

- 1- Contexte
- 2- Organisation et services
- 3- Europe
- 4 - Cycle données - stockage



Plateforme des données de Recherche Data Gouv

Love Data Week – ADOC Lorraine

Richard Moreno (INRAE) - 13 mars 2025



@RechercheDataGv

1/ Contexte



Objectifs

Recherche Data Gouv

Ne pas laisser d'équipes de recherche sans solution pour ouvrir ou partager leurs données

Faire de l'accompagnement un élément central du dispositif

Reconnaître, fédérer et amplifier les initiatives existantes

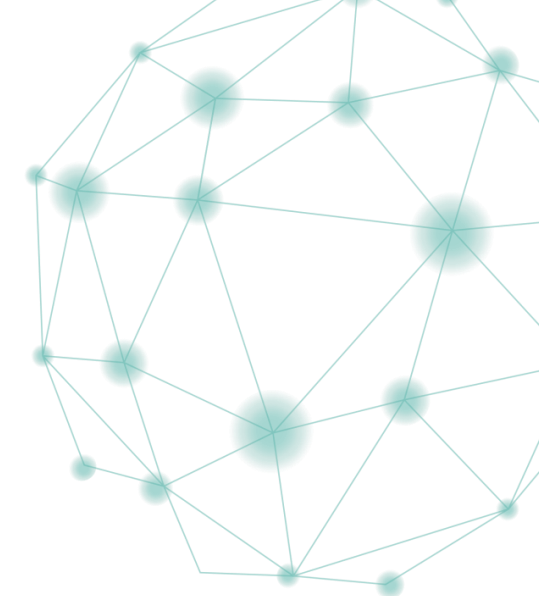
Un maillage d'offres d'accompagnement en proximité de la recherche

- Sensibiliser, former ou accompagner les équipes de recherche

Une offre souveraine de dépôt, publication et de signalement des données

Plateforme nationale pluridisciplinaire et fédérée des données de la recherche

- Complémentaire des entrepôts thématiques
- **Une alternative à la publication sur des plateformes commerciales ou ne répondant pas à des critères de confiance**
- Devenir un des services de *l'European open science cloud*





RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité

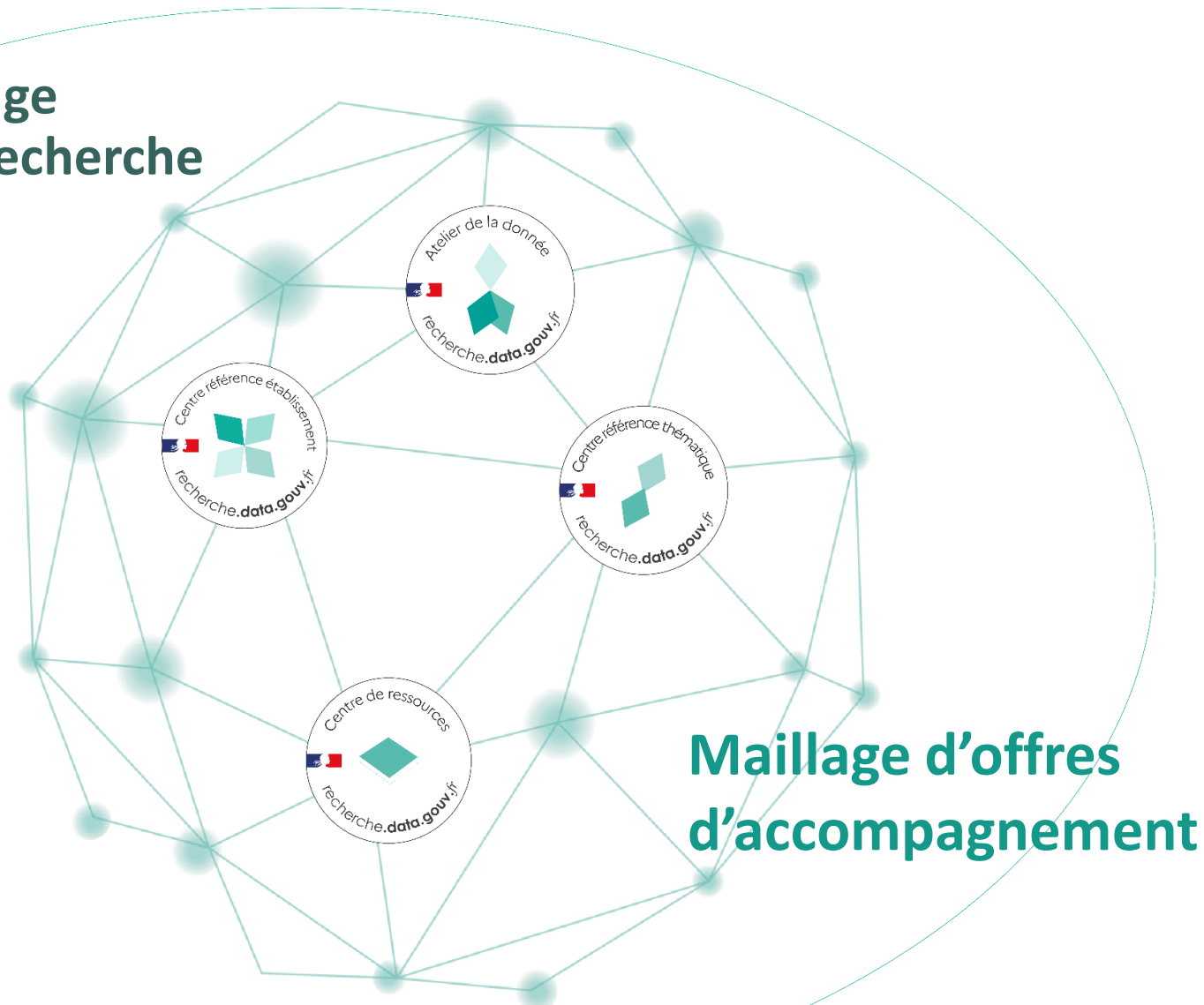
recherche.data.gouv.fr

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



Plateforme des données

Entrepôt-
catalogue



Maillage d'offres
d'accompagnement



@RechercheDataGv



Une plateforme des données mutualisée

entre les établissements d'enseignement supérieur et de la recherche

- Une offre nationale entrepôt et catalogue des données développée et maintenue par des établissements au service de toute la communauté de recherche
- Décharge des établissements du coût et du maintien d'une offre technique
- Une offre maintenue à l'état de l'art
- Masse critique pour atteindre une dimension internationale
 - Exemple : projet FIDELIS



Plateforme de données de Recherche Data Gouv

Est un des dispositifs opérationnels de Recherche Data Gouv

Elle regroupe les composants :

- Entrepôt de données
- Catalogue de données
- Outils du portail de Recherche Data Gouv
- Centre de ressources Plateforme

Est financée par le Ministère de l'Enseignement supérieur et de la Recherche, le ministère délègue la gestion du budget à INRAE



2/ Organisation et services

Une unité mixte de service Plateforme de données

Structure opérationnelle regroupe les personnels de l'équipe chargée de :

- Maintenir en conditions opérationnelles l'entrepôt, le catalogue et l'outil du portail
- Faire évoluer aux plans fonctionnels et techniques l'entrepôt et le catalogue de données
- Faire évoluer techniquement l'outil de portail sur la base des demandes validées en collège de coordination et portée par la responsable de la communication
- Coordonner les administrateurs des espaces institutionnels de donnée de l'entrepôt de données
- Former, accompagner, communiquer sur les services de la plateforme de données
- Jouer un rôle majeur dans l'écosystème national et international sur son périmètre d'action



RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité

recherche.data.gouv.fr

UMS Plateforme Nationale des données de la recherche

- Tutelles :
 - INRAE, Universités de Franche-Comté, Grenoble-Alpes, Lorraine, Strasbourg
- UMS opérationnelle depuis le 01/01/2025
- Organisation en 5 pôles d'activités
 - Portail
 - Entrepôt
 - Catalogue
 - Centre de ressources
 - Infrastructures

UNIVERSITÉ DE
FRANCHE-COMTÉ

UGA
Université
Grenoble Alpes

INRAE
la science pour la vie, l'humain, la terre

UNIVERSITÉ
DE LORRAINE

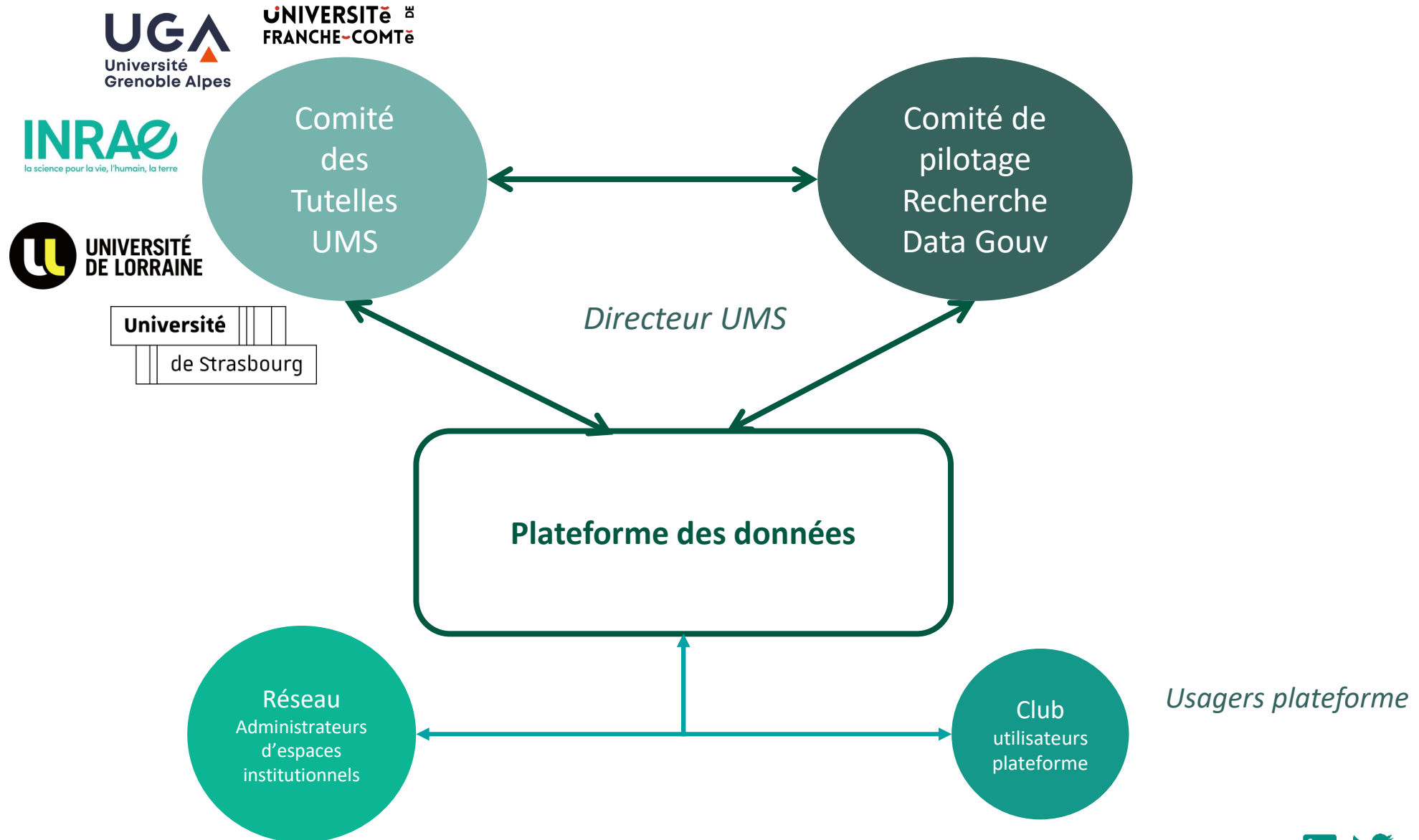
Université
de Strasbourg



@RechercheDataGv



Plateforme des données de Recherche Data Gouv





Gouvernance de l'UMS

Plateforme nationale des données de la recherche

Comité des tutelles

- Composé d'un représentant de chaque tutelle également membre du Comité de pilotage de *Recherche Data Gouv*
- Veille à la mise en cohérence des allocations des moyens de l'UMS vis-à-vis des orientations scientifiques et techniques validées par le Comité de pilotage de *Recherche Data Gouv*
- Se prononce sur le bilan financier et le budget prévisionnel
- Valide l'organisation de l'UMS proposée par son Directeur

Réseau des Administrateurs des espaces institutionnels de données

- Rassemble les administrateurs des espaces institutionnels de données
- Recense et organise les besoins et demandes communes des administrateurs
- Propose des actions dédiées
- Contribue à l'organisation des tests fonctionnels sur les nouvelles fonctionnalités de l'entrepôt et du catalogue

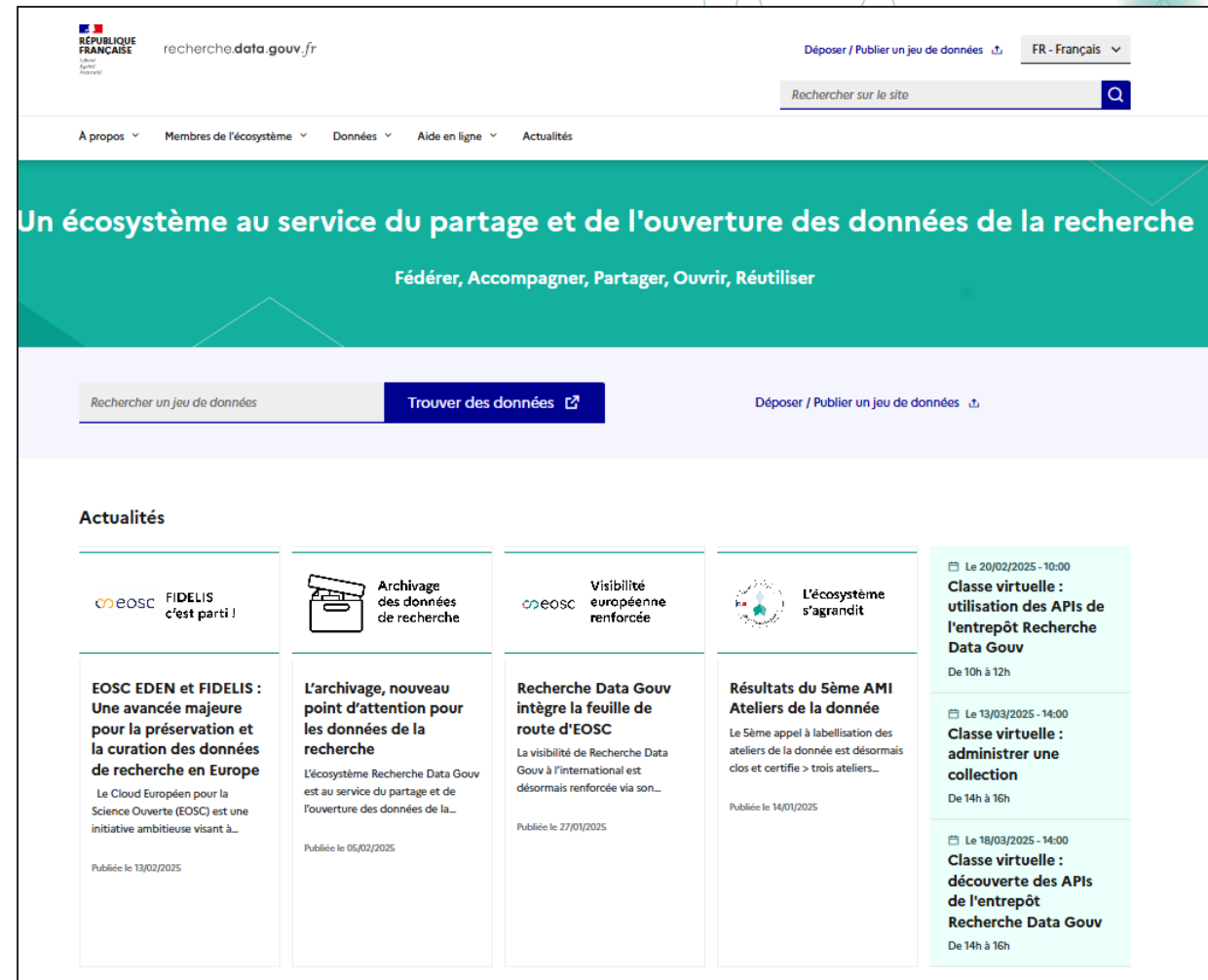
Club des utilisateurs

- Représente les différents usagers de la Plateforme : entrepôt, catalogue
- Recense et organise les besoins et demandes des utilisateurs
- Propose des suggestions d'évolution des services
- Propose des actions de communication en direction des utilisateurs



Pôle Portail

- Site WWW : <https://recherche.data.gouv.fr/fr>
- Missions
 - Maintenir le portail Recherche Data Gouv en condition opérationnelle
 - Développer et intégrer les besoins validés en collège de coordination et portée par la responsable de la communication
- Caractéristiques techniques
 - Technologie : framework PHP opensource Symfony
 - Infrastructures : cloud commercial ScaleWay qui sera migré vers des moyens cloud ESR : ADAGE (Strasbourg) et GRICAD (Grenoble)
- Activité en cours
 - Catalogue des entrepôts de confiance





Pôle Entrepôt

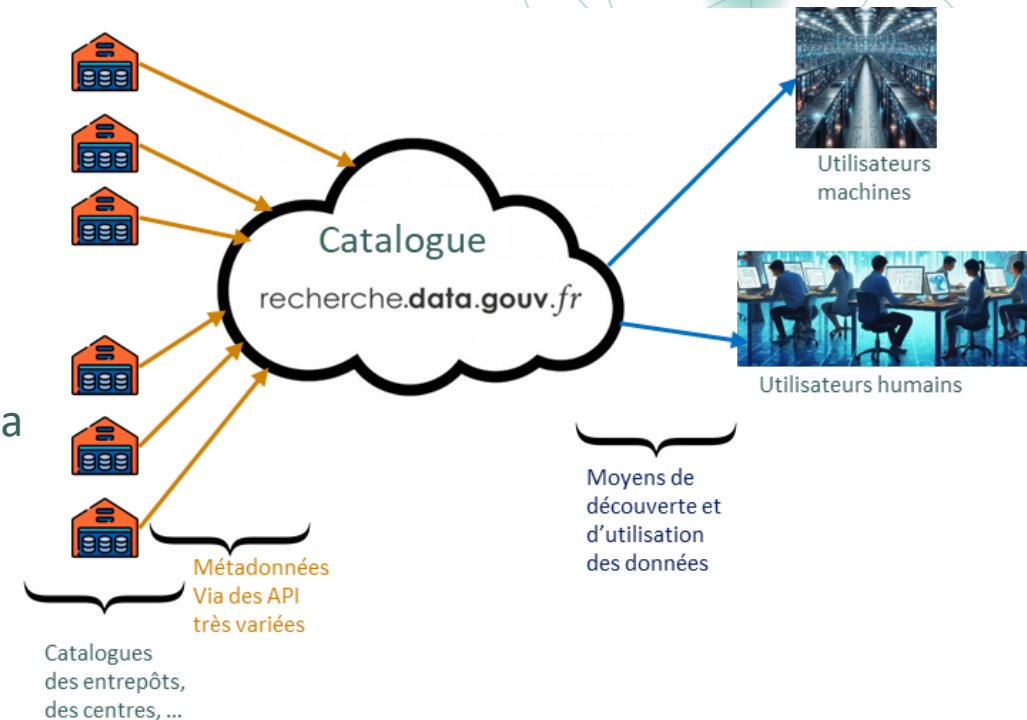
- Site : <https://entrepot.recherche.data.gouv.fr/>
- Missions
 - Maintenir l'entrepôt Recherche Data Gouv en condition opérationnel
 - Développer et intégrer les besoins exprimés par la direction, et le collège de coordination
- Caractéristiques techniques
 - Technologie : logiciel opensource Dataverse
 - Infrastructures : INRAE : Toulouse & Paris
- Quelques chiffres
 - ~ 70 espaces institutionnels
 - ~30 To de données
 - ~ 5700 jeux de données
 - ~ 1 900 000 téléchargements

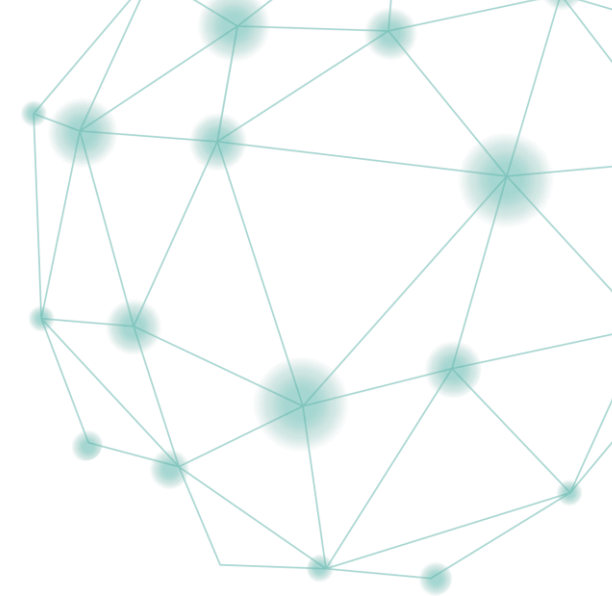
The screenshot shows the homepage of the recherche.data.gouv.fr website. The header includes the French Republic logo and the text 'Liberté Égalité Fraternité'. The main navigation bar contains links for 'Recherche', 'À propos', 'Guide d'utilisation', 'Support', 'Français', 'S'inscrire', and 'Se connecter'. Below the navigation bar, there are four featured logos: 'Espace Générique', 'Arts et Métiers Sciences et technologies - Arts et Metiers Institute of Technology', 'cirad Centre de Coopération Internationale en Recherche Agronomique pour le Développement', and 'CNRS Research Data'. A search bar is located below the logos, with the text 'Chercher dans cette collection...' and a magnifying glass icon. To the right of the search bar is a link for 'Recherche avancée'. Below the search bar, there are filters for 'Collections (1 069)', 'Jeux de données (5 509)', and 'Fichiers (97 455)'. The main content area displays search results for '1 à 10 de 6 578 résultats'. The first result is 'Data associated with the publication 'Microscopic and stochastic simulations of chemically active droplets'', dated 13 févr. 2025 - PHENIX. The second result is 'A dataset of food supplies and preferences of French university students', dated 13 févr. 2025 - ChemoSens. The third result is 'Tree Crown Data - MONTANE OSUG', dated 13 févr. 2025 - LESSEM. The website also features a sidebar with 'Catégorie de collection' (Projet de recherche (330), Laboratoire (313), Groupe de recherche (157), Organisation ou établissement (152), Chercheur (14)) and 'Source' (Recherche Data Gouv (4 580), Moissonné (1 998)).



Pôle Catalogue des données de la recherche

- Etat
 - Projet en cours de développement / prototypage
- Missions
 - Catalogue des données de la Recherche :
 - répertorier (signaler)
 - rechercher / découvrir des jeux de données
 - Développer et intégrer les besoins exprimés par la direction, la responsable communication et le collège de coordination
- Caractéristiques techniques
 - Technologie : A définir en privilégiant l'opensource
 - PoC en cours basé sur IA
 - Infrastructures : idem entrepôt





Catalogue des données de la recherche française

- Données de la recherche
 - Produite par une activité de recherche
- Données de la recherche française
 - Produites ou coproduites par au moins un auteur affilié à un établissement de recherche ou d'enseignement français ;
 - ou dans des entrepôts de données gérés par des établissements français de la recherche, sans faire de distinction sur l'affiliation des auteurs des données (français ou étrangers)
- Données hébergées dans
 - Potentiellement tout type d'entrepôt dans la mesure où elles sont signalées manuellement
 - —> Uniquement les entrepôts priorisés par les acteurs de Recherche Data Gouv concernant le référencement automatisé



Espaces institutionnels dans l'entrepôt

- Chaque établissement d'enseignement supérieur ou de recherche peut demander la création d'un espace institutionnel afin de déposer et publier ses données de recherche.
- Chaque établissement est responsable de son espace et de la curation des dépôts de jeux de données.
- La Plateforme assure la charge de maintien en condition opérationnelle de l'entrepôt => mutualisation
- Environ 70 établissements de l'ESR disposent d'un espace institutionnel
- Aujourd'hui le volume est limité à 5 To
- Evolution (stockage, curation) en cours d'analyse

Espaces institutionnels

Chaque établissement d'enseignement supérieur ou de recherche peut demandeur la création d'un espace institutionnel afin de déposer et publier ses données de recherche.

Ci-dessous les espaces déjà ouverts.

Vous n'avez pas encore d'espace institutionnel dans l'entrepôt de Recherche Data Gouv ? Un espace générique est proposé par défaut.

Aix Marseille Université (AMU)

+

Arts et Métiers Sciences et Technologies (ENSAM)

+

Avignon Université (AU)

+

Centre de coopération internationale en recherche agronomique pour le développement (CIRAD)

+

Centre national de la recherche scientifique (CNRS)

+

CY Cergy Paris Université (CYU)

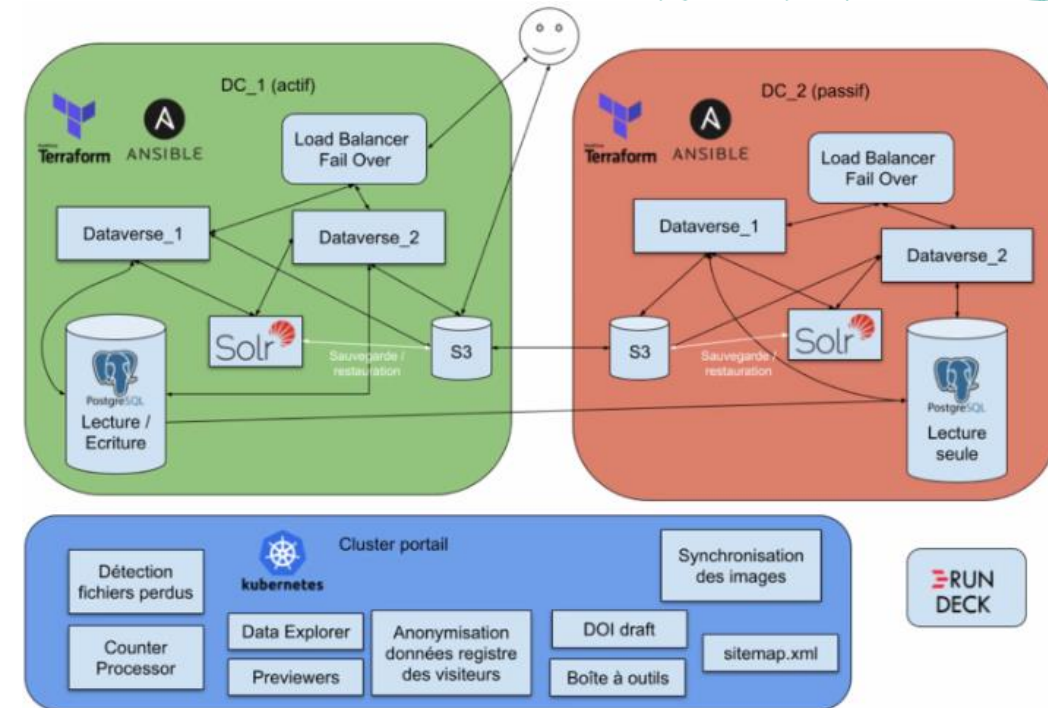
+

École des Ponts ParisTech

+

Pôle infrastructures de la plateforme

- Missions
 - Proposer des infrastructures permettant d'héberger les services de la plateforme et le stockage des données
 - Portail (ScaleWay) -& Entrepôt – Catalogue (INRAE)
 - Définition des architectures, accompagnement DevSecOPS des équipes applicatives, mise en œuvre des plateformes, MCO/MCS des plateformes, Exploitation
- Activité phare :
 - Migration vers moyens ESR (ADAGE & GRICAD)
 - Préparer les moyens pour le catalogue (GPU)



Centre de Ressources

Missions

- Accompagner et former les utilisateurs de la plateforme, en répondant à leurs questions et en dispensant des classes virtuelles.
- Prendre en charge l'administration fonctionnelle de l'entrepôt (ex création des espaces institutionnels de l'entrepôt)
- Animer le réseau métier des administrateurs et curateurs d'espaces institutionnels.
- Assurer la curation de l'espace générique.
- Contribuer au portail Recherche Data Gouv en rédigeant les pages de la rubrique Aide en ligne : guides, classes virtuelles, tutoriels, FAQ, glossaire.

Equipe multi-établissements :

- 2 copilotes
- 13 membres INRAE, CNRS, Universités de Grenoble Alpes, Strasbourg, Lorraine, Paris Nanterre.



Gestion de l'espace générique :

- Administration
- Curation des jeux de données déposés (56)

Aide en ligne

Comment partager ses données ?

Dépôts

Curateurs

Administrateurs

Chercher et explorer des données

Métadonnées

Boîte à outils



Les tutoriels

12 tutoriels réalisés en collaboration avec le Centre de ressources DoRANum.



Les classes virtuelles dédiées à la plateforme de données

<https://recherche.data.gouv.fr/fr/page/classes-virtuelles>

5 classes virtuelles dispensées régulièrement :

- Déposer un jeu de données
- Administrer une collection
- Découverte des APIs
- Utilisation des APIs
- FAIRE de la curation (Nouveauté !)

3/ Europe



Projet EOSC FIDELIS



Vise à établir et opérer un réseau européen d'entrepôts numériques de confiance (TDR) afin d'assurer la pérennité et la FAIRisation des données de recherche au sein de l'EOSC.

Harmonisation des pratiques, renforcement de l'interopérabilité et en développement les compétences des acteurs

- Kick-off le 04/02/25
- Implication en 2025 sur WP5 (Building common understanding of TDRs and a harmonized matrix of repository capabilities and characteristics)
 - Transparent Trustworthy Repository Attributes Matrix – TTRAM
 - Preparing for federation : Etat des lieux des dépôts numériques européens afin de déterminer comment FIDELIS peut au mieux les aider à s'améliorer, à s'aligner et à se fédérer avec l'EOSC en tant que TDR (Trusted Digital Repositories)



Projet EOSC LUMEN

LUMEN - Linked User-driven Multidisciplinary Exploration Network

Projet pluridisciplinaire pour développer et mettre en oeuvre des plateformes de découverte de données basées sur le web sémantique et faisant suite au projet GoTriple (Humanum)

Le catalogue Recherche Data Gouv sera un cas d'usage pour le projet LUMEN

Très complémentaire des travaux actuels basés sur recherche sémantique (~LLM/SLM) et modèle pivot DCAT classique



4/ Cycle données - stockage

Où publier vos données

Où publier vos données ?

Ce logigramme a pour objectif d'aider le déposant à identifier où il pourra publier son jeu de données.

Le jeu de données est-il communicable ?

Consultez ce guide pour en savoir plus.

OUI

NON

Le jeu de données peut être déposé dans un entrepôt de données.

Le jeu de données n'a pas vocation à être publié.

Existe-t-il un entrepôt thématique recommandé ?
Ou à défaut un entrepôt institutionnel ?

OUI

NON

Il est recommandé de déposer le jeu de données dans cet entrepôt.

Si celui-ci ne fournit pas de DOI, il est possible de décrire le jeu de données sans téléverser les fichiers dans l'entrepôt Recherche Data Gouv.

Le jeu de données peut être déposé dans l'entrepôt Recherche Data Gouv.



Il est recommandé de déposer le jeu de données dans cet entrepôt.

Si celui-ci ne fournit pas de DOI, il est possible de décrire le jeu de données sans téléverser les fichiers dans l'entrepôt Recherche Data Gouv.

Le jeu de données peut être déposé dans l'entrepôt Recherche Data Gouv.



Au moins un des co-auteurs est affilié à un établissement disposant d'un espace institutionnel.

Un seul espace est identifié.

Plusieurs espaces possibles sont identifiés.

Aucun espace n'est identifié.

Le jeu de données peut être déposé dans cet espace.

Pour tout renseignement complémentaire, contacter l'administrateur de l'espace.

Le jeu de données peut être déposé dans l'espace générique.

Pour tout renseignement complémentaire, contacter l'administrateur de l'espace.

Les co-auteurs s'accordent sur le choix de l'espace institutionnel.

Puis le jeu de données sera affiché dans les autres espaces concernés.





RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité

recherche.data.gouv.fr

Fiche chrono dépôt



Prérequis :

Avoir un ensemble de fichiers prêts : organisation logique, nommage clair et sans accents, présence d'un README (fichier texte permettant d'expliquer le contenu du jeu de données et d'éventuelles consignes d'utilisation)

Avoir à portée de main les références : de la publication associée, du projet lié, de la subvention (Optionnel)

Être connecté : privilégier la connexion institutionnelle via le lien en haut à droite.

Identifier son espace de dépôt : l'espace institutionnel de votre établissement s'il existe, sinon l'espace générique. Certains espaces institutionnels proposent des collections. Dans tous les cas, vérifier de bien avoir l'autorisation de déposer : le bouton « Ajouter des données » « Nouveau jeu de données » doit apparaître.

Projet, unité ou établissement

+ Ajouter des données

Nouvelle collection

Nouveau jeu de données

Dans le cas contraire, demander les droits de dépôt à l'équipe d'administration via le bouton [Contact](#) de l'espace.

Étape 1 : Déposer

La page de création de jeu
de données apparaît :

Modèle : Des modèles peuvent être proposés dans la collection hôte pour faciliter le dépôt.

Titre : Comme pour un titre d'article, le titre du jeu de données doit être suffisamment explicite et informatif sur les données déposées.

Auteurs : Conserver la forme (Nom, Prénom) pré-remplie et renseigner l'affiliation sous la forme **Acronyme du laboratoire ; Tutelle 1, Tutelle 2... ; Pays.** Ajouter l'ORCID (recommandé)
Note : Ajouter les éventuels coauteurs de la même façon.

Collection hôte

Changer la collection hôte effacera tous les champs dans lesquels vous pourriez avoir entré des données

Modèle de jeu de données

Modifier le modèle effacera tous les champs dans lesquels vous auriez entré des données.

Langue des métadonnées du jeu de données

Sélectionner la langue que vous souhaitez utiliser pour entrer les métadonnées.

Métadonnées générales

Titre

Un titre informatif

Ajouter « Données de réplication pour » au titre

Lien vers les données

(Facultatif) Un lien vers les données

Point de contact

Nom

Personne contact

Courriel

Un courriel valide

Auteur

Nom

Auteur du jeu de données

Type d'identifiant

ORCID

Affiliation

Unité ; Tutelle 1, Tutelle 2, etc. ; PAYS

Identifiant

0000-0000-0000-000X

Producteur

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels

Nom

Établissement tutelle

Nom abrégé

Acronyme ou sigle de la tutelle

URL du logo

Affiliation

Organisation XYZ

URL

Site internet de la tutelle

Collection hôte : Collection dans laquelle le jeu de données sera placé.
Note : il est possible de demander de déplacer un jeu de données par la suite en contactant les administrateurs de la collection.

Langue des métadonnées du jeu de données : La langue de saisie des métadonnées. Une fois la langue choisie, il n'est plus possible de la modifier dans les étapes suivantes.

Lien vers les données : à renseigner uniquement si aucun fichier n'est ajouté au jeu de données.

Point de contact : Contact pour ce jeu de données. Renseigner a minima un courriel.

Producteur : Renseigner tout organisme détenant la responsabilité administrative ou financière sur les données déposées. En général, les cotutelles.

Pays. Ajouter l'ORCID (recommandé)
Note : Ajouter les éventuels coauteurs de la même façon.

Description : A l'image d'un résumé d'article, préciser le contexte, les objectifs, la nature des données, l'approche utilisée, etc.

Sujet : Choisir un ou plusieurs sujets appropriés dans la liste.

Type de données : Choisir les éléments appropriés (au minimum : dataset).

Producteur

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels

Nom

Établissement tutelle

Nom abrégé

Acronyme ou sigle de la tutelle

URL du logo

Lien vers le logo de la tutelle

Ce champ prend en charge uniquement certaines balises HTML

Texte

Un descripteur écrit au format HTML, comportant des balises pour sa mise en forme, par exemple
 pour aller à la ligne, <i> pour écrire en gras ou en italique, etc. Une description littérale présente le contexte, les objectifs, les données et la méthode de façon à donner envie aux usagers de l'explorer, à l'image d'un résumé d'article.

Langue

French

Date

AAAA-MM-JJ

Terme

mot-clé

URI du terme

https://thesaurus.inrae.fr/

Nom du vocabulaire

Thésaurus INRAE

URI du vocabulaire

https://thesaurus.inrae.fr/

Type de données

Dataset

Publication associée

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels

Citation

Hubert Blanc, Laetitia Dracco, Recherche.data.gouv.fr : un écosystème au service du partage et de l'ouverture des données de recherche. Semaine Data-GRH, Centre de Données Socio-Politiques, Sciences Po, CHADES, Plateforme Universitaire de Données des Grands Miroirs, Université Paris Cité, Plateforme Universitaire de Données de Nanterre, MSH Mondes, Dec 2022, PARIS, France.

Type d'identifiant

halid

Identifiant

hal-04102550

URL

https://hal.science/hal-04102550v1

Producteur : Renseigner tout organisme détenant la responsabilité administrative ou financière sur les données déposées. En général, les cotutelles.

Mot-clé : Termes décrivant les aspects importants du jeu de données. De préférence, ces termes doivent être issus de vocabulaires de référence. Dans ce cas, préciser l'URI (identifiant) du terme, le vocabulaire employé ainsi que l'adresse à laquelle il est accessible en ligne.

Publication associée : Citation bibliographique de la référence d'un article lié aux données déposées. Pour les articles avec DOI : choisir le type d'identifiant DOI, puis copier-coller le DOI
Note : Pour HAL : choisir le type Halid ; puis copier l'identifiant sous la forme hal-xxxxxxx (ex. hal-00083055).



@RechercheDataGv



Fiche chrono dépôt



Informations sur la subvention

Agence: Organisation XYZ Identifiant: Ex: ANR-XX-XXXX-XXXX

Informations sur le projet

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Identifiant: Ex: ANR-XX-XXXX-XXXX Acronyme: EDSC-PHAR

Titre: Coordination and harmonisation of Natic Lien vers le site web: Adresse URL, commençant par https://

Work Package: Tâche:

Téléverser des fichiers: Téléverser ici les fichiers à déposer. Pour un grand nombre de fichiers, utiliser DV Uploader qui permet de reproduire l'arborescence de dossiers/sous-dossiers automatiquement.

Sauvegarder le jeu de données et passer à l'étape 2!

Étape 2 : Compléter

Origine des données: Possibilité de cocher plusieurs origines.

Jeu de données associé: Tout autre jeu de données lié à ce jeu de données. Pour les jeux de données avec DOI: choisir le type d'identifiant DOI, puis copier-coller le DOI.

Informations sur la subvention / sur le projet: Renseigner l'identifiant du projet ANR, un remplissage automatique des métadonnées de subvention, de titre et d'acronyme du projet est proposé.
Note: L'autocomplétion ne fonctionne qu'avec la première occurrence.

Modifier des fichiers: Le menu « Modifier » donne accès à des options supplémentaires sur les fichiers, notamment la mise en place d'embargo ou de restriction d'accès.

Contributeur: Toute personne ayant contribué au jeu de données, en dehors des auteurs. Renseigner sous la même forme que celle des auteurs et préciser le type de contribution. Note: à la différence d'un auteur, un contributeur n'apparaît pas dans la citation.

Logiciel: Nom et version de tout logiciel utilisé pour générer les données.

Période couverte: à remplir si pertinent

données lie à ce jeu de données. Pour les jeux de données avec DOI: choisir le type d'identifiant DOI, puis copier-coller le DOI.

Période couverte: à remplir si pertinent

Date de collecte: à remplir si pertinent

Des blocs de métadonnées complémentaires, spécialisées ou disciplinaires, sont disponibles. Exemples:
- Métadonnées géospatiales: si le jeu de données concerne un lieu ou une aire géographique précis
- Métadonnées sur les sciences sociales et les sciences humaines: pour les données d'enquête ou les données sociologiques, etc.

Étape 3 : Soumettre

Sauvegarder le jeu de données: Sauvegarder, la création est terminée!

Cliquer sur **Soumettre pour révision** pour envoyer le jeu de données au curateur. Après révision, le jeu de données sera:
• soit validé et disponible immédiatement en ligne,
• soit retourné pour propositions d'améliorations. Apporter les modifications suggérées avant de le soumettre à nouveau.
Note: dans le cadre d'une curation après publication, cliquer directement sur « Publier ».

Modalités d'accès au jeu de données

Soumettre pour révision

Modifier le jeu de données

Contact Partager



Convention avec institutions

- Quasi toutes : valides jusqu'en fin 2024
 - En cours de renouvellement
- Charte / Guide du curateur
 - Ex repérage de données non publiées
 - Bientôt, la plateforme proposera un service payant de curation pour ceux qui le souhaitent
 - Et un service d'analyse de la curation (a priori et a posteriori)

Niveau A : Contenu publié tel que déposé

- Ce niveau n'est pas accepté dans l'entrepôt Recherche Data Gouv sans curation a posteriori.

Niveau B : Curation de base

- Présence de métadonnées et de documentation suffisantes pour comprendre le contexte de production des données.

Niveau C : Curation avancée

- Conversion vers des formats ouverts ou largement utilisés si besoin ;
- Description précise de la production, de l'organisation et de la structuration des données :
 - Dictionnaire de données, README, plan de gestion de données, etc., documentant les schémas et vocabulaires contrôlés utilisés pour les variables et valeurs ;
 - Informations sur la provenance des données.

Niveau D : Curation au niveau des données

- Annotation et édition des données : requiert une expertise dans le domaine scientifique.

Choix du niveau de curation dans l'entrepôt

Le choix du niveau de curation relève de l'administrateur de la collection.

Néanmoins, tous les jeux de données déposés dans l'entrepôt Recherche Data Gouv font l'objet d'une curation. Cette curation porte sur la qualité documentaire et **garantit au minimum un niveau B**.

Les auteurs sont seuls responsables de la qualité scientifique des données et du respect du droit à diffuser ces données.



RÉPUBLIQUE
FRANÇAISE recherche.data.gouv.fr

Convention avec les institutions pour disposer d'un espace dédié sur l'entrepôt Recherche Data Gouv

Préambule

Le comité de pilotage de la science ouverte, présidé par la DGRI du Ministère de l'enseignement supérieur et de la recherche, a confié le développement et la gestion des modules « entrepôt » et « catalogue », ainsi que du portail web de recherche.data.gouv.fr à INRAE. L'entrepôt s'appuie sur Dataverse, la solution libre développée par Harvard et largement adoptée par la communauté française.

Les soussignés :

1. L'Institut National de la Recherche pour l'Agriculture, l'Alimentation et l'Environnement, Etablissement public à caractère scientifique et technologique, dont le siège est situé 147 rue de l'Université - 75338 PARIS CEDEX 07, représenté par son Président-directeur général, Monsieur Philippe MAUGUIN, représentant des modules Entrepôt et Catalogue Recherche Data Gouv, ci-après dénommé : « le représentant de l'Entrepôt - Catalogue Recherche Data Gouv » et par délégation par Madame Odile Hologne, Directrice de la Direction Science Ouverte;

ET

2. L'INSHEA, Institut national supérieur de formation et de recherche pour l'éducation des jeunes handicapés et les enseignements adaptés, établi au 58-60 avenue des Landes à Suresnes (92150), légalement représenté par Madame Murielle MAUGUIN, ci-après dénommée : « l'institution partenaire » ;
Les parties décrites ci-dessus aux points 1) et 2) sont ci-après dénommées les « Parties » et chacune une « Partie ».

Ont convenu de ce qui suit :



Convention avec institutions

- Stockage
 - Limite de volumétrie : 5 To
 - Evolution en cours
 - Une partie (qui sera > 5 To) gratuite
 - Possibilité de tarification au-dessus
- « Pérennité »
 - De la plateforme
 - Financements assurés par FNSO jusqu'en fin 2026
 - Au-delà, plans pour devenir Infrastructure de Recherche*
 - Recherche de financements complémentaires
 - DINUM, ANR, Europe, ...
 - Des données
 - Hébergement sécurisé mise à disposition des données pendant un minimum de 5 ans renouvelables après la publication.
 - Prévu d'être étendu à 15 ans (mais pas de notion d'archivage pérenne, plutôt)
 - Au-delà : GT archivage du collège des données du COSO

