

# GÉRER SES DONNÉES SUR LE PLATEAU DE SACLAY

11 FÉVRIER 2025, 9H À 18H

BÂTIMENT D'ENSEIGNEMENT MUTUALISÉ, INSTITUT POLYTECHNIQUE DE PARIS, PALAISEAU



## Gestion des données de la recherche politique nationale, réalisations de terrain et perspectives

*Gilles Mathieu – Ministère chargé de  
l'enseignement supérieur et de la recherche*



# Ambition

posée par la loi pour une république numérique (2016)  
rappelée par la science ouverte

Ouvrir l'ensemble des produits et des méthodes de la recherche

- >> Les publications
- >> Les données
- >> Les codes sources

*Une science plus cumulative, plus robuste, plus reproductible*

*Une science transparente et accessible à toutes et tous*

# Concevoir nativement des données réutilisables

## Travail scientifique de FAIRisation des données, renseigner :

- Les origines et les transformations des données
- Les guides de codification
- Les modèles et la description détaillée des méthodes de collecte, de nettoyage, d'analyse
- Le code spécifique à un instrument
- Les logiciels conçus pour des applications propres au domaine

**Plus la réutilisation est loin du point d'origine, en termes de temps, de problématique de recherche, de discipline... plus il devient difficile de retrouver et d'interpréter des données**

# Concevoir nativement des données réutilisables

## Travail scientifique de FAIRisation des données, renseigner :

- Les origines et les transformations des données
- Les guides de codification
- Les modèles et la description détaillée des méthodes de collecte, de nettoyage, d'analyse
- Le code spécifique à un instrument
- Les logiciels conçus pour des applications propres au domaine

**Plus la réutilisation est loin du point d'origine, en termes de temps, de problématique de recherche, de discipline... plus il devient difficile de retrouver et d'interpréter des données**

**Le premier bénéficiaire d'une donnée bien décrite est celui qui l'a produite !**

**Retrouver et réutiliser ses données, mieux les gérer, gain de temps et réduction de l'empreinte numérique**

**Taux de citations des publications dont les données sont partagées : + 25% à + 75%**

**Les données partagées reconnues comme une production de la recherche**

# Les engagements

*Deuxième plan national pour la science ouverte*  
*Politique des données, des algorithmes et codes sources*

**Soutenir la structuration, la préservation, le partage, l'ouverture,  
la découverte des données de la recherche**

**Favoriser les pratiques de gestion raisonnée et écoresponsable des données**

**Favoriser les pratiques de réutilisation des données de la recherche**

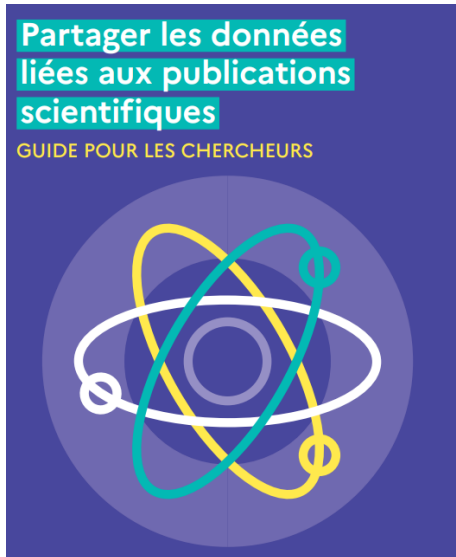
**Fédérer les initiatives locales et thématiques**



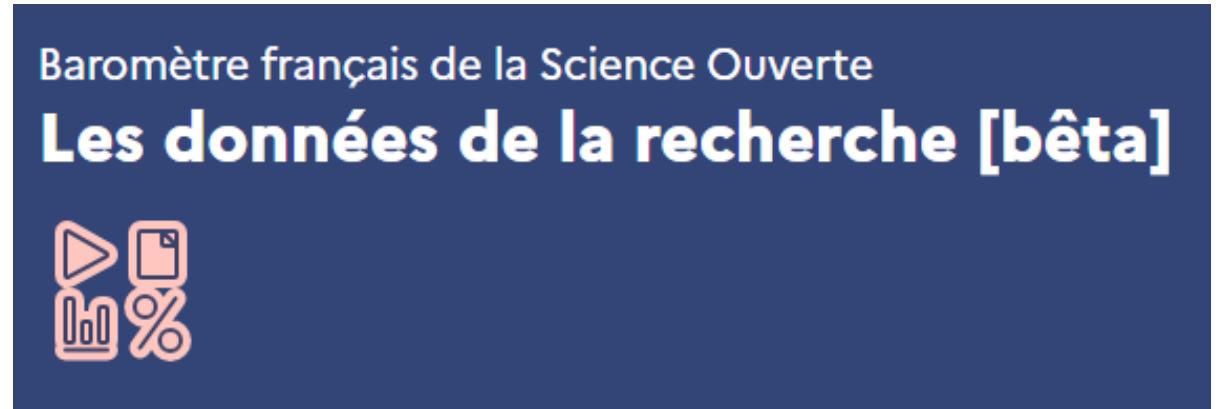
# Données de la recherche : guides, actions...



**Partager les données  
liées aux publications  
scientifiques**  
GUIDE POUR LES CHERCHEURS



<https://www.ouvrirlascience.fr/comment-partager-les-donnees-liees-a-vos-publications-scientifiques-un-guide-vous-accompagne/>



<https://barometredelascienceouverte.esr.gouv.fr/donnees-de-la-recherche/general/>



<https://www.ouvrirlascience.fr/wp-content/uploads/2024/03/24-02-22-Donnees-FR-WEB.pdf>



## Qui cela concerne-t-il ?



**Communauté de recherche**  
*Chercheurs, ITA...*



**Soutien à la recherche**  
*Personnels accompagnants, SCD,  
services informatiques...*



**Etablissements**  
*Tutelles, financeurs, état...*



## Qui cela concerne-t-il ?



Communauté de recherche  
Chercheurs, ITA

**Ecoute du besoin**



**Mise en réseau**

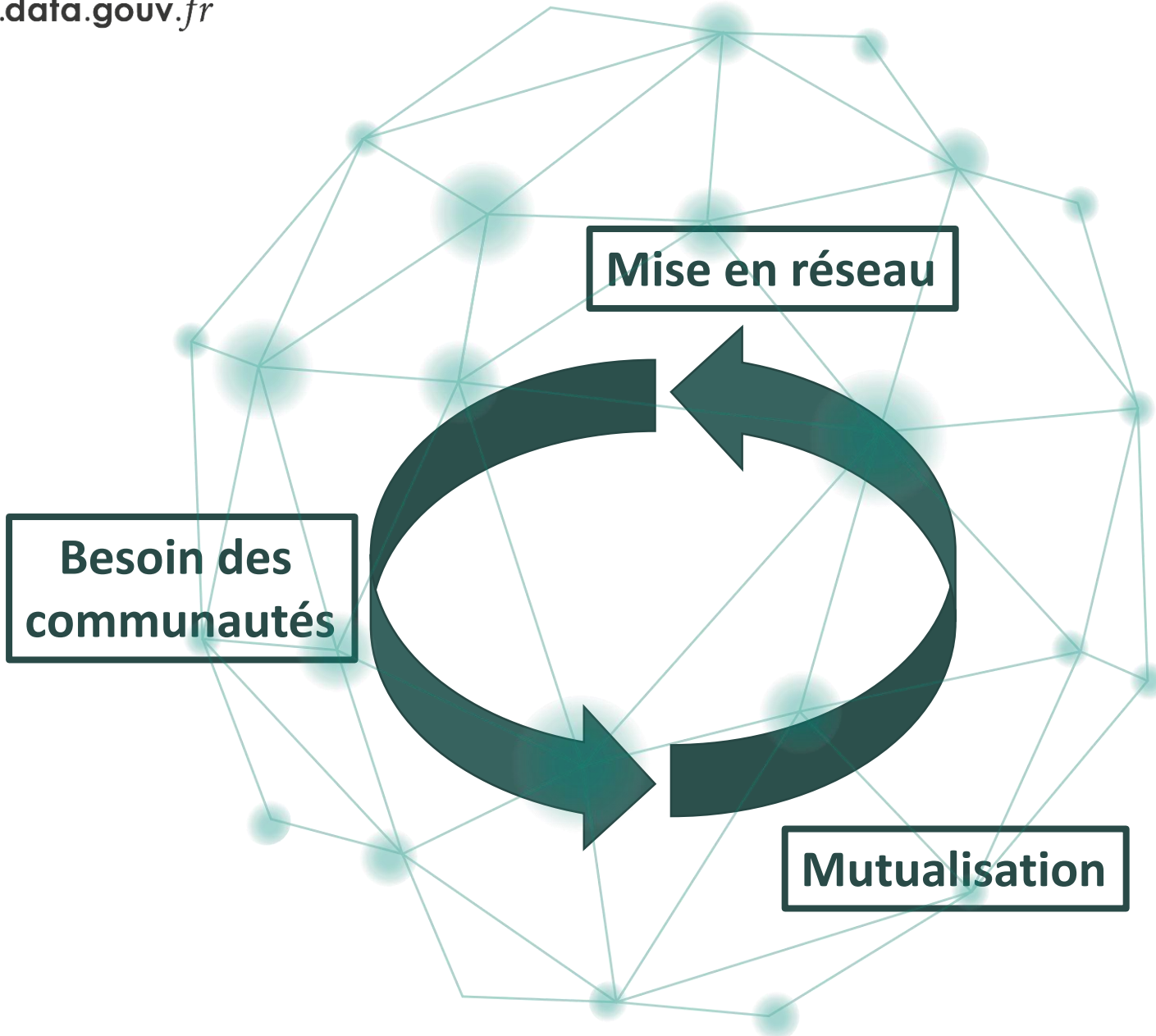
Soutien à la recherche  
Personnels accompagnants, SCD,  
services informatiques...

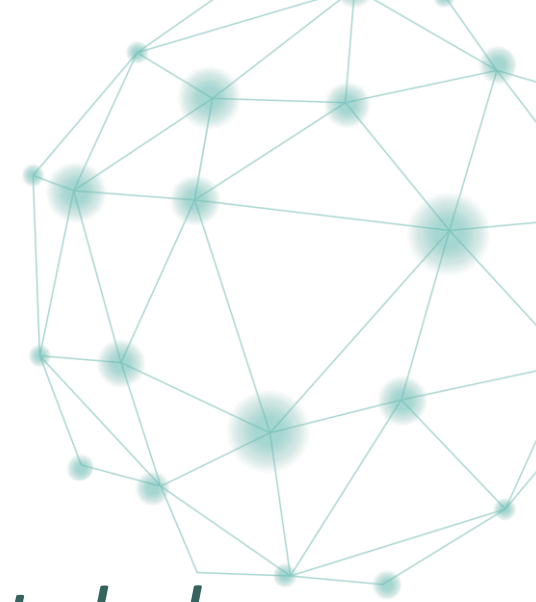


Etablissements  
Tutelles, financeurs, état...

**Mutualisation**







# Recherche Data Gouv

*un écosystème construit par et pour les établissements de l'enseignement supérieur et de la recherche pour vous accompagner à la gestion, au partage et à l'ouverture de vos données de recherche*

## Objectifs

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

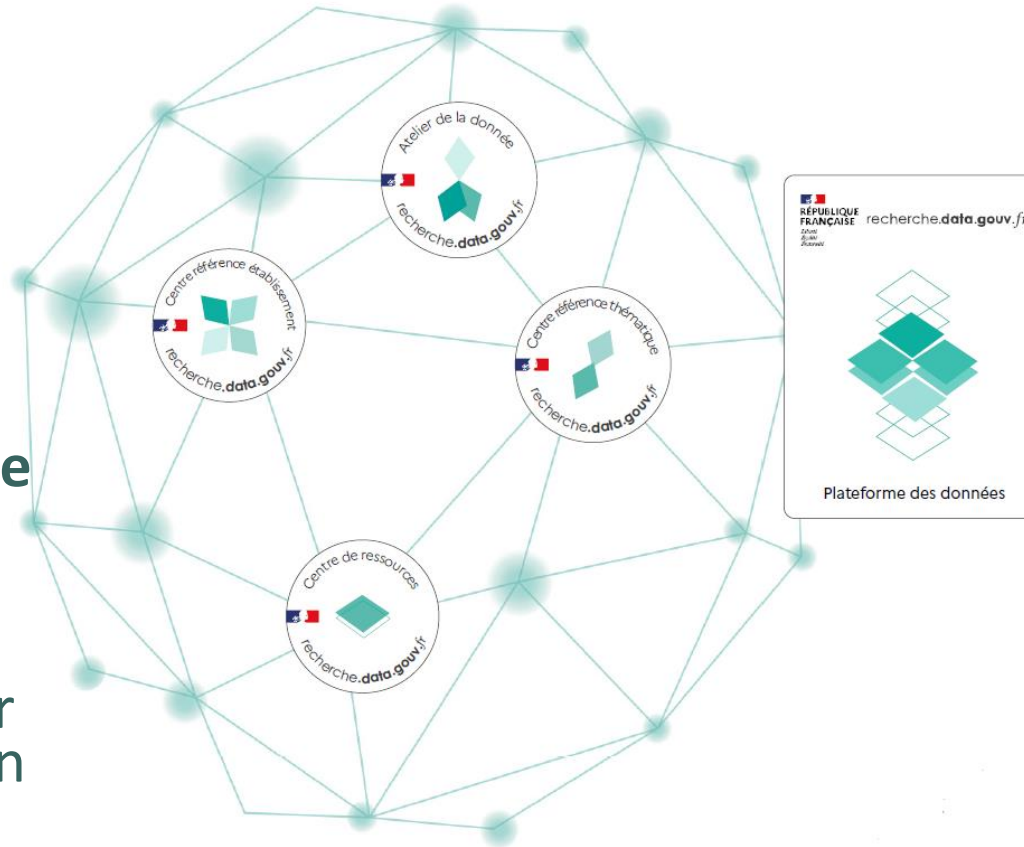
Faire de l'accompagnement un élément central du dispositif

Reconnaître, fédérer et amplifier les initiatives existantes

## Concrètement, Recherche Data Gouv, c'est :

### une fédération nationale de centres de compétences

via un maillage de dispositifs d'accompagnement sur la gestion et la diffusion des données de recherche, au plus près des équipes de recherche.



### une plateforme nationale des données de recherche

c'est-à-dire une solution souveraine pour déposer, publier, signaler, découvrir, rendre accessible et réutilisable les données de recherche grâce à :

un entrepôt national de données de recherche ;

un catalogue des données publiées sur l'entrepôt lui-même ou sur des entrepôts externes.

# Accompagnement de proximité via les ateliers de la donnée

*Expertise généraliste en proximité des équipes de  
recherche pour toute question relative à la donnée*

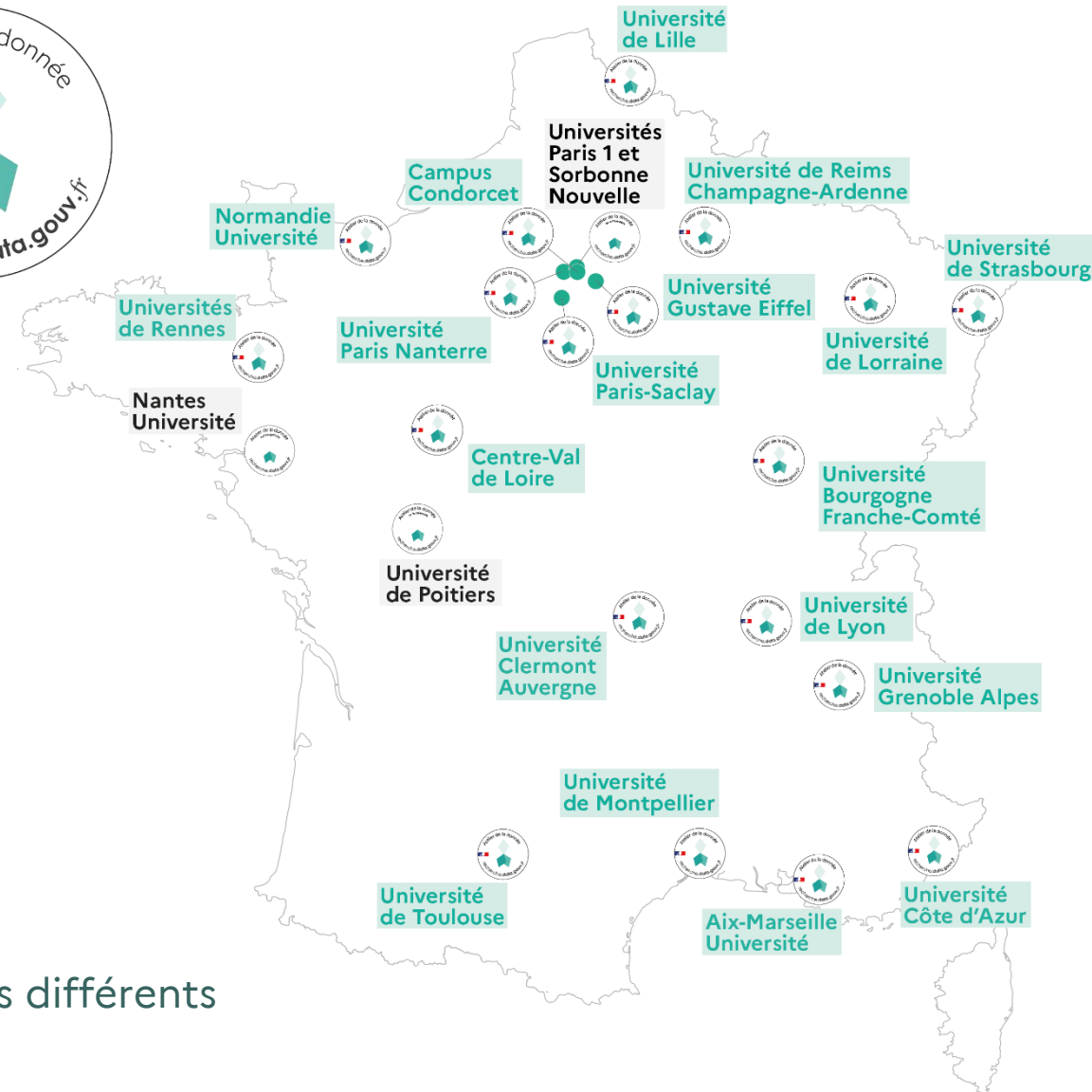
Regroupement d'expertises complémentaires de  
différents établissements

Un réseau d'entraide, de mutualisation et de  
capitalisation des services actifs

**22 ateliers de la donnée opérationnels**

Plusieurs ateliers en projet

100+ ETP  
400+ personnes  
90+ établissements différents



# Les centres de référence établissement



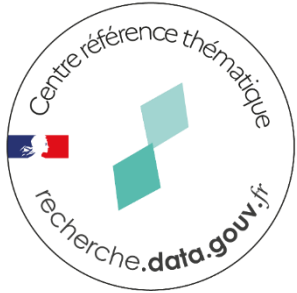
## *Offre d'accompagnement propre aux orientations de recherche d'un établissement*

Etablissements nationaux de l'Enseignement supérieur et/ou de la Recherche

Proposent depuis plusieurs années une offre d'accompagnement propre à leurs orientations de recherche et à destination de leurs équipes de recherche



## Les centres de référence thématiques



### *Expertise par domaine scientifique*

Définition et diffusion des **bonnes pratiques et des standards internationaux de gestion, traitement et diffusion des données** par domaine scientifique

**Premier paysage :**  
infrastructures de recherche ayant une activité structurante de gestion et diffusion de données pour leur communauté scientifique

Réflexion en cours entre toutes les infrastructures nationales en biologie et santé




## Les centres de ressources au service de tous les autres modules de Recherche Data Gouv



Centre de ressources  
supports pédagogiques, e-formations  
 Recherche Data Gouv


*Mise en commun et certification de **supports pédagogiques** et élaboration et diffusion de **e-formations généralistes***



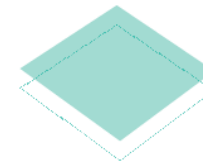
Centre de ressources  
compétences  
 Recherche Data Gouv

*Dispositif de **montée en compétences**  
pour les ateliers de la donnée*



Centre de ressources  
outils communs aux ateliers  
 Recherche Data Gouv

*Élaboration et maintien d'**outils communs aux ateliers***



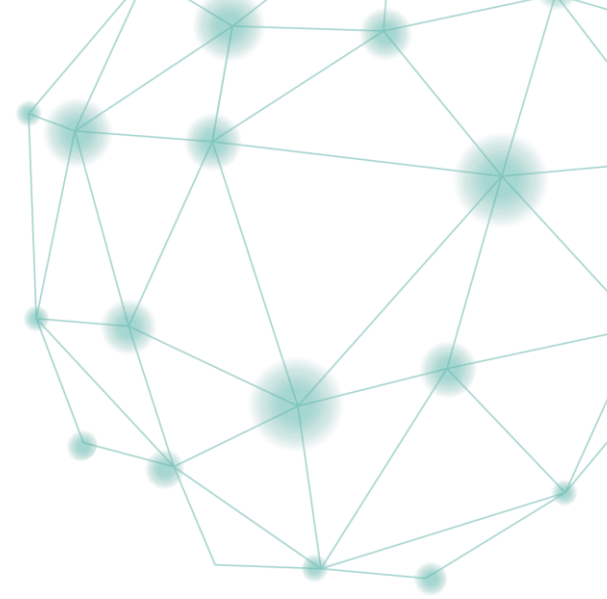
Centre de ressources  
entrepôt-catalogue  
 Recherche Data Gouv

*Développe et maintien l'**entrepôt-catalogue** et  
accompagne les utilisateurs*

# Un entrepôt de données mutualisé

## entre les établissements d'enseignement supérieur et de la recherche

- Une offre nationale développée et maintenue par des établissements au service de toute la communauté de recherche
- Les équipes de recherche déposent et publient directement leurs données
- Des espaces pour les établissements qui assurent la gestion et la modération de leurs données
- 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



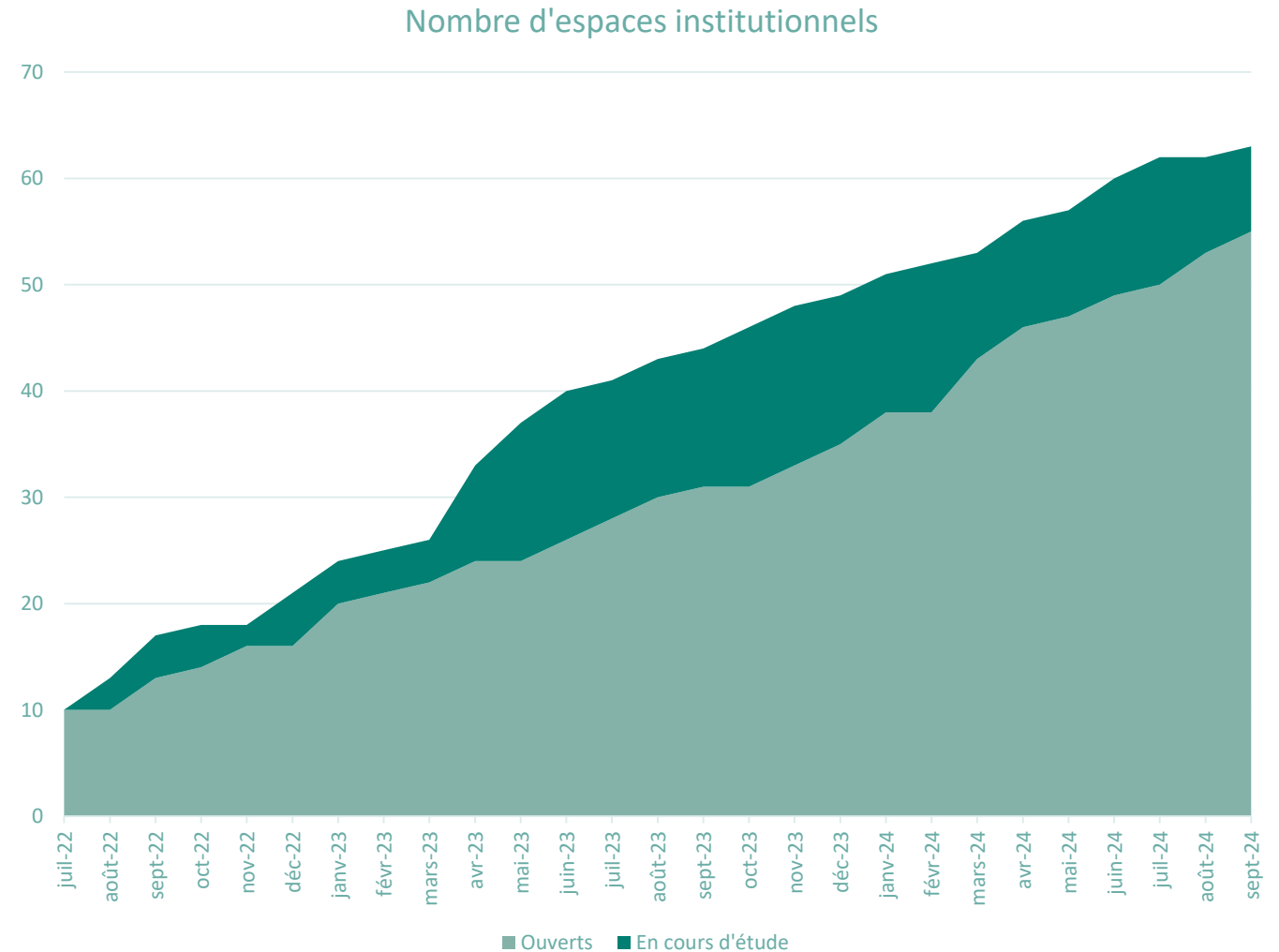


## Un écosystème en expansion

- Poursuite des appels à labellisation des ateliers de la donnée
  - <https://recherche.data.gouv.fr/fr/page/appels-a-manifestation-dinteret-ateliers-de-la-donnee>
  - 4<sup>ème</sup> appel ouvert en décembre 2023, terminé au printemps 2024
  - 5<sup>ème</sup> appel ouvert en juin 2024, terminé en octobre 2024
  - Un 6<sup>ème</sup> appel ouvrira en mars 2025, puis sur une base annuelle
- Plusieurs nouveaux projets d'ateliers en réflexion

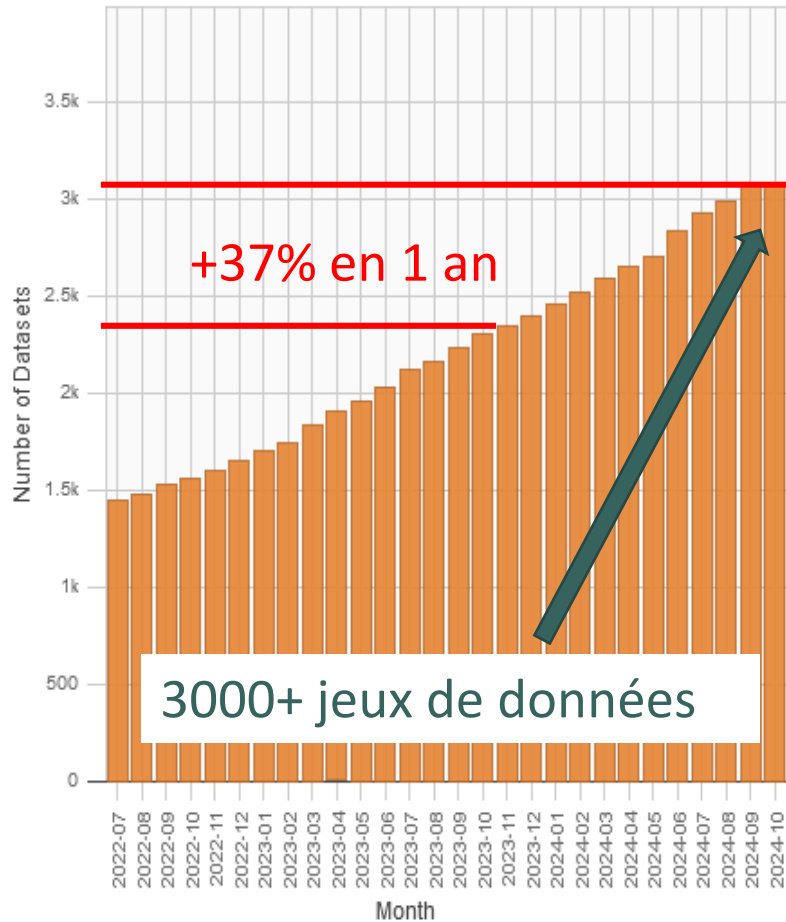
## De plus en plus d'établissements impliqués

- 55 espaces institutionnels ouverts
- 20 nouveaux depuis un an (+57%)

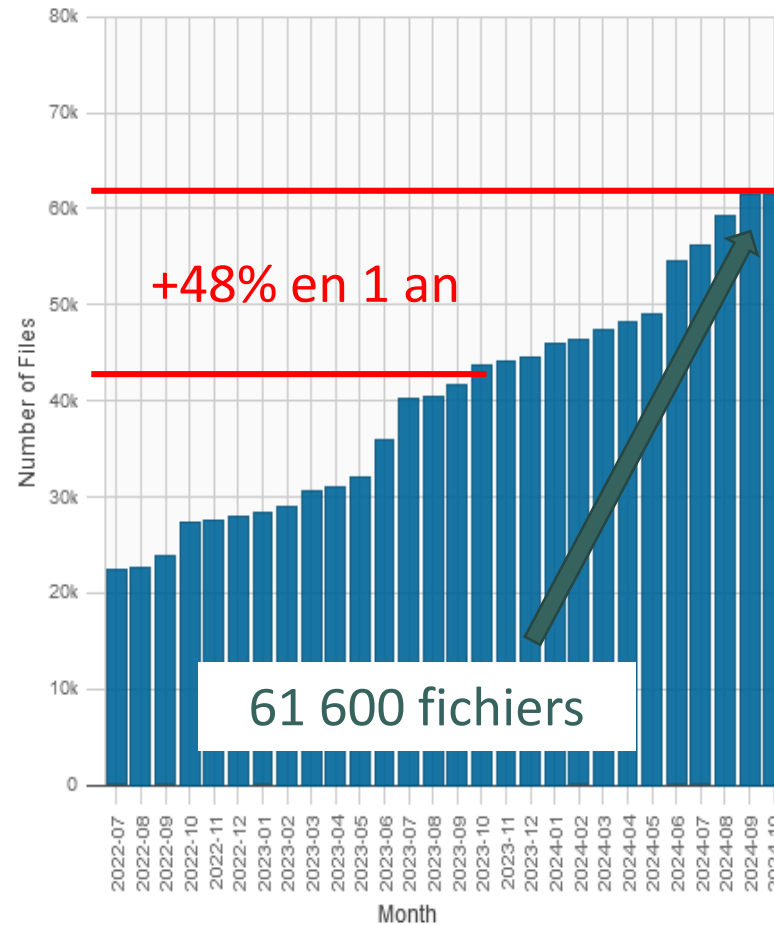


# Un usage toujours croissant

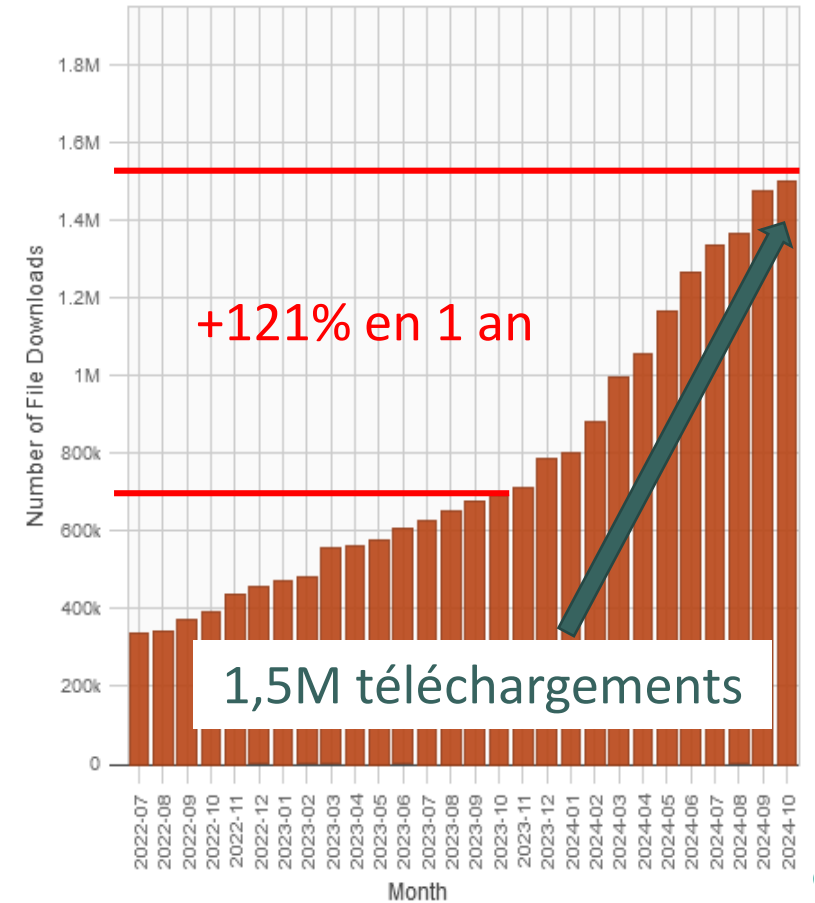
Total Datasets



Total Files



Total File Downloads



## Entrepôts de données disciplinaires de confiance

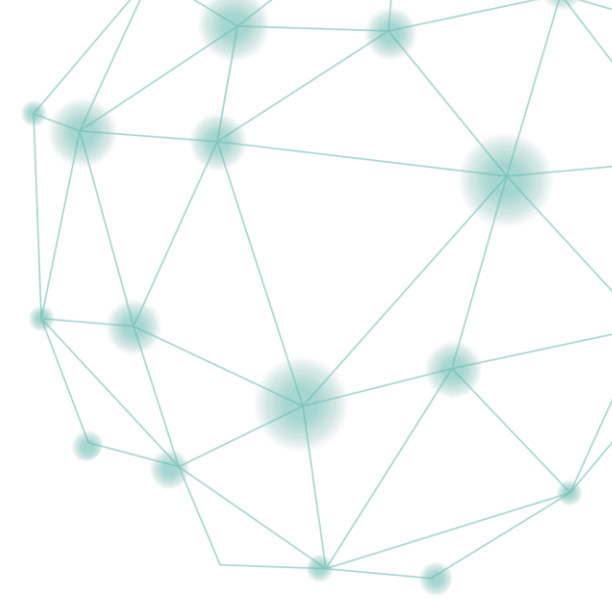
- Publication au printemps 2024 d'une liste établie par le comité pour la science ouverte
  - <https://www.ouvrirlascience.fr/selectionner-un-entrepot-thematique-de-confiance-pour-le-depot-de-donnees-methodologie-et-analyse-de-loffre-existante/>
- Mise en ligne sur le portail Recherche Data Gouv d'un outil de visualisation des entrepôts de cette liste
  - <https://recherche.data.gouv.fr/fr/entrepots>
- Travail en cours pour l'amélioration de la fonctionnalité

## Travail sur la mise à disposition d'informations juridiques

- Mise en place d'un outil de ressources juridiques sur le droit des données de la recherche
  - Mission confiée à l'Université de Montpellier (Agnès Robin, Laura Tomasso)
  - Début de l'implémentation du portail en juillet 2024



# Ouverture des objets de la recherche



Publications

Une nécessaire  
complémentarité

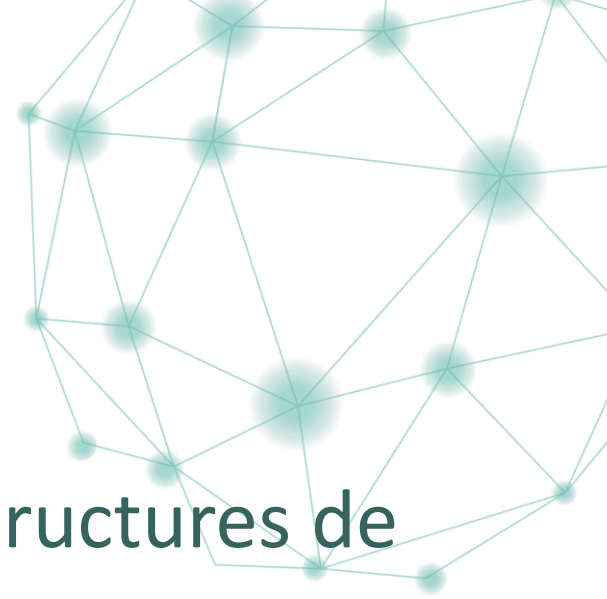
Données



Codes sources



Software Heritage



## Stratégie internationale : objectifs

- **Initier ou renforcer** les partenariats avec des infrastructures de recherche similaires dans d'autres pays
- **Contribuer à l'harmonisation des normes d'interopérabilité** des données partagées de la recherche
- **Favoriser le développement des pratiques** de production de données réutilisables
- **Créer un réseau d'entrepôts de confiance** pour la diffusion des données de la recherche.

## Stratégie internationale : ambition

- **Offrir un accès universel et une meilleure compréhension des données de la recherche française à tous** : chercheurs de toutes disciplines, acteurs socio-économiques et citoyens.
- **Être reconnu comme une référence mondiale** pour l'accès et la gestion des données de recherche.





# Stratégie internationale : actions concrètes

## En tant que plateforme pour l'ouverture et le partage des données

- Co-construire le réseau européen d'entrepôts de confiance
- Travailler au-delà des frontières sur les sujets transdisciplinaires



## En tant que réseau d'accompagnement

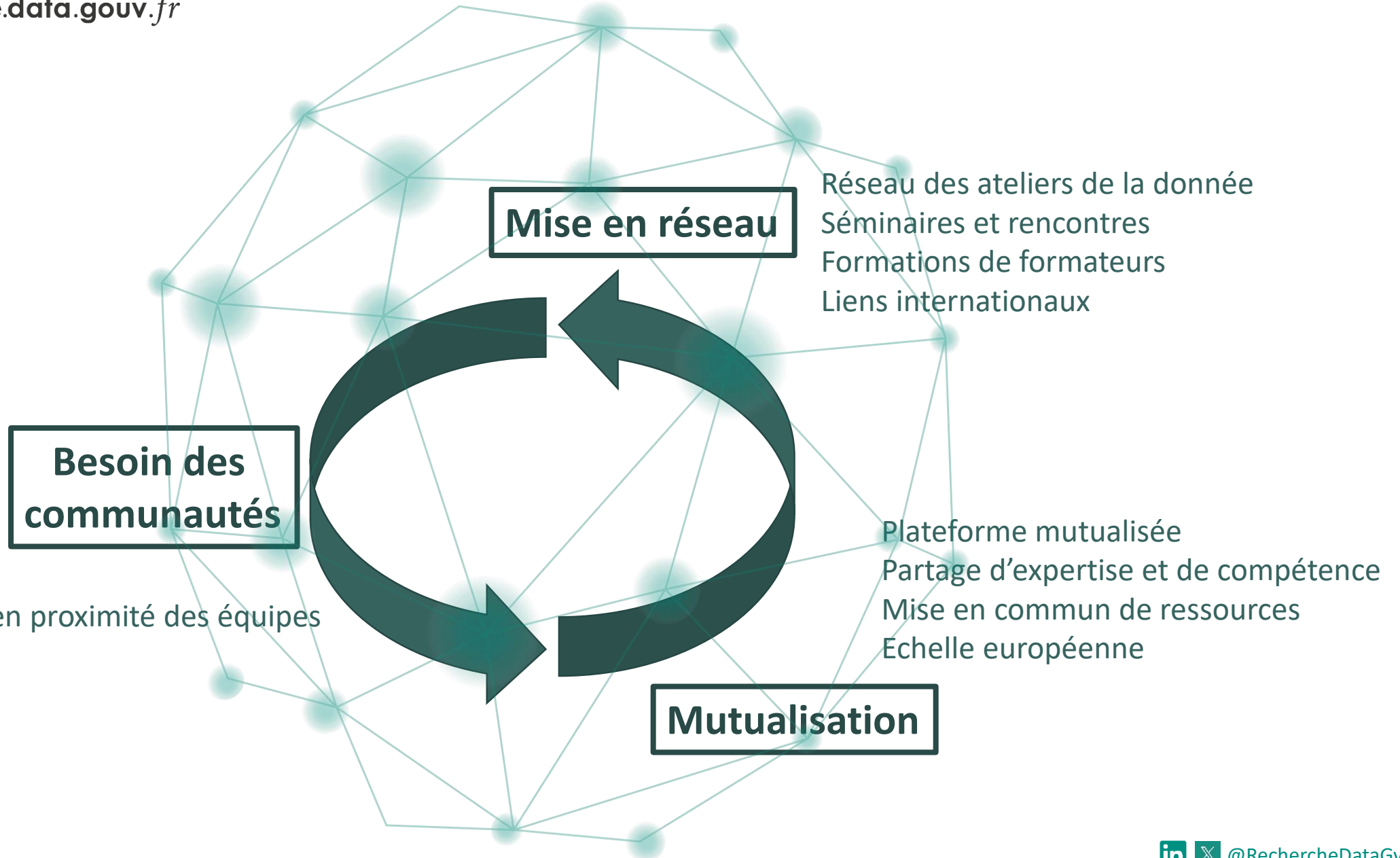
- Première brique d'un centre de compétences national
- Partage des pratiques d'accompagnement à l'international

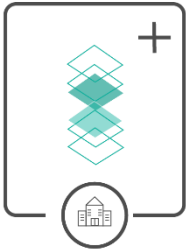


## En tant que service national de données

- Partage de bonnes pratiques
- Liaison avec d'autres groupes internationaux







## Créer un espace institutionnel

Gérer l'ouverture et le partage des données de son établissement

[http://recherche.data.gouv.fr/uploads/general/Documents/Demande\\_creation\\_espace\\_institutionnel.pdf](http://recherche.data.gouv.fr/uploads/general/Documents/Demande_creation_espace_institutionnel.pdf)



## Devenir atelier de la donnée

Pour accompagner en proximité les équipes de recherche dans la gestion, la structuration et la diffusion de leurs données

<https://recherche.data.gouv.fr/fr/page/rejoindre-lecosysteme>



## Soutenir le dispositif et participer à la gouvernance

Mobilisation de ressources d'établissements au sein d'une unité mixte de service qui maintient et fait évoluer l'entrepôt mutualisé des données

**Un écosystème  
ouvert à toute la  
communauté de  
recherche et  
construit par les  
établissements  
partenaires**



RÉPUBLIQUE  
FRANÇAISE

*Liberté  
Égalité  
Fraternité*

recherche.data.gouv.fr

## Pour suivre toutes les actualités

<https://recherche.data.gouv.fr/fr>



@RechercheDataGv



<https://www.linkedin.com/company/recherche-data-gouv/>

