



Entrepôts : diffuser les données de la recherche

Science ouverte



Violaine Louvet, GRICAD

Administratrice des Données, Algorithmes et Codes de l'UGA

Journée « Gestion des données de recherche en SHS »

23 novembre 2021





Service en ligne permettant le dépôt, la description, la conservation, la recherche et la diffusion des jeux de données en vue de leur réutilisation.

A ne pas confondre avec des plateformes de stockage ou d'archivage.

Il existe des milliers d'entrepôts !

Différents types d'entrepôts de données

- Généralistes
- Thématiques
- Institutionnels
- Propres à un éditeur



- Des **entrepôts généralistes**

- › Zenodo
- › Dryad
- › Figshare
- › etc.



- Des **entrepôts institutionnels**

- › Data INRAE
- › Datasuds (IRD)
- › Data.sciencepo
- › Dorel (U Lorraine)
- › etc.



- Des **entrepôts thématiques**

- › Inter-university Consortium for Political and Social Research (ICPSR)
- › Nakala
- › Réseau Quetelet
- › Ortolang
- › etc.

- Des **entrepôts propres à un éditeur**

- › Mendeley Data
- › Gigascience / GigaDB
- › etc.



Pourquoi déposer ses données ?



- Répondre aux exigences des **financeurs**
- Augmenter la **visibilité** de ses recherches, rendre les données accessibles et citables indépendamment de l'article scientifique
- Garantir la **reproductibilité** des résultats de publications, en assurer la **transparence**
- Répondre aux exigences des **éditeurs**

Quel entrepôt choisir ?



- En premier lieu si il en existe, dans un **entrepôt disciplinaire**
- Recommandations du **financeur, de l'éditeur**
- Recommandations du **projet de recherche**
- Recommandations de l'**établissement ou l'organisme de rattachement**
- Utilisation d'un **annuaire** :
 - re3data (Registry of Research Data Repositories)
 - OAD (Open Acces Directory/Data repositories)
 - FAIRsharing (sciences de la vie et biomédecine)
 - OpenDOAR
 - ROAR
 - CoreTrustSeal (entrepôts certifiés)
 - etc.

- **Obligations / Recommandations**
 - Financeurs
 - Projet
 - Institution
 - Editeur
- **Statut de l'entrepôt**
 - Reconnaissance disciplinaire
 - Public/privé (pérennité)
 - Lieu d'hébergement des données
 - Certification
 - Modération des dépôts

- **Modalités de dépôt**

- › Types de données acceptés
- › Type de formats acceptés
- › Gestion des versions
- › Lien avec la publication
- › Durée d'accès - Préservation sur le long terme des données
- › Volume accepté
- › Coût du dépôt
- › Facilité du dépôt

- **Questions liées aux données**

- › Type d'accès possibles aux données
- › Embargo / restriction de l'accès
- › Licences
- › Modalités de description
- › Standards de métadonnées
- › Identifiant pérenne



- **Autres services**

- Moissonnage vers d'autres catalogues / entrepôts
- Aide au dépôt
- Simplicité d'utilisation
- Existence de statistiques d'utilisation, de consultation, de téléchargements

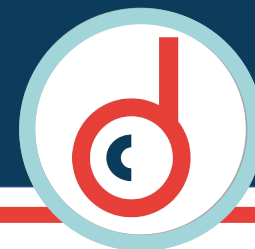
- Choix des **données pertinentes** (brutes, traitées, analysées)
 - Avec un point de vigilance concernant les données personnelles ou confidentielles
 - En s'assurant que les volumes sont en adéquation avec ce que permet l'entrepôt choisi
- Choix des **formats** (ouverts)
- Choix des **éléments de description** : métadonnées générales et disciplinaires, Readme ...
- **Organisation et nommage** des fichiers
- Choix de la **licence et des modalités d'accès** (ouvert, restreint, avec embargo)

- En France, la **loi pour une république numérique** prévoit une **liste de licences** (gérée par décret) qui peuvent être utilisées par les administrations pour la réutilisation à titre gratuit de leurs informations publiques
<https://www.data.gouv.fr/fr/pages/legal/licences/>
 - **Etalab 2.0**
 - **ODC Open Database License** (ODbL) version 1.0 (avec obligation de partage à l'identique)
- Compatibilité assurée avec les licences **Open Government Licence** (OGL, Royaume-Uni), **Open Data Commons Attribution** (ODC-BY, Open Knowledge Foundation)et **Creative Commons Attribution** (CC-BY, Creative Commons)

Où déposer ses données quand on est dans un laboratoire UGA ?



- Stratégie de l'établissement : proposer un **entrepôt institutionnel** pour répondre aux besoins des scientifiques qui n'ont pas de solution disciplinaire
- Un calendrier de mise en œuvre qui avait été prévu pour une entrée en production au 1^{er} semestre 2022
- Donc comparable au calendrier du projet national [recherche.data.gouv](https://recherche.data.gouv.fr/)
- Le choix a été fait de ne rien déployer en production en local mais de **rejoindre la plateforme nationale** dès son démarrage.



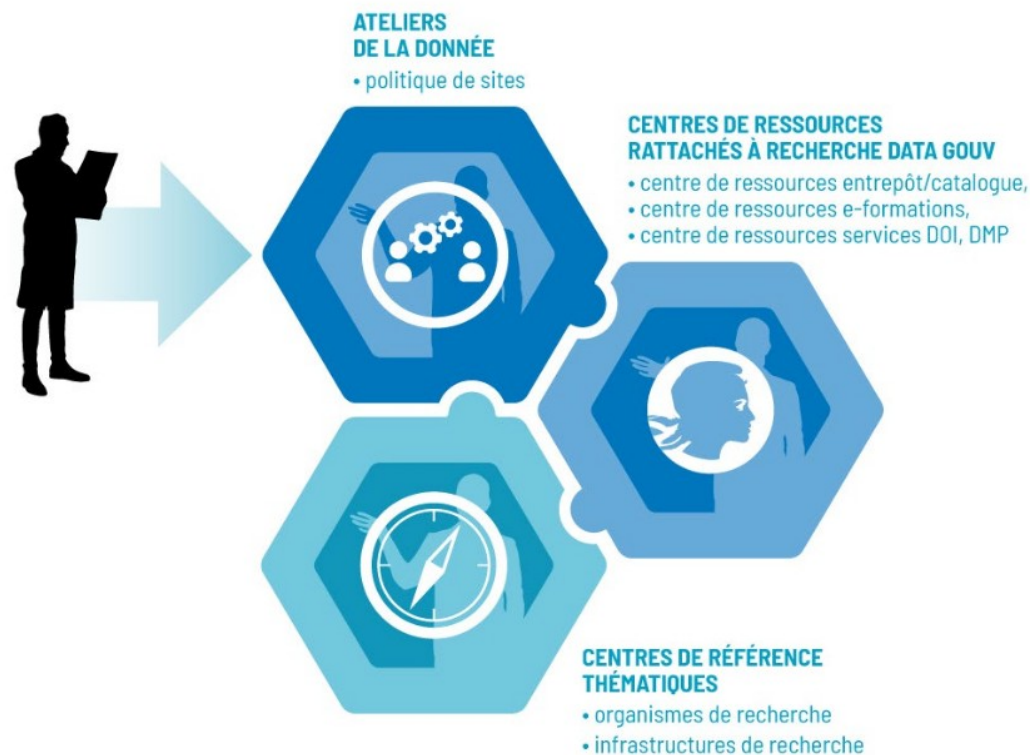
Une plateforme à cinq modules

RECHERCHE DATA GOUV



- Modules « **entrepôt - catalogue** »
- Projet confié par le MESRI à l'INRAE avec la participation d'établissements volontaires dont l'UGA
- Module « **ateliers de la donnée** »
- AMI en cours. Objectif de labelliser des initiatives d'accompagnement des communautés recherche sur l'ensemble du cycle de vie des données
- Réponse en cours de construction à l'UGA basée sur les actions de la Cellule Data Grenoble Alpes (CDGA)

- Première version de l'entrepôt en **mars 2022**
- Première vague de labellisation des ateliers de la donnée en **avril – mai 2022**
- Le dispositif d'accompagnement va aussi s'appuyer sur :
 - Des **centres de ressources thématiques**, spécialisés sur des aspects disciplinaires
 - Des **centres de ressources rattachés à Recherche Data Gouv** (techniques, mutualisation de formations, ...)



- En cours : travail sur l'organisation pour l'**accompagnement des dépôts** sur Recherche Data Gouv
 - Un groupe de travail composé de membres de laboratoires de différentes disciplines
- En cours : **construction de la réponse** pour labelliser la CDGA comme atelier de la donnée
 - Discussions à venir avec les référents données des laboratoires pour assurer l'adéquation de la réponse avec les besoins



- Sur le site science ouverte de l'UGA
 - <https://scienceouverte.univ-grenoble-alpes.fr/donnees/stocker/plateformes-locales/>
- Si vous avez des questions, contacter la cellule data :
 - sos-data@univ-grenoble-alpes.fr

The screenshot shows the 'Science ouverte' website header with navigation links: FAQ, Glossaire, Support, Contact, À propos, and Connexion. Below the header is a search bar and a menu with links to SCIENCE OUVERTE, PUBLICATIONS - HAL, DONNÉES, SERVICES, and ACTUALITÉS. The main content area is titled 'Cellule Data Grenoble Alpes et réseau des référents données'. It includes a description of the Cellule Data Grenoble Alpes, a list of competencies (Informatique, Information scientifique et technique, Juridiques), and a list of members (UAR GRICAD, bibliothèques universitaires, laboratoires, DIST CNRS). There is also a section for 'Prochaines Formations' with dates and times.

FAQ Glossaire Support Contact À propos Connexion

Science ouverte
Université Grenoble Alpes

SCIENCE OUVERTE PUBLICATIONS - HAL DONNÉES SERVICES ACTUALITÉS

Cellule Data Grenoble Alpes et réseau des référents données

Les données de la recherche > Cellule Data Grenoble Alpes et réseau des référents données

Cellule Data Grenoble Alpes

La Cellule Data Grenoble Alpes est une structure opérationnelle d'accompagnement des communautés scientifiques du site grenoblois sur toutes les problématiques liées aux données de la recherche.

Elle regroupe des compétences complémentaires afin de pouvoir répondre à tous les questionnements et déployer les outils nécessaires pour satisfaire les besoins des chercheurs dans le domaine des données de la recherche :

- Informatique
- Information scientifique et technique
- Juridiques

Elle est composée de membres de l'UAR GRICAD, des bibliothèques universitaires, de laboratoires et de la DIST CNRS.

Elle a pour mission de :

- Répondre aux questionnements des scientifiques. Ce site est également là pour vous aider.
- Animer, former et constituer une base de connaissances sur les thématiques liées aux données de la recherche
- Faire de la veille juridique et technique, et suivre les initiatives nationales, européennes et internationales

Déposer dans HAL

A La Une

Coup de toner sur la science ouverte (Université Toulouse Capitole)

Semaine Data-SHS 2021 ~ PUD Grenoble Alpes

Ma thèse en 180 secondes - Edition 2022

Prochaines Formations

NOV 9h00 - 18h00
15 Journée Science Ouverte - CNRS

NOV 14h00 - 15h30