



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

recherche.data.gouv.fr



Recherche Data Gouv

Rapport d'activité 2020-2024

Sommaire

1.		
Aux origines de Recherche Data Gouv (2020-2021)	04	
Contexte et enjeux	05	
Phase d'étude et conception	06	
2.		
Mise en place et lancement : principaux jalons	08	
3.		
Organisation et gouvernance	10	
Comité de pilotage	10	
Structuration d'une gouvernance perenne	11	
Création d'une Unité Mixte de Service pour la plateforme	12	
4.		
Développement de la plateforme	14	
Portail web	14	
Entrepôt	15	
5.		
Réseau de centres de compétences	16	
Ateliers de la donnée	16	
Centres de référence thématique	19	
Centres de référence établissement	20	
Centre de ressources	20	
6.		
Mise en place du portail et de la communication	22	
7.		
Réalisation et impacts	26	
Bilan quantitatif	26	
Partenariats (inter)nationaux	26	
Événements, rencontres et présentations publiques	27	
8.		
Perspectives	30	

Éditorial



En juillet 2022, Recherche Data Gouv voyait officiellement le jour. Cette inauguration, point d'orgue d'un long travail de co-construction, marquait le passage d'une ambition à une réalité : offrir à la communauté scientifique un

soutien de proximité pour la gestion FAIR des données de recherche et un service national souverain et pérenne pour leur ouverture et leur partage.

Quatre ans plus tard, le chemin parcouru est considérable. Plus qu'un entrepôt ou un portail, Recherche Data Gouv s'est affirmé comme un dispositif fédérateur de centres de compétences, rassemblant plus de 400 personnels au sein de 81 établissements partenaires.

Ce rapport retrace les principales étapes de cette aventure collective : l'ambition, la genèse, les réalisations, les partenariats — mais aussi les défis à venir. Il témoigne de l'engagement de nombreux acteurs, au plus près

des équipes de recherche, dans de très nombreux établissements, pour faire du partage des données un levier de qualité, de transparence et de circulation des connaissances.

À l'heure où Recherche Data Gouv se prépare à franchir une nouvelle étape dans son développement, en s'inscrivant durablement dans le paysage de la recherche en France et en Europe, ce document se veut à la fois la mémoire d'un chemin parcouru et le socle sur lequel s'appuyer pour construire, ensemble, les prochaines avancées.

Isabelle Blanc

Directrice de Recherche Data Gouv



Aux origines de Recherche Data Gouv



Figure 1 : couverture du 1^{er} plan national pour la science ouverte.



Figure 2 : couverture du 2^{ème} plan national pour la science ouverte.

Contexte et enjeux

La mise en place de Recherche Data Gouv s'inscrit dans le cadre du développement de la science ouverte en France. Elle vise à répondre à des besoins concrets identifiés par les chercheurs, les institutions et les pouvoirs publics en matière de gestion, de diffusion et de valorisation des données de recherche.

Un besoin structurel d'un service national générique et souverain

La France ne disposait pas d'un entrepôt national générique permettant d'accueillir les données de recherche issues de disciplines ne disposant pas d'infrastructures adaptées. Ce manque représentait un frein majeur pour la diffusion des données, en particulier dans un contexte de montée en puissance des exigences d'ouverture. La création d'un service souverain, pérenne et de qualité constituait donc une nécessité stratégique.

Un projet porté par une forte volonté politique

Recherche Data Gouv s'inscrit dans les objectifs du 2^{ème} Plan national pour la science ouverte (2021-2024), notamment l'action visant à créer un service d'accueil et de diffusion des données. Ce projet répond également à une évolution réglementaire progressive qui tend à l'ouverture des données de recherche :

- Loi pour une République numérique (2016) ;
- Programme Horizon 2020 (2017) ;
- 1^{er} Plan national pour la science ouverte (2018) ;
- Obligation des plans de gestion des données par l'Agence nationale de la recherche (2019).

Il s'agissait ainsi de répondre aux demandes croissantes des financeurs en matière de transparence et de réutilisation des données scientifiques.

Une réponse aux besoins exprimés par les chercheurs et les établissements

Recherche Data Gouv est pensé pour répondre à plusieurs attentes

opérationnelles des acteurs de la recherche, à savoir :

- Accompagnement personnalisé : de nombreux utilisateurs, peu familiers avec le gestion et la diffusion des données, ont besoin d'un soutien humain, technique et juridique ;
- Simplicité du dépôt : les chercheurs expriment un besoin fort de solutions simples et peu chronophages pour la diffusion de leurs données, afin de se recentrer sur leur activité scientifique ;
- Valorisation et visibilité des données : la visibilité et la citabilité des données constituent un levier de reconnaissance scientifique. Le service doit assurer leur accessibilité dans des conditions favorables à leur réutilisation.

Optimisation du paysage national

Le paysage des entrepôts de données est marqué par la multiplication d'initiatives hétérogènes, souvent redondantes, et difficilement soutenables à long terme, ce qui génère une dispersion des ressources peu efficiente. Un entrepôt national permet de fédérer les efforts, de mutualiser les ressources et de proposer une alternative publique cohérente face aux plateformes privées internationales.

Qualité et interopérabilité

Recherche Data Gouv ambitionne de structurer et diffuser des données selon des standards de qualité élevés, notamment en suivant les principes FAIR (Findable, Accessible, Interoperable, Reusable), pour accroître la visibilité de la recherche française. Elle contribue à l'intégrité et à la réutilisation des travaux scientifiques.

Phase d'étude et conception

Un projet par et pour les communautés scientifiques

Le Collège des Données du comité pour la science ouverte a conduit une étude visant à évaluer les conditions de mise en place d'un service générique de diffusion des données de recherche.

Trois livrables en sont issus :

- **Analyse des besoins des utilisateurs¹** : une méthodologie participative a identifié les exigences des chercheurs (producteurs et réutilisateurs) et des gestionnaires. Les spécifications prioritaires comprennent une interface ergonomique, la saisie assistée des métadonnées via l'intégration de référentiels (ORCID, IdRef, AurÉHAL) et un accompagnement humain pour alléger la charge administrative liée à la gestion des données. Le rôle central de l'accompagnement a été mis en avant comme un facteur déterminant d'adhésion au service.
- **Comparaison internationale²** : l'analyse comparative de sept services nationaux (ARDC Australie, DANS Pays-Bas, UKDS Royaume-Uni, RADAR Allemagne, DFDR Canada, DNO Norvège, JORH Royaume-Uni) a documenté leurs modèles de gouvernance, fonctionnalités, organisation et financement. Les échanges directs avec leurs opérateurs ont confirmé des objectifs convergents : dépôt, archivage pérenne, accompagnement à la gestion des données, visibilité et accessibilité des jeux de données. L'étude souligne l'importance de la co-construction avec les acteurs clés et de l'intégration dans les réseaux nationaux et internationaux.
- **Scenarios de déploiement³** : trois modèles d'architecture ont été proposés, évaluant différents degrés de mutualisation, de faisabilité technique et de contraintes d'interopérabilité. L'analyse SWOT a privilégié le modèle national centralisé (comparable à DataverseNO) pour sa simplicité de mise en œuvre et la mutualisation des développements.

La dernière phase de l'étude a spécifié les rôles, les responsabilités des acteurs du

dispositif, les procédures de curation et les conditions de gouvernance pour garantir la qualité du service.

Le projet s'est appuyé sur une démarche participative, intégrant les retours des acteurs de l'écosystème scientifique pour ajuster la feuille de route, les fonctionnalités et l'organisation (espaces institutionnels, gestion des demandes).

Définitions des missions

Recherche Data Gouv a pour vocation de faciliter l'ouverture et la réutilisation des données scientifiques, de soutenir les chercheurs dans leurs pratiques de gestion des données, de garantir leur accessibilité et leur pérennité, et de promouvoir l'adoption des principes FAIR. Elle contribue ainsi à améliorer la visibilité, l'impact et la qualité de la recherche française, en offrant un socle commun robuste, aligné sur les standards internationaux.

Recherche Data Gouv repose sur une approche écosystémique unique, combinant une infrastructure technique de pointe et un réseau national de dispositifs d'accompagnement. À travers cet écosystème fédéré, Recherche Data Gouv propose une offre de services globale et mutualisée, à la fois disciplinaire, thématique et géographique. Cette structuration permet de répondre à des besoins hétérogènes, en particulier dans les disciplines ou les contextes encore peu couverts par les entrepôts de données spécialisés.

MISSIONS

- Faciliter l'ouverture, la découverte et la réutilisation des données scientifiques
- Accompagner les équipes de recherche dans la gestion et la diffusion de leurs données
- Garantir la qualité, l'accessibilité et la pérennité des données
- Promouvoir l'adoption des principes FAIR
- Renforcer la visibilité de la recherche française à l'échelle nationale et internationale

¹ Collège des Données de la recherche. *Étude de faisabilité d'un service générique d'accueil et de diffusion des données simples : recueil des besoins et des contraintes des usagers*. Comité pour la science ouverte. 2020, 62 p. DOI : 10.52949/13

² Hugo Catherine. *Étude comparative des services nationaux de données de recherche Facteurs de réussite*. Comité pour la science ouverte. 2021, 88 p. DOI : 10.52949/6

³ Collège des Données de la recherche. *Étude de faisabilité d'un service générique d'accueil et de diffusion des données simples : ambitions du service et scénarios de mise en œuvre*. Comité pour la science ouverte. 2021, 42 p. DOI : 10.52949/12

2

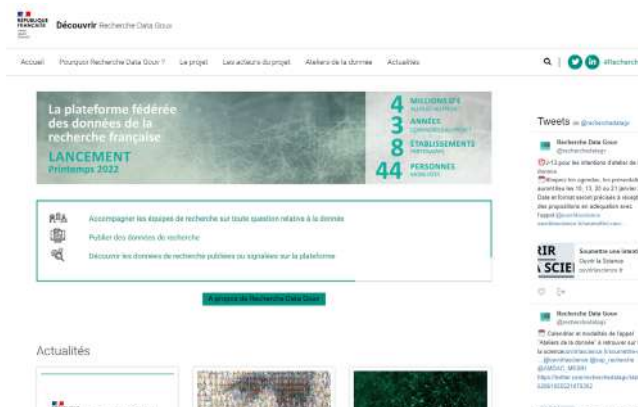
Mise en place et lancement : principaux jalons

Septembre 2021

Installation du comité de pilotage et de l'équipe multi-établissements pour le développement de l'entrepôt et du portail.

Décembre 2021

Validation de la mise à disposition d'offres de création d'espaces institutionnels pour tous les établissements, incluant des saisies anticipées pour les partenaires du projet.



Janvier 2022

Communication sur le projet Recherche Data Gouv. Lancement du 1^{er} AMI Ateliers de la donnée.



8 juillet 2022

Inauguration de Recherche Data Gouv :
portail recherche.data.gouv.fr,
entrepôt de données et
premiers centres de compétences.
Première rencontre des acteurs de
l'écosystème.

Printemps 2022

Lancement des premiers services
de Recherche Data Gouv.

3

Organisation et gouvernance

Comité de pilotage

Phase initiale

Créé pour une durée de trois ans dans le cadre du déploiement de Recherche Data Gouv, le comité de pilotage relevait initialement par délégation du comité de pilotage de la science ouverte du MESR. Chargé de l'orientation stratégique du projet, il joue un rôle central dans la définition et le suivi du projet, en veillant à la cohérence de l'écosystème des données de la recherche. Son rôle, sa composition et sa fréquence ont évolué au fil du temps, passant d'un mode projet à une structure de gouvernance plus pérenne.

Ses principales missions étaient les suivantes :

- Définir la feuille de route et les priorités du projet ;
- Proposer les orientations générales, les moyens à mobiliser et les solutions à mettre en œuvre ;
- Garantir la cohérence globale et le respect des objectifs fixés ;
- Contribuer à la définition du modèle structurel cible du dispositif ;
- Valider les objectifs et la stratégie de communication ;
- Assurer le lien avec le comité de pilotage de la science ouverte.

Le comité instruisait également les demandes de création d'espaces institutionnels et a délégué certaines décisions, comme la labellisation des

ateliers de la donnée ou la validation des demandes standards, à des instances dédiées.

Présidé par l'administratrice ministérielle des données, algorithmes et codes source, le comité de pilotage en mode projet se composait principalement de :

- Représentants du MESR ;
- Représentants des parties prenantes du projet (INRAE, CNRS, Universités Grenoble-Alpes, Lille, Lorraine, Paris, Paris Nanterre, Strasbourg) ;
- Représentants de UDICE ;
- Représentants de France Universités (Paris-Saclay, Toulouse Jean-Jaurès) ;
- Représentants d'organismes de recherche (CEA) ;
- Experts invités (HAL, Collège des Données de la recherche du Comité pour la science ouverte).

Le comité de pilotage s'est réuni mensuellement pendant la phase de développement de la plateforme (4 réunions de septembre à décembre 2021, 8 en 2022). À compter de 2023, les séances sont passées à un rythme bimestriel (6 réunions en 2023, 6 en 2024).



Découvrir les membres du comité de pilotage recherche.data.gouv.fr/fr/page/comite-de-pilotage

Structuration d'une gouvernance pérenne

Début 2024, une réflexion a été engagée sur la nécessaire évolution de la gouvernance. Cette démarche collective a été structurée autour d'ateliers de travail organisés au sein d'un comité de pilotage élargi pour l'occasion. En mai 2024, une proposition d'organisation et de gouvernance cible pour l'Infrastructure de Recherche Data Gouv a été validée, avec un impact direct sur la future composition du comité de pilotage.

L'écosystème Recherche Data Gouv a vocation à devenir une des infrastructures inscrites sur la feuille de route nationale des infrastructures de recherche. Celle-ci reposera sur cinq instances, qui coordonneront et orienteront les actions de la future infrastructure :

- Une **direction**, qui définira la stratégie globale et pilotera les actions de Recherche Data Gouv ;
- Un **comité de pilotage renouvelé**, qui validera la feuille de route, les

orientations politiques et budgétaires, assurera la cohérence du dispositif et supervisera la mise en œuvre des décisions ;

- Un **conseil scientifique international**, qui garantira la qualité et de la pertinence des orientations scientifiques de Recherche Data Gouv ;
- Un **conseil des membres**, qui rassemblera l'ensemble des établissements responsables d'un ou plusieurs services de l'écosystème au sein d'un conseil ouvert à tous les membres, afin d'assurer le partage de la stratégie globale, la coordination des actions et mesurer la satisfaction des usagers ;
- Un **comité des utilisateurs** représentera les besoins des utilisateurs de la plateforme des données en faisant remonter leurs attentes, difficultés et suggestions.

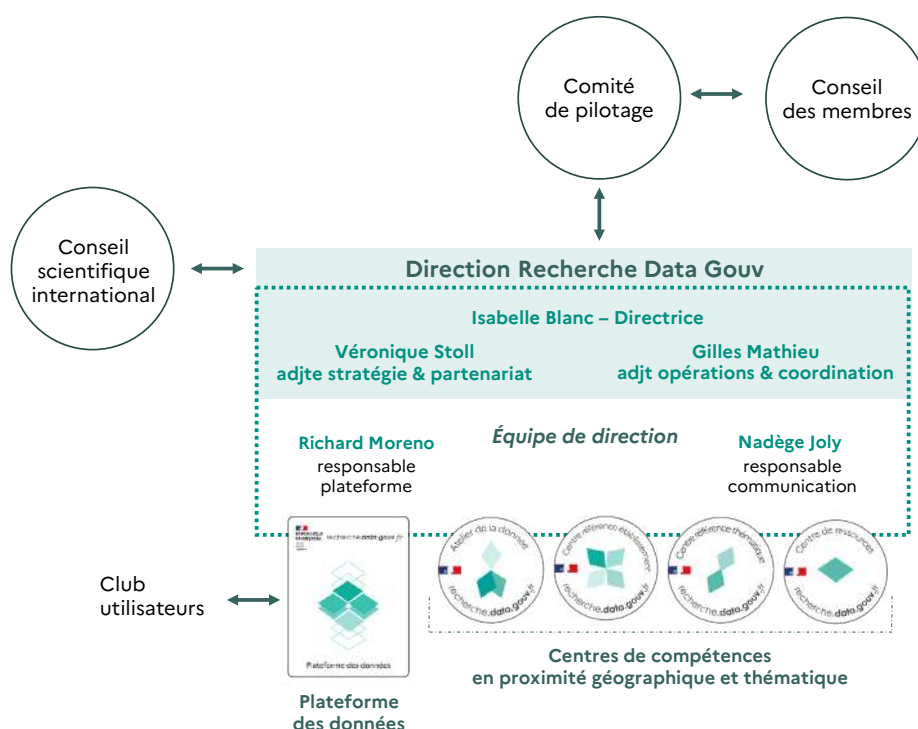


Figure 3 : schéma de la gouvernance Recherche Data Gouv (fin 2024).

Fin 2024, la direction de Recherche Data Gouv se compose d'Isabelle Blanc (directrice), de Véronique Stoll (directrice adjointe stratégie et partenariats), et de Gilles Mathieu (directeur-adjoint opérations et coordination).

Les responsables des unités fonctionnelles, Richard Moreno (responsable de la plateforme) et Nadège Joly (responsable de la communication) forment l'équipe de direction élargie de Recherche Data Gouv.

Enfin le collège de coordination rassemble

l'équipe de direction et le bureau des ateliers de la donnée. Il facilite le partage d'informations et assure une concertation globale.

Une structure pérenne pour la plateforme

Dans le cadre de la structuration de l'infrastructure Recherche Data Gouv, il a été décidé dès 2023 de créer une unité mixte de service (UMS) pour assurer l'exploitation de la plateforme nationale des données de recherche. INRAE, l'Université Grenoble Alpes, l'Université de Franche-Comté, l'Université de Lorraine et l'Université de Strasbourg en constituent les tutelles.

L'UMS regroupe le portail, l'entrepôt, le catalogue ainsi que le centre de ressources associé à la plateforme. Elle a pour mission d'assurer le fonctionnement opérationnel, l'évolution technique, le support et la coordination autour de la plateforme. Elle prend notamment en charge :

- Le pilotage opérationnel et la maintenance des services (portail,

entrepôt, catalogue) ;

- La gestion de l'infrastructure technique et applicative ;
- La veille technologique ;
- L'accompagnement, la formation et la communication autour des outils ;
- La curation de l'espace générique de l'entrepôt ;
- La coordination des administrateurs d'espaces institutionnels.

L'UMS assure également la gestion des ressources humaines et budgétaires nécessaires à ses missions.

Depuis le 1^{er} septembre 2024, la direction de l'UMS a été confiée à Richard Moreno. La convention constitutive de l'UMS sera signée en 2025.

En résumé, Recherche Data Gouv, c'est :

- Une **fédération nationale des 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 accessibles et réutilisables les données de recherche.



Figure 4 : les prémices de l'équipe Plateforme Recherche Data Gov, fin 2021.

4

Développement de la plateforme

Recherche Data Gouv repose sur une architecture technique avancée, structurée autour de deux services majeurs conçus pour assurer une gestion optimisée des données de la recherche et renforcer leur accessibilité : le portail et l'entrepôt.

Portail

Le portail constitue le point d'entrée de l'ensemble des services de Recherche Data Gouv : centres de compétences, entrepôt pour que les équipes de recherche puissent venir diffuser leurs données de recherche et interface de recherche tout public pour accéder aux données publiées.



recherche.data.gouv.fr



The screenshot shows the homepage of the Recherche Data Gouv portal. At the top left is the French Republic logo and the text 'recherche.data.gouv.fr'. To the right are links for 'Déposer / Publier un jeu de données' and a language selector set to 'FR - Français'. Below this is a search bar with the placeholder 'Rechercher sur le site' and a magnifying glass icon. A navigation menu includes 'À propos', 'Données de recherche', 'Centres de compétences', 'Plateforme de données', 'Aide en ligne', and 'Actualités'. The main banner features the text 'Un écosystème au service du partage et de l'ouverture des données de la recherche' and 'Fédérer, Accompagner, Partager, Ouvrir, Réutiliser'. At the bottom, there is a search bar with 'Rechercher un jeu de données', a blue button 'Trouver des données', and a link 'Déposer / Publier un jeu de données'.

Actualités

Entrepôt

Fondé sur le logiciel open source Dataverse, cet entrepôt mutualisé avec plus de 70 institutions propose un espace dédié aux communautés sans solution thématique spécifique. Il garantit une gestion sécurisée et centralisée des données.



entrepot.recherche.data.gouv.fr



The screenshot shows the homepage of the Recherche Data Gouv website. At the top, there is a navigation bar with links for 'Recherche', 'À propos', 'Guide d'utilisation', 'Support', 'Français', 'S'inscrire', and 'Se connecter'. Below the navigation bar, a yellow banner contains a welcome message and a link to the 'espace générique'. The main section features a large teal banner with the text 'Un écosystème au service du partage et de l'ouverture des données de la recherche' and the tagline 'FÉDÉRER, ACCOMPAGNER, PARTAGER, OUVRIR, RÉUTILISER'. Below this, there are sections for 'Recherche Data Gouv' and 'Génération datapaper'. A statistics section shows '1 014 797 téléchargements'. At the bottom, there is a paragraph describing the platform as a sovereign solution for sharing and opening research data, based on the Dataverse software.

EN CHIFFRES (2024)

83

Actualités

publiées entre l'ouverture du portail,
en juillet 2022 et décembre 2024.

194 K

Visites et contributions

sur Recherche Data Gouv en 2024
contre 124 K en 2023.

4,9 K

Jeux de données

déposés sur recherche.data.gouv.fr
en 2024 contre 3,4 K en 2023.

1,65 M

Données téléchargées

sur entrepot.recherche.data.gouv.fr
en 2024 contre 789 K en 2023.

5

Réseau de centres de compétences

La mise en œuvre des services d'accompagnement en gestion et diffusion des données s'appuie sur un réseau national de centres de compétences. Sous diverses formes (ateliers de la donnée, centres de référence thématiques...), ces structures apportent un soutien de proximité aux équipes de recherche, en tenant compte à la fois des spécificités géographiques et disciplinaires.

Ateliers de la donnée

Les ateliers de la donnée sont des structures d'accompagnement ayant pour vocation d'être le point d'entrée à destination des équipes de recherche pour toute question relative à la gestion et à la diffusion de leurs données. Ils agissent en proximité géographique, sur un site au périmètre défini, pour apporter aux équipes de recherche un premier niveau d'expertise.

Appels à labellisation « Ateliers de la donnée »

Depuis sa création, Recherche Data Gouv soutient la mise en place et le développement de ces ateliers au niveau national, au travers du label « Ateliers de la Donnée Recherche Data Gouv ». Cette labellisation répond à un cahier des charges précis et permet de reconnaître et valoriser la qualité des services et de l'accompagnement proposés. Elle est accordée depuis 2022 au moyen de campagnes nationales d'appels à labellisation permettant l'appréciation sur dossier des ateliers candidats, et sollicite

un comité d'évaluation composé d'experts indépendants. Ces appels à manifestation d'intérêt ont pour ambition de labelliser des dispositifs, d'accompagner dans leur réflexion les établissements pendant les phases de maturation et de préfiguration de leur projet, et de structurer sur le territoire le réseau des ateliers de la donnée. Ils visent à faire éclore et encourager des dispositifs qui se développeront progressivement dans la durée grâce à des compétences de personnels mises en commun entre établissements.

Évolution des ateliers de la donnée

Entre l'inauguration de Recherche Data Gouv en juillet 2022 et décembre 2024, le nombre de personnes impliquées dans l'opération et l'animation des ateliers a été quasiment multiplié par 8 (le nombre d'ETP correspondants a été, lui, multiplié par 5). Au-delà de l'augmentation du nombre d'ateliers qui



implique nécessairement une croissance du personnel, on note une stabilité des ressources au sein de chaque atelier, dont la plupart n'enregistre au cours du temps que des variations minimales.

L'offre de service des ateliers de la donnée s'est également étoffée et a gagné en maturité : de nombreux services déclarés comme « en projet » lors des dossiers de candidature à la labellisation sont enregistrés comme étant fonctionnels à la fin 2024.

Le réseau des ateliers de la donnée

Le réseau des ateliers de la donnée est composé de l'ensemble des ateliers de la donnée labellisés, en trajectoire de labellisation ou en projet. Ce réseau est animé par un bureau mis en place en octobre 2022, et dont les missions principales sont :

- Animer le réseau des ateliers de la donnée ;
- Favoriser l'échange de bonnes pratiques et d'expérience entre les différents membres du réseau ;
- Assurer la circulation de l'information entre le collège de coordination et le réseau des ateliers ;
- Représenter le réseau près de la gouvernance de Recherche Data Gouv.

D'octobre 2022 à décembre 2024, le bureau était composé de Stéphanie Castex (Université de Reims Champagne-Ardenne), Stéphanie Cheviron (Université de Strasbourg) et Alexis Arnaud (Université Grenoble Alpes).

Pour faire vivre ce réseau, mutualiser les efforts et favoriser une montée en compétences globale, plusieurs groupes de travail sont actifs fin 2024 :

- **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 interaction données et codes** ;
- **GT6 interaction entre les ateliers de la donnée et les centres de référence thématiques** ;
- **GT7 archivage pérenne des données scientifiques**.

Chaque groupe interagit avec l'ensemble des dispositifs d'accompagnement de l'écosystème Recherche Data Gouv, en particulier les centres de ressources et les centres de référence thématiques, ainsi qu'avec les acteurs présents sur ces questions, comme le GTSO Données de Couperin ou le Collège des Données du comité pour la science ouverte.

Les ressources développées sont mises à la disposition de tout l'écosystème. L'objectif est de fédérer les réflexions à l'ensemble des acteurs de la donnée sur le territoire national.



Les documents publiés et différents livrables du réseau des ateliers de la donnée sont disponibles sur la page <https://recherche.data.gouv.fr/fr/page/reseau-des-ateliers-de-la-donnee>.



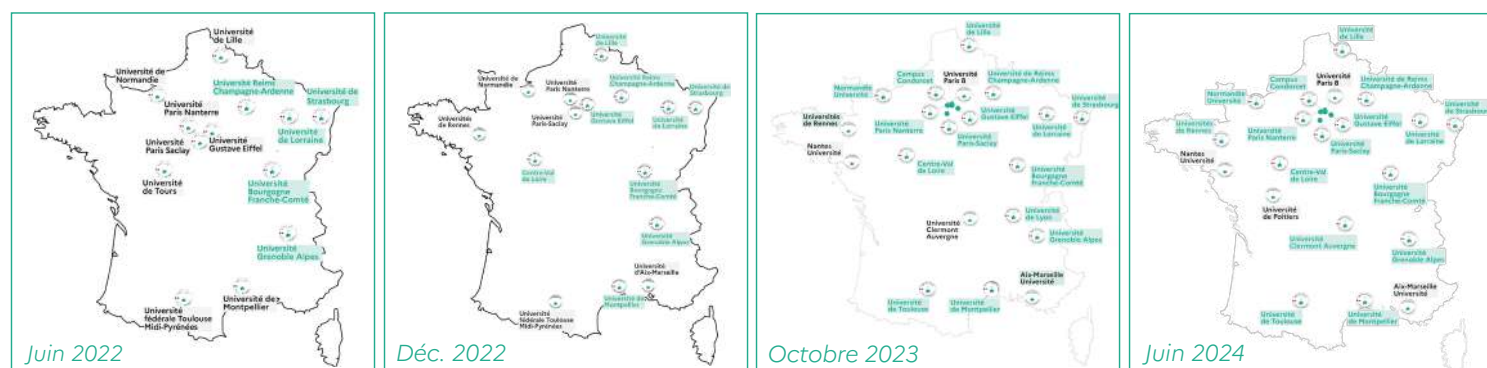
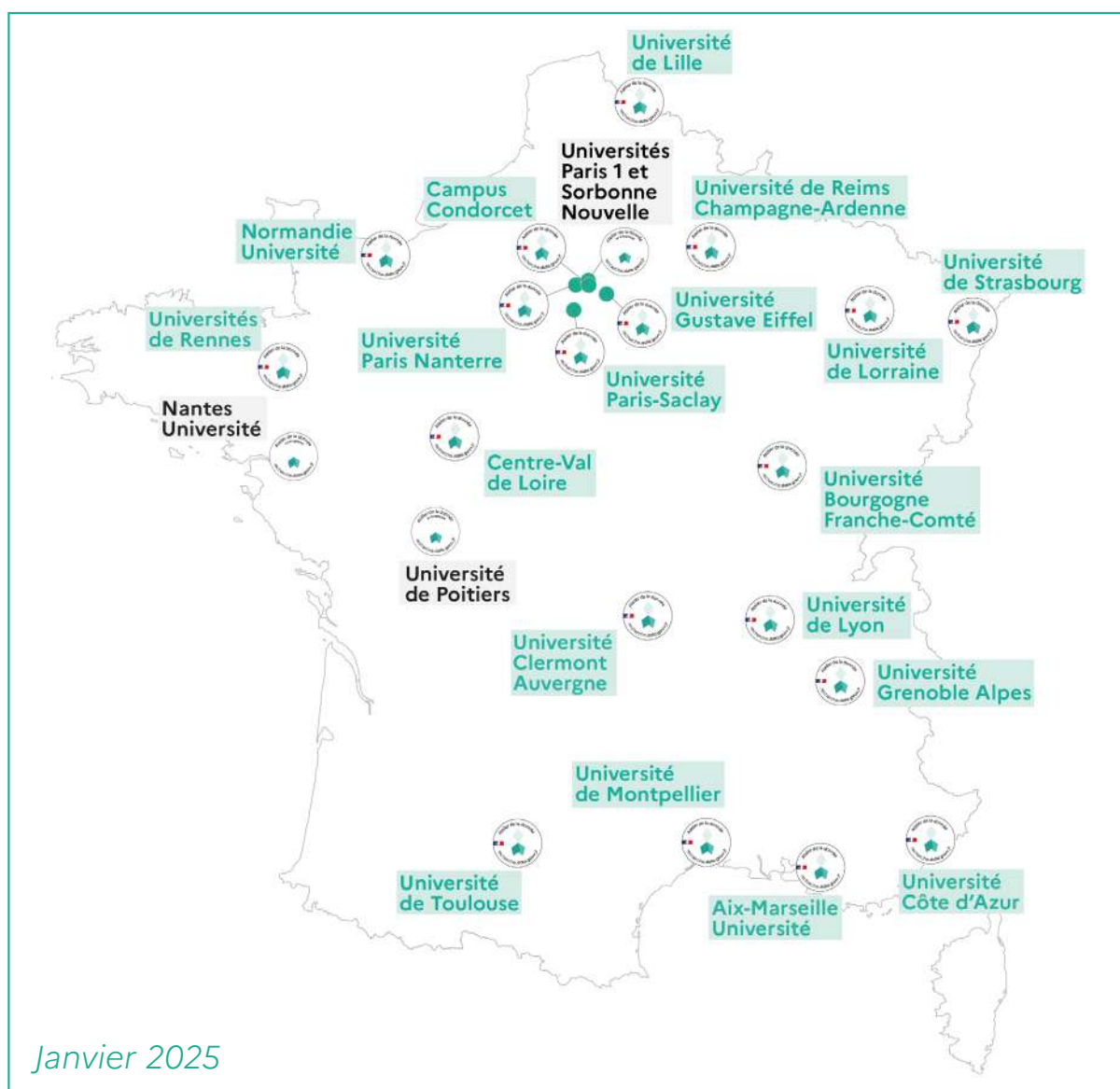


Figure 5 : rétrospective des cartes des ateliers de la donnée au fil des appels. (en vert : labellisés ; en noir : placés sur la trajectoire)

Dates de l'appel			Nombre d'ateliers candidats par appel		Nombre total d'ateliers post appel (retenus + ultérieurs)	
	Ouverture	Clôture	T*	L**	T	L
1	22/11/2021	28/02/2022	8	5	8	5
2	24/06/2022	14/09/2022	2	4	6	9
3	15/03/2023	10/07/2023	3	5	5	14
4	04/12/2023	29/03/2024	1	2	4	16
5	17/06/2024	25/10/2024	1	3	3	19

Figure 6 : appels réalisés entre 2021 et 2024 - *T : Ateliers de la donnée placés en trajectoire de labellisation. **L : Ateliers de la donnée labellisés.

	Nombre de personnes impliquées	ETP correspondants
Juin 2022	71	23,3
Déc. 2022	199	52,1
Sept. 2023	430	106,9
Juin 2024	441	110,5
Déc. 2024	557	128,6

Figure 7 : évolution des ressources humaines impliquées au sein des ateliers de la donnée.

Centres de référence thématiques

Les centres de référence thématiques ont été établis afin d'apporter une expertise propre à un domaine scientifique particulier.

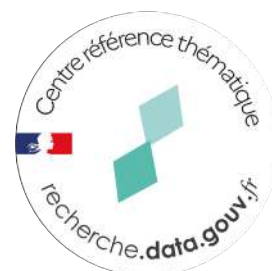
Ils sont au nombre de cinq et sont portés par :

- Institut Français de Bioinformatique (IFB) ;
- Data Terra ;
- Progedo ;
- Huma-Num ;
- Centre de Données astronomiques de Strasbourg (CDS).

Le centre de référence thématique porté par le Pôle National de Biodiversité (PNDB) a été fusionné en 2024 avec le Centre de référence thématique Data Terra, du fait de l'inclusion du PNDB comme l'un des pôles de Data Terra.

Ces centres de référence thématiques s'appuient sur des infrastructures de recherche nationales disposant d'une expertise reconnue et proposant à leurs communautés un accompagnement à la gestion FAIR des données.

Afin d'ouvrir l'expertise proposée à d'autres disciplines, une démarche d'élargissement est entreprise depuis 2024, visant à intégrer de nouveaux dispositifs. En particulier, cette dynamique est à l'œuvre dans le domaine de la biologie-santé, où plusieurs infrastructures nationales en biologie santé (INBS) s'associent à l'IFB pour proposer un point de référence dans les disciplines associées.



Centres de référence établissement

Lors de la première campagne de labellisation, quelques structures d'accompagnement ont été identifiées comme « centre de référence établissement ». Il s'agit de dispositifs proposant de longue date un service d'accompagnement au sein de leur établissement, et dont l'intégration dans le réseau des ateliers de la donnée de Recherche Data Gouv a été proposée suivant le principe de valorisation des structures existantes, l'un des principes fondateurs de l'écosystème.

Recherche Data Gouv compte 4 centres de référence établissement : École des Ponts – Paristech, INED, INRAE et SciencesPo.

Depuis leur création et jusqu'à la fin de la période concernée, les quatre centres de référence établissement ont contribué au réseau des ateliers de la donnée en participant aux différents groupes de travail.



Centres de ressources

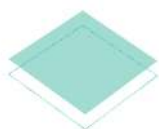
Les centres de ressources sont des dispositifs apportant un appui aux ateliers de la donnée et aux centres de référence thématiques. Construits autour de services existants, ils permettent une capitalisation à l'échelle nationale et aux bénéfices de tous.

Quatre centres de ressources ont été mis en place au démarrage de Recherche Data Gouv :

- Le centre de ressources « supports pédagogiques et e-formation », porté par la plateforme DoRANum de l'INIST du CNRS ;
- Le centre de ressources « outils communs aux ateliers », porté par l'équipe OPIDoR de l'INIST du CNRS ;

- Le centre de ressources « Montée en compétences des personnels des ateliers », porté par le GIS URFIST ;
- Le centre de ressources plateforme de Recherche Data Gouv, porté par l'équipe dédiée à l'opération et au support de la plateforme nationale des données.

En 2024, ces centres de ressources sont désignés par leur porteur. On parle désormais des centres de ressources **DoRANum**, **OPIDoR**, **URFIST** et centre de ressources **Plateforme Recherche Data Gouv** (CReP).



Centre de ressources
Plateforme
Recherche Data Gouv



Centre de ressources
outils communs aux ateliers
Recherche Data Gouv



Centre de ressources
Urfist
Recherche Data Gouv



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

Mise en place du portail et de la communication

Dès 2021, la communication a été placée au centre du projet, avec en objectifs majeurs :

- Faire connaître et informer autour de Recherche Data Gouv ;
- Construire, étoffer et faire vivre la communauté autour de Recherche Data Gouv ;
- Attirer davantage d'utilisateurs sur la plateforme et de candidatures pour les dispositifs d'accompagnement (ateliers de la donnée etc.) ;
- Accompagner les membres de l'écosystème dans leur communication en lien avec Recherche Data Gouv ;
- Être un point de contact, de support le cas échéant.

Instaurer une identité graphique

Le logo et l'univers graphique

Les premières étapes en termes de communication furent de penser une charte graphique de Recherche Data Gouv en cohérence avec :

- le système design de l'Etat ;
- la notion de branche sectorielle de Data Gouv.



Figure 8 : logo de Recherche Data Gouv qui reprend la Marianne à sa gauche et l'URL du portail à sa droite.

L'univers graphique de Recherche Data Gouv se compose

d'une palette de couleurs identitaires :

- Le vert menthe (issu du nuancier du système design de l'Etat) et ses atténuations ;
- Le noir et le blanc, pour un contraste optimal.



de graphismes limpides :

- Une sphère symbole de l'écosystème, fédérateur des acteurs ;
- Des losanges verts, représentant les données de recherche, sujet coeur de Recherche Data Gouv.



Valoriser l'écosystème

Le besoin d'identifier les différents modules et de développer le sentiment d'appartenance a provoqué la création de macarons. Ceux-ci permettent aux établissements de les apposer directement sur leur propre communication, sans règle précise tel un label.

Les macarons sont des ronds sobres, échos aux puits de données, qui portent la Marianne d'envergure nationale, renvoient vers le portail pour découvrir Recherche Data Gouv et se différencient via l'empilement spécifique des losanges pour chaque dispositif.

Dans un second temps, en complément

des macarons entrepôt et catalogue, un macaron rectangulaire de la Plateforme des données s'est ajouté à l'univers graphique en vue de marquer la notion d'outils, de plateforme à contrario des autres macarons distinguant des centres de compétences.

Enfin, des macarons ont été déclinés pour chaque centre de ressources.

La principale appropriation par les établissements hors intégration sur des affiches est l'autocollant.



L'ensemble des éléments graphiques se trouvent dans la charte graphique disponible sur le portail recherche.data.gouv.fr

Se doter d'outils

Afin d'accompagner les activités propres à Recherche Data Gouv, plusieurs outils de communication ont été créés en déclinaison de la charte graphique, notamment :

- Un motion design diffuser lors de l'inauguration ;
- Des capsules vidéos disponibles sur la chaîne Canal U ;
- Une brochure 3 volets pour ancrer les principaux éléments de discours ;
- Des kakemonos, pour en un regard identifier Recherche Data Gouv lors d'une manifestation physique ;
- Des gabarits de document texte, de présentation, etc. ;
- Des badges nominatifs pour s'identifier lors des rencontres ;
- Des bannières numériques pour de nombreux relais d'informations.



Figure 9 : de haut en bas, la brochure, une bannière et des autocollants Recherche Data Gouv.

Optimiser la présence numérique

La difficulté des premiers pas de Recherche Data Gouv était de diffuser au mieux et à chacun des informations sur les évolutions du projet, son calendrier et sa trajectoire en amont de l'inauguration. Par conséquent, un site web vitrine a offert un lieu numérique de convergence des informations et des actualités gravitant autour du projet à l'ensemble de la communauté. L'univers graphique étant sous embargo, seul le logo et un vert menthe ont été utilisés.



Découvrir le motion design créé en 2022 pour l'inauguration de Recherche Data Gouv
<https://www.canal-u.tv/chaines/rdg/recherche-data-gouv-un-ecosysteme-au-service-du-partage-et-de-l-ouverture-des-donnees>



En parallèle, le portail recherche.data.gouv.fr a été construit tel un liant entre chaque acteur de l'écosystème mais aussi et surtout comme une porte d'entrée vers la plateforme Recherche Data Gouv et vers les centres de compétences.

Aujourd'hui, divers acteurs interviennent sur le portail :

- Le pôle portail de la plateforme poursuit la gestion du développement et de la maintenance du portail ;
- Le comité éditorial fait vivre les contenus du portail selon la charte éditoriale relative ;
- Le centre de ressources Recherche Data Gouv alimente les tutoriels et autres ressources disponibles dans l'aide en ligne.

En 2024, plus de 194 000 visites sont recensées pour le site web recherche.data.gouv.fr, contre 124 000 en 2023, et 38 000 pour le second semestre de 2022 (inauguration en juillet 2022).

Recherche Data Gouv existe également sur les réseaux sociaux numériques, avec un souhait particulier à terme de développer sa présence sur Linked In, Bluesky et Mastodon. L'enjeu est d'instruire, de participer à la e-réputation de Recherche Data Gouv et surtout de faire communauté, d'encourager les acteurs de la recherche dans l'ouverture et le partage des données de recherche.

La force de Recherche Data Gouv se trouve aussi dans les acteurs de l'écosystème qui relaient dans leur propre communication (mailings, newsletters, réseaux sociaux).

plus prononcée d'un profil non francophone et d'une vision plus étendue de Recherche Data Gouv ;

- La conformité avec le Référentiel général d'amélioration de l'accessibilité.

Proposer une communication à long-terme

La communication de Recherche Data Gouv, désormais instaurée, se plonge dans des problématiques d'envergures différentes, notamment :

- L'expansion naturelle de l'écosystème et le défi de maintenir une communication grande échelle engageante, bienveillante et juste ;
- Le développement de la place de Recherche Data Gouv sur la scène européenne, et donc l'inclusion

Réalisations et impacts

Bilan quantitatif

La plateforme héberge fin 2024 environ 5 000 jeux de données, témoignant d'une forte adoption par la communauté scientifique. Ces données ont été téléchargées près de 2 millions de fois. Couvrant un large éventail de domaines — des sciences agricoles à l'ingénierie, en passant par les sciences humaines, la santé, les sciences naturelles, sociales et l'exploration de l'univers — cette diversité

illustre non seulement le rôle fédérateur de la plateforme à travers des disciplines variées, mais aussi sa capacité à favoriser des usages transversaux. Elle permet en effet la découverte et la réutilisation de jeux de données bien au-delà des champs de recherche d'origine des chercheurs, ouvrant la voie à de nouvelles perspectives interdisciplinaires.

Partenariats (inter)nationaux

À l'échelle internationale, Recherche Data Gouv joue un rôle clé en tant que service national dédié au partage des données de la recherche française. Elle a initié le groupe d'intérêt « National Data Services Interest Group » au sein de la Research Data Alliance (RDA), lancé officiellement à l'automne 2024. Ce groupe réunit les infrastructures nationales de données de plusieurs pays, tels que le Canada, l'Australie, l'Afrique du Sud, le Japon, l'Allemagne, la Finlande, les Pays-Bas, et la France via Recherche Data Gouv. L'objectif principal est de favoriser le partage d'expériences, l'échange de bonnes pratiques en matière de gestion et d'organisation des données, ainsi que de contribuer aux travaux de standardisation de la description d'un « service national de données », conformément aux recommandations du GORC Interest Group⁴.

Recherche Data Gouv est un acteur majeur au sein de l'EOSC, avec deux projets phares :

FIDELIS⁵ : en tant qu'entrepôt national de confiance, Recherche Data Gouv participe, via INRAE et le ministère chargé de l'enseignement supérieur et de la recherche, à la création d'un réseau européen d'entrepôts de données interopérables et sécurisés. Ce projet, financé par le programme Horizon Europe et lancé en janvier 2025, vise à renforcer l'écosystème européen pour le dépôt et l'ouverture des données de la recherche.

Skills4EOSC⁶ : en tant que centre de compétence national, Recherche Data Gouv contribue à la consolidation du soutien à la science ouverte en Europe. Labellisé centre de compétence d'EOSC depuis février 2024, Recherche Data Gouv met l'accent sur la formation et le partage d'expertise pour soutenir les équipes de recherche européennes dans la gestion des données.

⁴ The Global Open Research Commons International Model, <https://zenodo.org/records/10040098>

⁵ Site web de EOSC-EDEN et FIDELIS : <https://eden-fidelis.eu/>

⁶ Site web de Skills4EOSC : <https://www.skills4eosc.eu>

Événements, rencontres et présentations publiques



Figure 10 : inauguration de l'écosystème Recherche Data Gouv (juillet 2022).

Journée d'inauguration

La ministre de l'Enseignement supérieur et de la Recherche a inauguré, le 8 juillet 2022, Recherche Data Gouv, un écosystème au service du partage et de l'ouverture des données de recherche, devant 200 acteurs de la recherche à l'université Paris Cité.

Au côté de la naissance de l'écosystème Recherche Data Gouv, l'autre point d'orgue de l'événement fut l'ouverture de la plateforme, à savoir l'entrepôt et le portail, branche sectorielle de data.gouv.fr.

Cette inauguration a été précédée, le 7 juillet 2022, de la première rencontre de l'écosystème à la BULAC. Les objectifs de la journée étaient de découvrir la palette complète des acteurs de Recherche Data Gouv, de faire se rencontrer les membres de l'écosystème et de proposer des mises en situation variées d'ouverture et de partage de données de recherche.

Aussi, Recherche Data Gouv a été présenté à de nombreuses occasions et dans des contextes différents, que ce soit au niveau national (universités, réseaux métiers, journées science ouverte, événements nationaux ou locaux) ou international (conférences, projets européens, interventions invitées).



À lire : communiqué de presse du 8 juillet 2022, MESR, <https://www.enseignementsup-recherche.gouv.fr/fr/la-ministre-de-l-enseignement-superieur-et-de-la-recherche-sylvie-retailleau-inaugure-recherche-data-86206>



Figure 11 : première rencontre de l'écosystème Recherche Data Gov (juillet 2022).

Séminaires de l'écosystème

Depuis 2023, un séminaire réunit deux fois par an l'ensemble des acteurs de l'écosystème.

Ainsi, quatre séminaires ont été organisés sur la période 2023-2024.

Ces séminaires sont l'occasion de regrouper les contributeurs de l'écosystème autour d'un programme composé de sessions plénières et d'ateliers de travail. Ils permettent le partage d'informations et la réflexion commune, et renforcent la culture et l'identité de groupe de l'infrastructure.



**6-8 juin 2023,
Lyon**, 101 participants

<https://seminairerdg2023.sciencesconf.org/>



**14-15 décembre
2023, Reims**,
106 participants

<https://seminaire2rdg2023.sciencesconf.org/>





**4-6 juin 2024,
Poitiers, 135 participants**

<https://seminairerdg2024.sciencesconf.org/>



**20-22 novembre
2024, Nancy,
140 participants**

<https://rdg-seminaire4.sciencesconf.org/>



Perspectives

Évolution de la plateforme

Le catalogue de données : repose sur des technologies de big data et d'intelligence artificielle permet une recherche sémantique pour trouver des jeux de données sur la base de leur sens plutôt que par mots-clés stricts. Son déploiement en 2025.

Le service d'évaluation de la qualité de la curation : basé sur l'IA et exploitant des LLM, cet outil analyse la qualité des descriptions des jeux de données. Il fournit une aide aux déposants de données pour améliorer la qualité de leur dépôt au sein de l'entrepôt Recherche Data Gouv et d'autres entrepôts de l'ESR.

Le logiciel Dataverse sur lequel est basé l'entrepôt des données connaîtra une forte évolution début 2026 pour s'appuyer sur des principes architecturaux et des technologies à l'état de l'art.

Les moyens informatiques de la plateforme vont migrer d'INRAE vers des centres de données labellisés de l'ESR.

Une refonte du portail est prévue entre fin 2025 et début 2026.

Une phase d'analyse a démarré pour obtenir à terme une certification Core Trust Seal de l'entrepôt, pour l'ancrer encore plus dans une démarche de qualité reconnue par la communauté scientifique.

La plateforme est impliquée dans l'initiative AI Factory de la Commission Européenne, dans le cadre d'EuroHPC (European High-Performance Computing Joint Undertaking). Cette participation active place Recherche Data Gouv au cœur des futures AI Giga Factories européennes, comme le prévoit l'initiative InvestAI de la Commission, marquant ainsi son engagement dans l'écosystème européen de l'IA.

Candidature à la feuille de route nationale des infrastructures de recherche

Au cours du premier semestre 2025 s'ouvrira le processus de la mise à jour de la feuille de route nationale des infrastructures de recherche. Document stratégique, celle-ci identifie et soutient les grandes infrastructures et services scientifiques essentiels à la recherche française. Elle permet de renforcer la visibilité des infrastructures et favorise leur articulation avec les dynamiques européennes.

Recherche Data Gouv candidate au titre des infrastructures de recherche en information scientifique.

Renforcement des partenariats internationaux

À l'international, la stratégie de Recherche Data Gouv contribue au déploiement de l'EOSC (European Open Science Cloud). Recherche Data Gouv participe à la création d'un réseau européen d'entrepôts de données interoperables et sécurisés dans le cadre du projet Fidelis. Les équipes de recherche d'autres pays européens impliquées dans le réseau des centres de compétences de l'EOSC bénéficieront de formations et d'accompagnements communs fournis par Recherche Data Gouv. Cette extension renforcera la collaboration internationale, l'échange d'expériences et la montée en compétences collective dans la gestion des données de recherche. Recherche Data Gouv ambitionne ainsi de devenir un point de visibilité central pour les données de la recherche française en s'intégrant pleinement au sein de l'EOSC et d'EuroHPC.

