



21^{ème}

ÉCOLE INTER-ORGANISMES

QUALITÉ ET RESPONSABILITÉ SOCIÉTALE

EN RECHERCHE ET EN ENSEIGNEMENT SUPÉRIEUR

du 12 au 14 Septembre 2023

Recherche Data Gouv : Un écosystème au service du partage et de l'ouverture des données de recherche

Gilles Mathieu - *Chargé de Mission « données de la Recherche »*

Isabelle Blanc - *Administratrice Ministérielle des Données, Algorithmes et Codes sources*

Ministère de l'Enseignement Supérieur et de la Recherche

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

Septembre 2023



Sujets abordés

- Recherche Data Gouv : ambition et objectifs
- Les différents dispositifs et leur fonctionnement
- Les clés de la réussite collaborative





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 !

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 réutilisation des données de la recherche

Créer Recherche Data Gouv

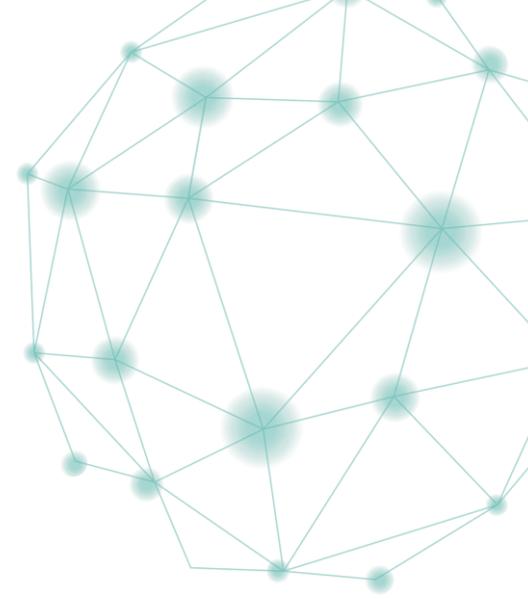




Objectifs

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

Faire de l'accompagnement un élément central du dispositif





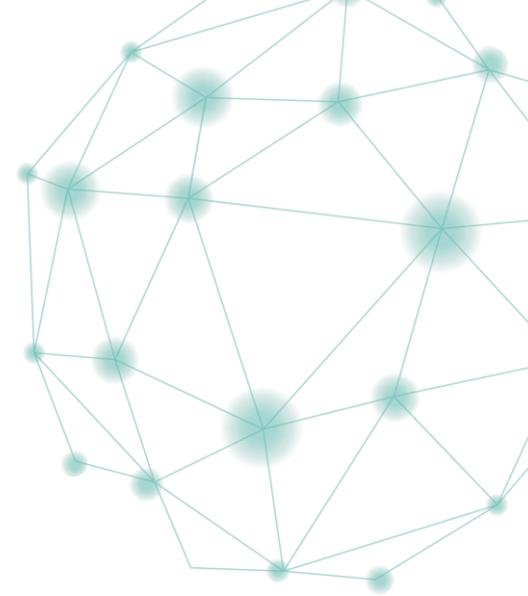
Objectifs

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

Faire de l'accompagnement un élément central du dispositif

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

- Sensibiliser, former ou accompagner les équipes de recherche

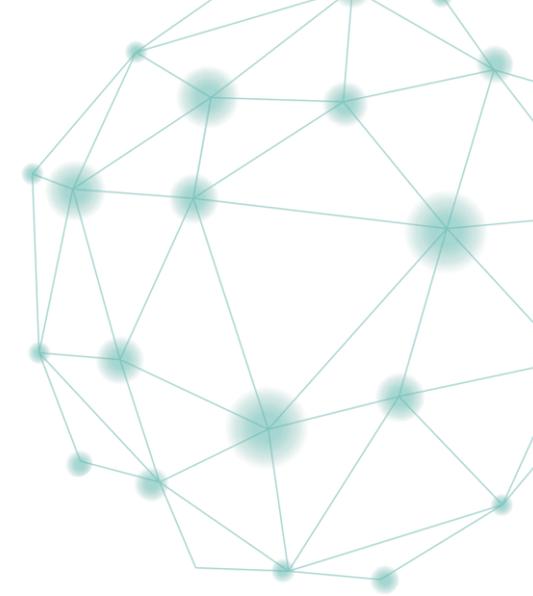




Objectifs

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

Faire de l'accompagnement un élément central du dispositif



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*



Les ateliers de la donnée s'installent progressivement sur tout le territoire



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

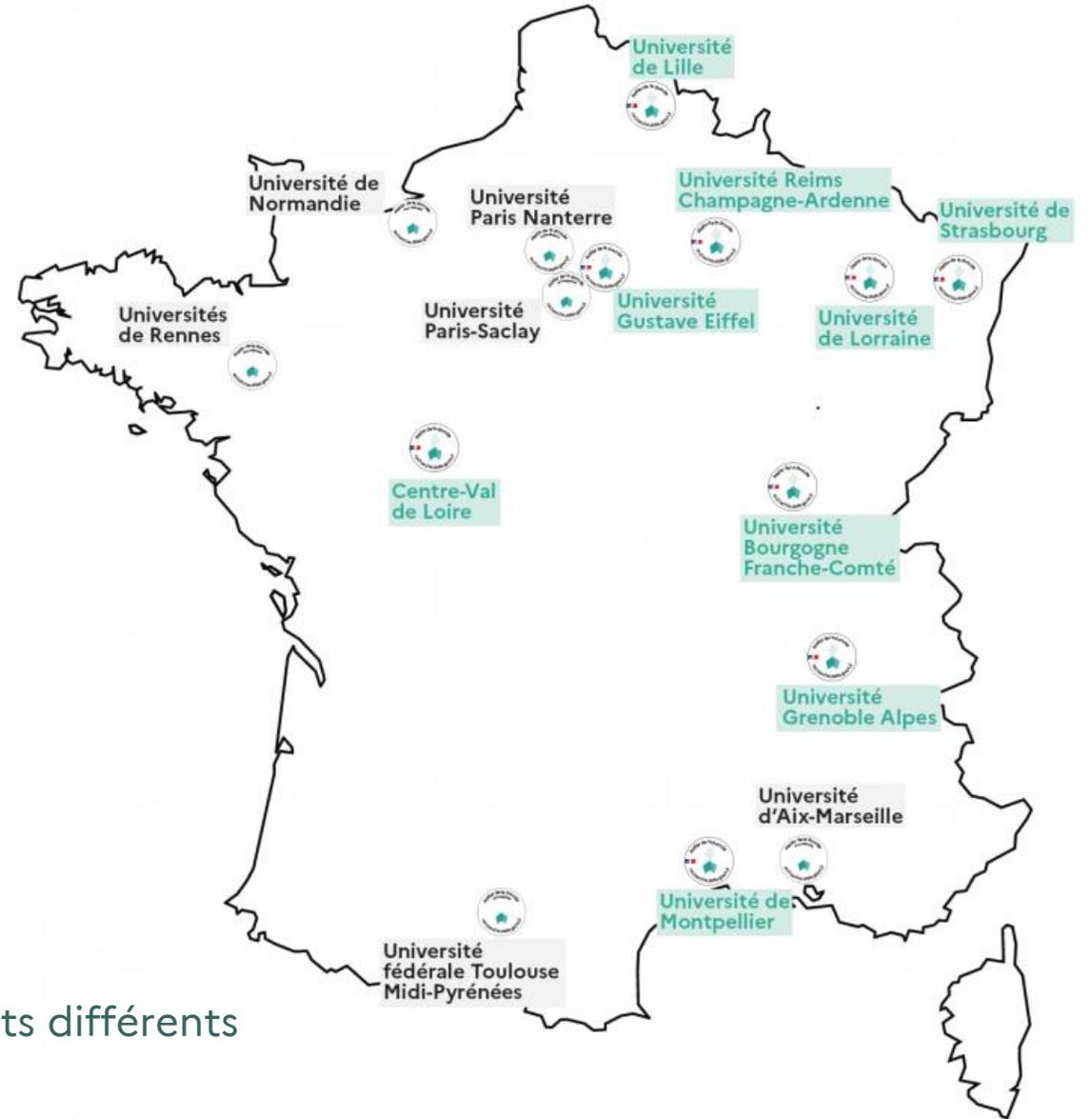
Regroupement d'expertises complémentaires de différents établissements

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

15 ateliers de la donnée opérationnels

10 ateliers en projet

85 ETP
342 personnes
73 établissements différents



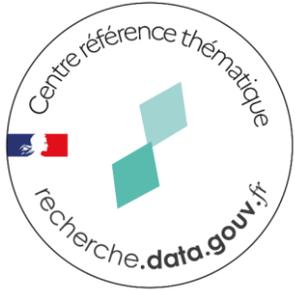


RÉPUBLIQUE
FRANÇAISE

Liberté
Égalité
Fraternité

recherche.data.gouv.fr

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

novembre 2022





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 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**



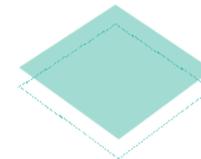
Centre de ressources compétences
Recherche Data Gouv

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

15 + ATELIERS DE LA DONNÉE

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



6 + CENTRES DE RÉFÉRENCE THÉMATIQUES

Expertise par domaine scientifique

4 CENTRES DE RÉFÉRENCE ÉTABLISSEMENT

Accompagnement propre aux orientations de l'établissement



4 CENTRES DE RESSOURCES

Pour soutenir les ateliers et capitaliser leurs pratiques

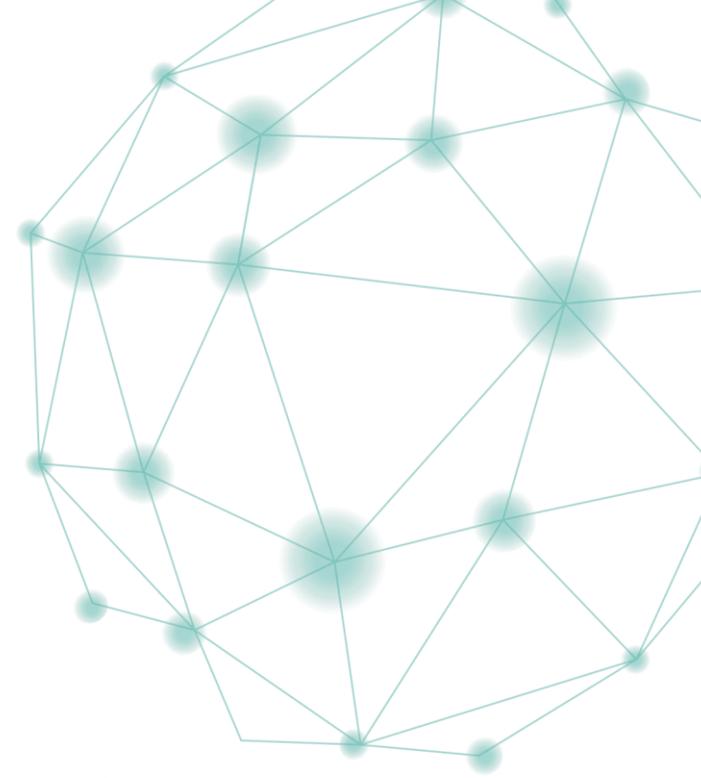
1 ENTREPÔT DE DONNÉES

Offre mutualisée pour tous les établissements pour le dépôt et la publication des données



1 CATALOGUE DES DONNÉES

Repérer et signaler les données des entrepôts externes de confiance



Une gouvernance collective

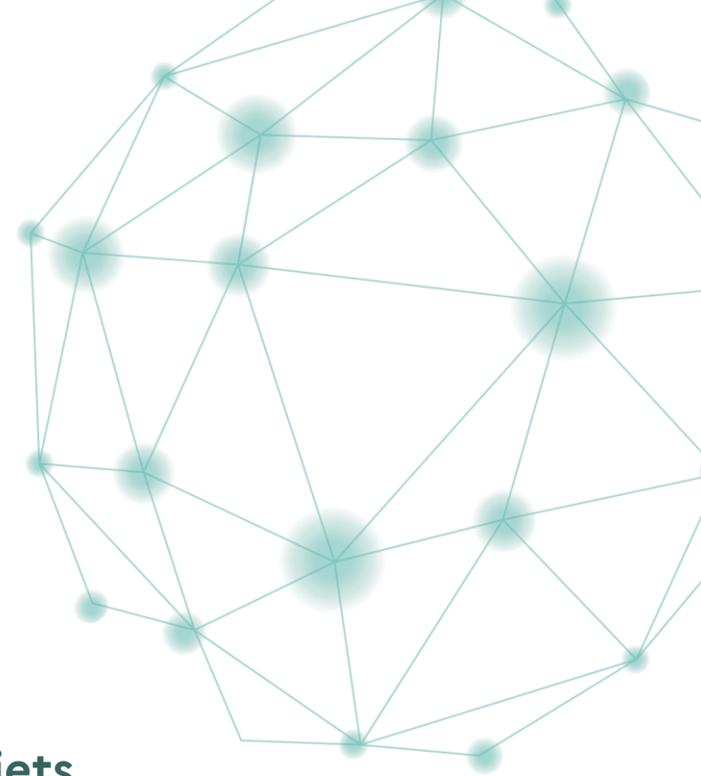
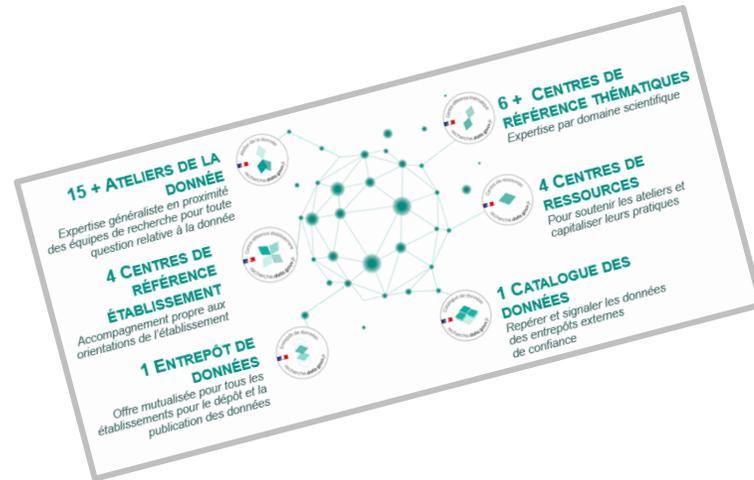
- Partenaires du projet



- Représentants de la communauté ESR



Diversité...



- Plus de 35 groupes d'acteurs différents
- 1200+ possibilités d'interactions
- Quelques centaines de personnes impliquées
- Des dizaines de sujets communs à traiter
- Des périmètres corrélés parfois en recouvrement
- Des tutelles différentes, porteuses de leur propre politique

Comment ça marche ?

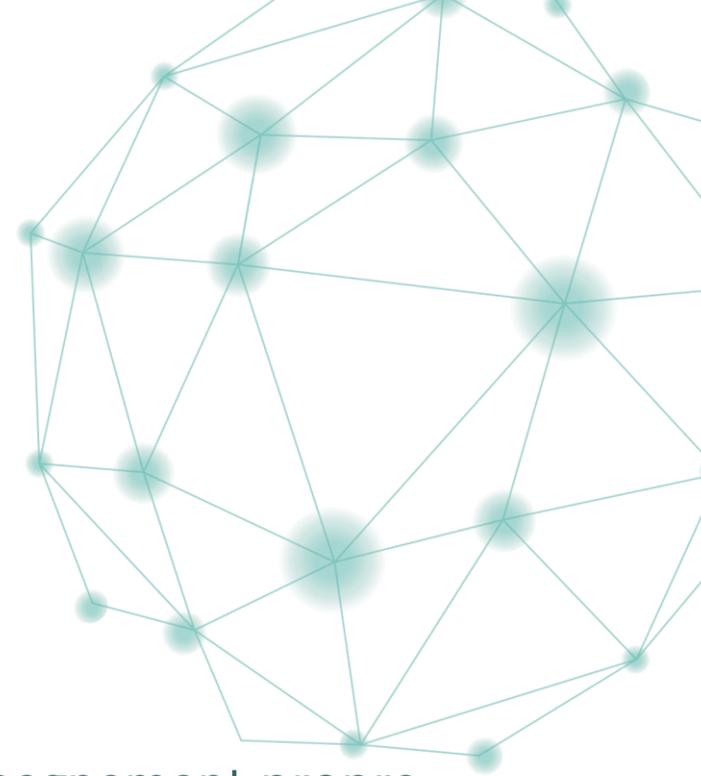


Exemple 1 : Intégration dans le paysage

La communauté ESR n'a pas attendu *Recherche Data Gouv* pour s'impliquer sur les données

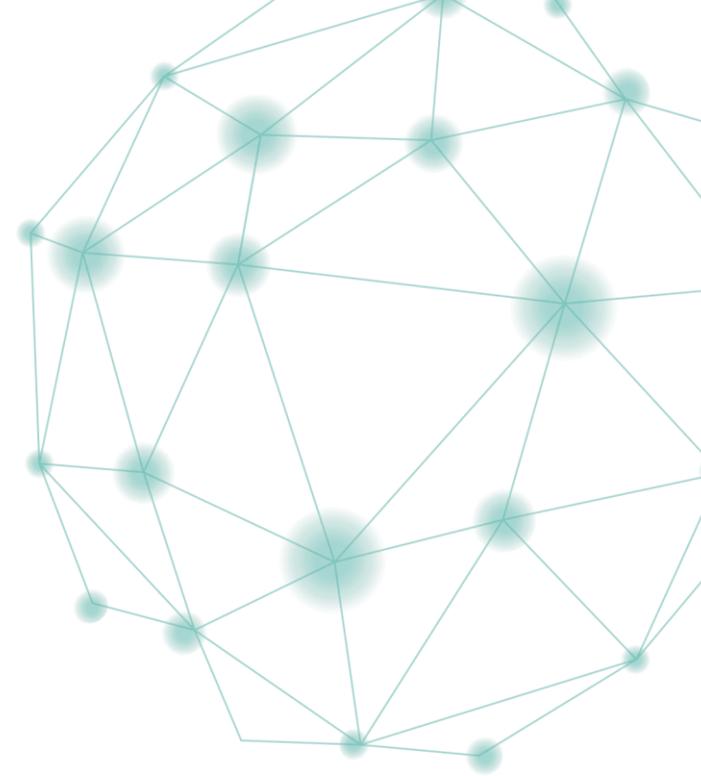
- **Centres de référence thématiques**
 - Infrastructures de Recherches proposant déjà des services
- **Centres de référence établissements**
 - Etablissement proposant depuis plusieurs années une offre d'accompagnement propre
- **Centres de ressources**
 - Structures existantes intervenant sur un domaine particulier
- **Principe de labellisation des ateliers de la donnée**
 - Reconnaissance d'un service déjà en place

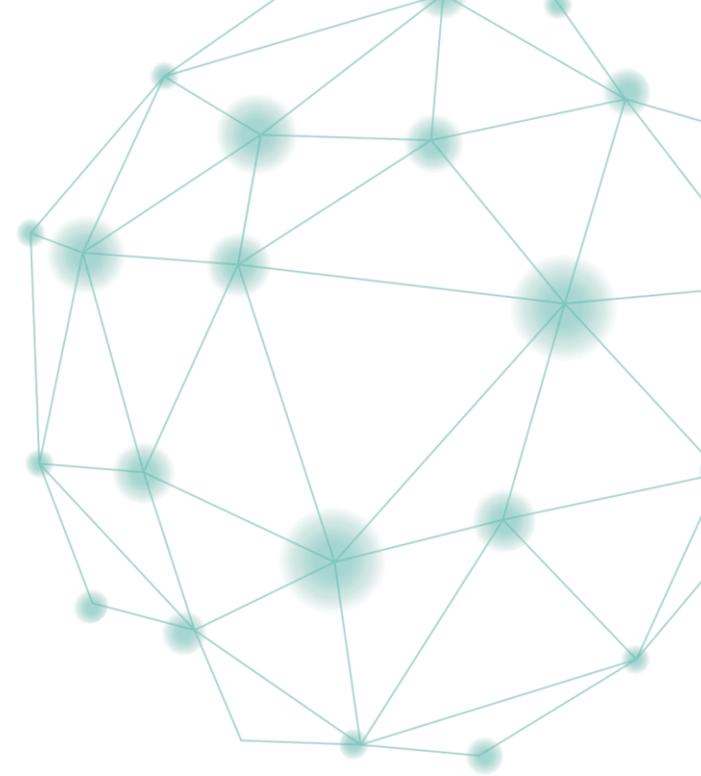
L'idée n'est pas de tout refaire, mais de permettre aux dispositifs existants de mieux fonctionner ensemble



Exemple 2 : le réseau des ateliers de la donnée

- **Les ateliers de la donnée**
 - Diversité de profils : chercheurs, enseignants, ingénieurs, bibliothécaires, etc.
 - Des compétences complémentaires
 - Taille très variable : de quelques personnes à plusieurs dizaines.
- **Réseau des ateliers**
 - Partager connaissances et compétences
 - Créer du lien et enrichir les interactions entre ateliers
 - Assurer une transmission uniforme des informations au sein des ateliers





Exemple 2 : le réseau des ateliers de la donnée

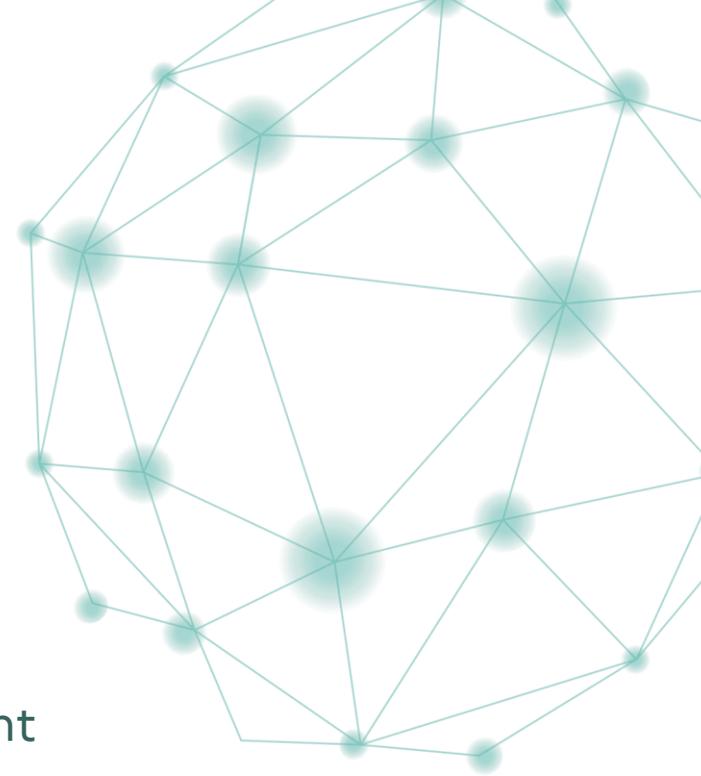
Les groupes de travail du réseau

- **Objectif** : mettre en commun les ressources développées par les ateliers de la donnée, et produire de nouvelles ressources utiles à l'ensemble des ateliers.
 - GT1 outils applicatifs communs pour le fonctionnement en réseau des ateliers
 - GT2 formation professionnelle des personnels des ateliers
 - GT3 boîte à outils et fiches pratiques à destination des chercheurs
 - GT4 mise en œuvre d'un réseau de correspondants données dans les laboratoires
 - GT5 accompagnement sur les codes et logiciels



Exemple 2 : le réseau des ateliers de la donnée

- **Le réseau est autonome**
 - Formé par les ateliers, géré par les ateliers
 - Piloté par un bureau composé de 3 membres qui changent régulièrement
- **Permet d'homogénéiser les interactions avec les autres dispositifs**
 - Centres de références
 - Centres de ressources
 - Autres acteurs et organes de pilotage





Exemple 3 : le module "portail – entrepôt – catalogue"

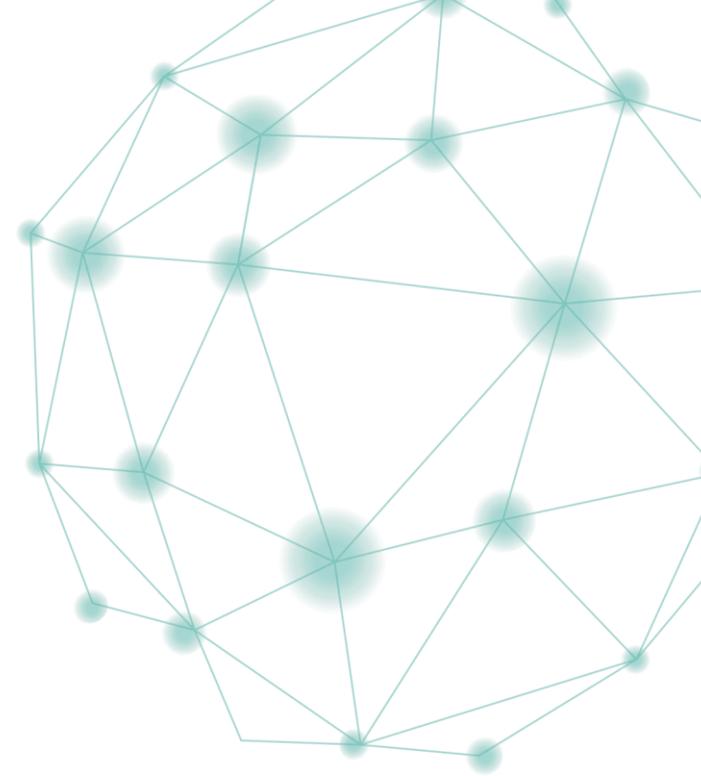
- **Une organisation en mode projet, avec une cible pérenne à terme**
 - 6 work packages, gestion agile
 - Pilotage par INRAE, capitalisant sur l'expertise de data.inrae.fr
 - Impliquant de nombreux partenaires
- **Un service global et mutualisé pour la communauté**
 - Les entrepôts institutionnels : gains financiers et RH pour les établissements
 - Permet à chacun d'appliquer ses règles de gestion
 - Une cohérence technique pour un environnement national
- **Masse critique pour atteindre une dimension internationale**



Exemple 3 : le module "portail – entrepôt – catalogue"

Quelques chiffres (en date de juillet 2023)

- **Audience cumulée du portail <https://recherche.data.gouv.fr/fr>**
 - 25 950 visiteurs uniques
 - 102 542 pages vues
- **Flux dans l'entrepôt (depuis l'ouverture début juillet 2022) :**
 - 30 espaces institutionnels actifs
 - 15 demandes d'espaces en cours d'instruction
 - 808 jeux de données déposés, 622 jeux de données moissonnés
 - 2 030 disponibles au total
 - 29 031 fichiers déposés ou moissonnés (36 000 disponibles)
 - 288 569 téléchargements de fichiers
- **Activité support (depuis l'ouverture début juillet 2022) :**
 - 30 formations organisées
 - 614 participants aux formations
 - 827 réponses par e-mail par l'équipe support

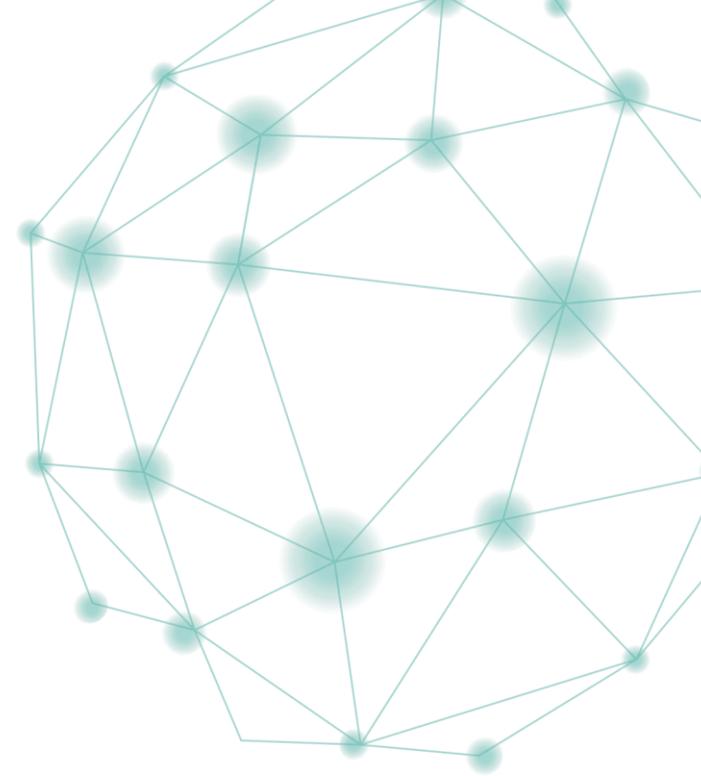




Comment ça marche ?

3 clés de réussite

- **Rassembler sans uniformiser, homogénéiser sans appauvrir**
 - Prendre en compte l'existant
 - Capitaliser sur les dispositifs en place
 - Privilégier le liant
- **Eviter les pyramides et les silos**
 - S'organiser de manière transversale
 - Travailler en réseau
 - Favoriser et valoriser les initiatives *bottom-up*
- **Devenir démonstrateur d'efficacité**
 - Rendre un service concret et mesurable
 - Axer la mutualisation sur les intérêts communs
 - Préférer l'agilité au monolithique





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/>



LICENCE OUVERTE
OPEN LICENCE



@RechercheDataGv