les résultats de recherche en informatique sont directement intégrés dans la solution finale
- C'est un projet *open source* qui est appelé à fédérer plus largement



Merci !



emilie.lerigoleur@univ-tlse2.fr

ORCID: [0000-0002-0864-659X](https://orcid.org/0000-0002-0864-659X)

[CV HAL](#)

pascal.dayre@irit.fr

ORCID: [0000-0002-2351-8991](https://orcid.org/0000-0002-2351-8991)

mathieu.massaviol@univ-amu.fr

ORCID: [0000-0002-1358-1567](https://orcid.org/0000-0002-1358-1567)

kristell.michel@ens-lyon.fr

[CV](#)

corinne.pardo@univ-amu.fr

ORCID: [0000-0001-5273-7117](https://orcid.org/0000-0001-5273-7117)

[CV HAL](#)

jean-claude.raynal@univ-amu.fr

marie-laure.tremelo@univ-amu.fr

ORCID: [0000-0002-4802-7179](https://orcid.org/0000-0002-4802-7179)

[CV HAL](#)

Toutes les productions de l'équipe SO-DRIIHM sont dans la [collection HAL SO-DRIIHM](#)



E. Lerigoleur, P. Dayre, et al. - 23/09/21