

L'accompagnement des données de la recherche au CNRS

Depuis [juin 2023](#), le CNRS dispose d'un [entrepôt institutionnel](#) pour partager les données de la recherche. Il s'adresse aux communautés scientifiques qui ne disposent pas d'un entrepôt thématique ou disciplinaire de confiance. L'ouverture de cet entrepôt s'inscrit dans les objectifs du [Plan « Données de la recherche »](#) du CNRS pour diffuser une culture des données FAIR - faciles à trouver, accessibles, interopérables et réutilisables -.

L'engagement du CNRS en faveur des données de la recherche

Le partage des données de la recherche constitue l'un des quatre piliers fondamentaux de la [feuille de route science ouverte](#) du CNRS, avec pour objectif de « développer une culture de la gestion et du partage des données chez tous les acteurs du cycle de vie de la donnée [...] basée sur la mise en œuvre de principe FAIR ».

En novembre 2020, le CNRS met en place un [Plan](#) dédié aux données de la recherche afin de structurer son action. « Ce plan s'inscrit dans la nécessité actuelle d'accélérer le développement vers la science ouverte, pour assurer la préservation et la réutilisation de ces données », déclare Alain Schuhl, directeur général délégué à la science du CNRS, dans une [interview](#) qui en présente les actions.

« Ce Plan s'inscrit dans la nécessité actuelle d'accélérer le développement vers la science ouverte, pour assurer la préservation et la réutilisation de ces données. »

Alain Schuhl

L'ambition du Plan Données se traduit en trois objectifs : dynamiser et diffuser une culture des données FAIR, faire connaître les services et outils existants, et accompagner l'évolution des pratiques et le développement de nouveaux services et outils. « Une bonne gestion des données, en respectant les principes d'ouverture et de partage, offre la possibilité à d'autres chercheurs et chercheuses de réutiliser des données existantes sans avoir à effectuer de nouvelles opérations de collectes ou de manipulations qui peuvent être longues et coûteuses », souligne Christine Hadrossek, chargée d'études sur les données de la recherche à la Direction des données ouvertes de la recherche (DDOR) du CNRS.

Dans le cadre de sa politique science ouverte, le CNRS fait partie des sept établissements partenaires réunis pour porter le développement et la maintenance de l'écosystème national [Recherche Data Gouv](#). À l'inauguration en juillet 2022, Antoine Petit, Président-directeur général du CNRS, a salué la collaboration exemplaire entre les équipes de l'enseignement supérieur et de la recherche (ESR) et INRAE qui ont réussi à créer la plateforme en moins d'un an.

Hébergé sur la plateforme Recherche Data Gouv, l'entrepôt [CNRS Research Data](#) propose une solution institutionnelle pour entreposer les données de la recherche pour les scientifiques du CNRS qui ne disposent pas d'un entrepôt thématique ou disciplinaire adapté à leur domaine de recherche. Un [logigramme « Où publier vos données »](#) est disponible pour aider le déposant à identifier où déposer son jeu de données.



Sensibilisation et formation à la gestion et au partage des données, des éléments essentiels de la politique de l'établissement

Dans le cadre du Plan données de la recherche, le CNRS mène « un important travail de sensibilisation et d'acculturation des scientifiques sur l'intérêt du partage des données et la promotion des solutions existantes », affirme Sylvie Rousset, directrice de la DDOR, dans une [interview](#) publiée en novembre 2022.

« Nous menons un important travail de sensibilisation et d'acculturation des scientifiques sur l'intérêt du partage des données et la promotion des solutions existantes. »

Sylvie Rousset

Cette sensibilisation s'appuie sur l'expérience et l'expertise des communautés pionnières dans la gestion et le partage des données de la recherche afin d'accompagner au mieux les autres disciplines sur la mise en place de bonnes pratiques. Certaines infrastructures de recherche, les plus avancées sur le sujet des données, constituent les premiers [centres de référence thématiques](#) de l'écosystème Recherche Data Govv.

Les [ateliers de la donnée](#), à proximité géographique des laboratoires et des équipes scientifiques, apportent une première expertise sur la gestion des données de la recherche. Le CNRS participe en tant que partenaire, aux diverses initiatives d'ateliers, qu'elles soient locales ou régionales.

Membre du Comité pour la science ouverte (CoSO), le CNRS s'implique également dans l'élaboration de guides destinés aux chercheurs, doctorants et personnels d'appuis. Disponibles en libre accès sur le site du CoSO « [Ouvrir la science](#) », ces guides partagent des bonnes pratiques, des outils et des ressources pour permettre de se familiariser aux différentes étapes de publication des résultats de la recherche.

Dans la lignée du Plan données de la recherche, l'Inist se mobilise pour accompagner les équipes de recherche et répondre aux besoins d'expertise et de formation sur les questions de FAIRisation et de partage des données.

La plateforme d'autoformation [DoRANum](#), développée en 2015 avec le [GIS « Réseau Urfist »](#), propose des ressources pédagogiques numériques autour de la gestion des données. Ce service permet à tous les utilisateurs d'adapter le niveau de formation en fonction de leurs besoins. DoRANum est labellisée [centre de ressources, supports pédagogiques, e-formations](#) au sein de l'écosystème Recherche Data Govv.

Par ailleurs, l'Inist met à disposition de la communauté de l'ESR le portail [OPIDoR](#) (optimisation du partage et de l'interopérabilité des données de la recherche) qui propose trois outils et services pour faciliter la gestion et la valorisation des données de la recherche.

DMP OPIDoR	Cat OPIDoR	PID OPIDoR
Planifier Outil d'aide à la création en ligne de plans de gestion de données	Repérer Espace wiki des services dédiés aux données de la recherche	Identifier Service d'attribution d'identifiants pérennes DOI



L'outil [DMP OPIDoR](#) dispose de modèles standardisés de plans de gestion de données, ainsi qu'une aide à la conception de modèles personnalisés. Depuis octobre 2023 le modèle de plan de gestion de données établi par le CNRS ([PGD CNRS](#)) est disponible sur l'outil DMP OPIDoR. Ce modèle comprend des recommandations propres à la politique de l'établissement en matière de science ouverte et de protection des données. Pour une harmonisation internationale de la gestion des données, PGD CNRS est basé sur le modèle de Science Europe, modèle également préconisé par certaines agences de financement comme l'ANR.

Sous forme d'un wiki sémantique, l'outil collaboratif [Cat OPIDoR](#) recense et décrit les services français qui contribuent à la gestion, au partage et la réutilisation des données scientifiques tout au long de leur cycle de vie. Depuis [juillet 2023](#), cet espace héberge le [catalogue CNRS des entrepôts et des services des données](#) dont le CNRS a la responsabilité ou auxquels il contribue en y affectant des moyens.

Porteur du consortium français membre de DataCite, l'Inist accompagne les scientifiques dans l'identification de leurs productions scientifiques. Le service [PID OPIDoR](#) contribue à l'attribution des identifiants pérennes (DOI) pour la publication des résultats de recherche : article, données, logiciels, etc.

Les réseaux technologiques transverses assurent une meilleure compréhension des métiers de la donnée

Pilotés par la Mission pour les initiatives transverses et interdisciplinaires ([MITI](#)) du CNRS, les réseaux métiers favorisent les échanges autour des pratiques et assurent une veille sur l'évolution des métiers et des compétences. Ils contribuent également à des actions de formation afin d'accompagner les personnels et les unités du CNRS tout au long de leur carrière.

La MITI soutient et pilote plus d'une vingtaine de réseaux nationaux et une soixantaine de réseaux régionaux associés.

Parmi eux, le [réseau Renatis](#) accompagne les scientifiques dans le domaine de l'information scientifique et technique (IST) et plus particulièrement en science ouverte. À travers les [C@fés Renatis](#) et les ateliers [DialoguIST](#), le réseau facilite les échanges et le partage d'expériences autour des métiers de l'IST et des actualités de la science ouverte.

Le groupe de travail inter-réseaux « [Atelier Données](#) », composé de représentants de 12 structures, s'attache à mener une réflexion sur la gestion des données porteuse d'une vision « métiers » et « réseaux ». En janvier 2023, le groupe de travail actualise le [Guide de bonnes pratiques sur la gestion des données de la recherche](#). « Ce guide est intéressant parce qu'il présente différents métiers et fonctions au travers des actions de formation menées par les réseaux, autour de la gestion des données. » précise Joanna Janik, chargée de mission compétences et formation à la DDOR. Des actions sont menées pour définir les compétences liées à la FAIRisation des données et mieux reconnaître la spécificité des métiers de gestion, de préservation et de diffusion des données ([data librarian](#), [data curator](#), [data steward](#), etc.).

Au niveau national, le CNRS participe au groupe de travail de l'[Étude sur les métiers de la donnée](#) pilotée par le ministère de l'Enseignement supérieur et de la Recherche. Cette étude vise à



comprendre les dynamiques en action et à qualifier les activités, les métiers et les compétences qui ont pris forme au sein des communautés françaises et internationales les plus avancées.

Le CNRS est également investit au niveau européen, dans le projet [Skills4EOSC](#) dont l'objectif est d'identifier les compétences métier et d'harmoniser le paysage actuel de formation, en science ouverte et en science des données, à l'échelle européenne.

