

Déposer ses données sur Recherche Data Gouv



1

L'ÉCHO-SYSTÈME RECHERCHE DATA GOUV

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

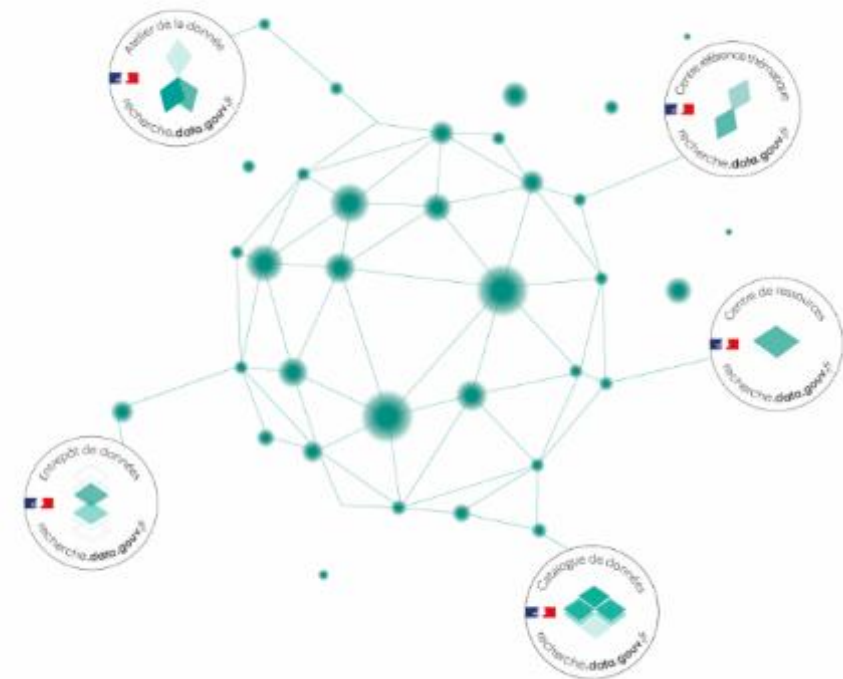
Plusieurs modules :

- [Des ateliers de la donnée](#)
- [Des centres de référence thématiques](#)
- [Des centres de ressources](#)
- [Des centres de référence établissement](#)

Deux modules pour déposer, publier et signaler des données :

- [Un entrepôt pour déposer et utiliser des données](#)
- Un catalogue pour rechercher les données publiées sur l'entrepôt ou sur des entrepôts externes

<https://recherche.data.gouv.fr/fr/page/a-propos-de-recherche-data-gouv>



2

CHOISIR SON ENTREPÔT DE CONFIANCE

Qu'est-ce qu'un jeu de données ?

« Agrégation, sous une forme lisible, de données brutes ou dérivées présentant une certaine "unité", rassemblées pour former un ensemble cohérent. »

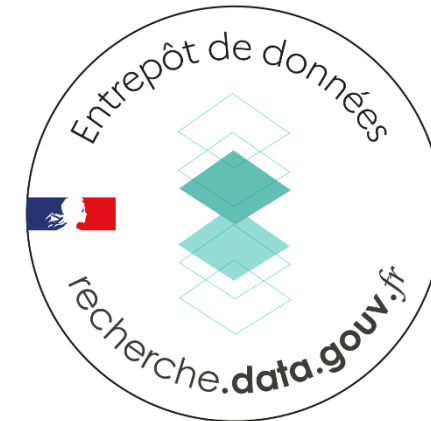
Gaillard R.- 2014. [De l'Open data à l'Open research data : quelle\(s\) politique\(s\) pour les données de recherche ?](#)

- **Ensemble cohérent, composé d'un ou plusieurs fichiers**
- **Décrit par des métadonnées**
- **Identifié par un DOI**
- **Comportant une licence (si possible ouverte), pour indiquer les possibilités de réutilisation**

Périmètre de l'entrepôt Recherche Data Gouv

Recherche Data Gouv :

- Est dédié aux données associées aux publications
- N'a pas vocation à se substituer au paysage des entrepôts disciplinaires existants. Il est destiné aux disciplines n'ayant pas encore d'entrepôts disciplinaires de confiance
- Accepte projets pluridisciplinaires
- N'accepte pas les données à caractère personnel



Trouver un entrepôt de confiance

Avant de déposer votre jeu de données il est donc indispensable de vérifier s'il n'existe pas un entrepôt de confiance dans votre discipline.

Comment trouver un entrepôt de confiance ?

Consultez la liste du ministère



Pour vous faciliter la tâche, le Comité pour la Science Ouverte maintient à jour une [liste](#). Il propose aussi une [méthode](#) pour identifier soi-même les entrepôts de confiance.

Contactez LabRADO



Nous pouvons vous accompagner pour trouver un entrepôt adapté à vos besoins.

labrador@sorbonne-universite.fr

Consultez vos pairs



C'est un bon moyen de trouver l'entrepôt utilisé par votre communauté. Mais veillez à ce qu'il soit bien de confiance (cf encadré ci-contre).

Les entrepôts certifiés



CoreTrustSeal certifie les entrepôts de confiance. Mais attention à leur localisation, surtout si vous avez des données personnelles.

Les critères d'un entrepôt de confiance :

- Effectuer une modération (vérification / validation) des jeux de données déposées
- Apporter des garanties sur sa pérennité
- Attribuer des identifiants pérennes aux jeux de données
- Ne pas pratiquer la cession de droit
- Ne pas pratiquer pas une politique tarifaire excessive
- Être localisé en UE si le type de données l'exige (données personnelles)

Si vous n'avez pas d'entrepôt de confiance dans votre discipline ou si votre projet est pluridisciplinaire, vous pouvez déposer vos données sur Recherche Data Gouv



<https://entrepot.recherche.data.gouv.fr/dataverse/sorbonne-univ>

Un entrepôt conforme aux principes FAIR



- Attribution d'un DOI Datacite
- Métadonnées riches conformes au standard DataCite
- Indexé dans Google, Be2Find, DataCite Commons...
- Les métadonnées au standard DataCite spécifient l'identifiant DOI du dataset



- Les données et/ou métadonnées sont accessibles via les protocoles https, OAI-PMH, et un ensemble d'API
- Accès restreint possible (à la demande, sous authentification, embargo...)
- Les métadonnées resteront accessibles via DataCite même les données ne le sont plus



- Les métadonnées sont disponibles dans de multiples standards (JSON, DublinCore...)
- Possibilité d'ajouter des URI de termes de vocabulaires externes
- Possibilité de lier des métadonnées à d'autres métadonnées (données et publication associées, ID projets, ORCID..)



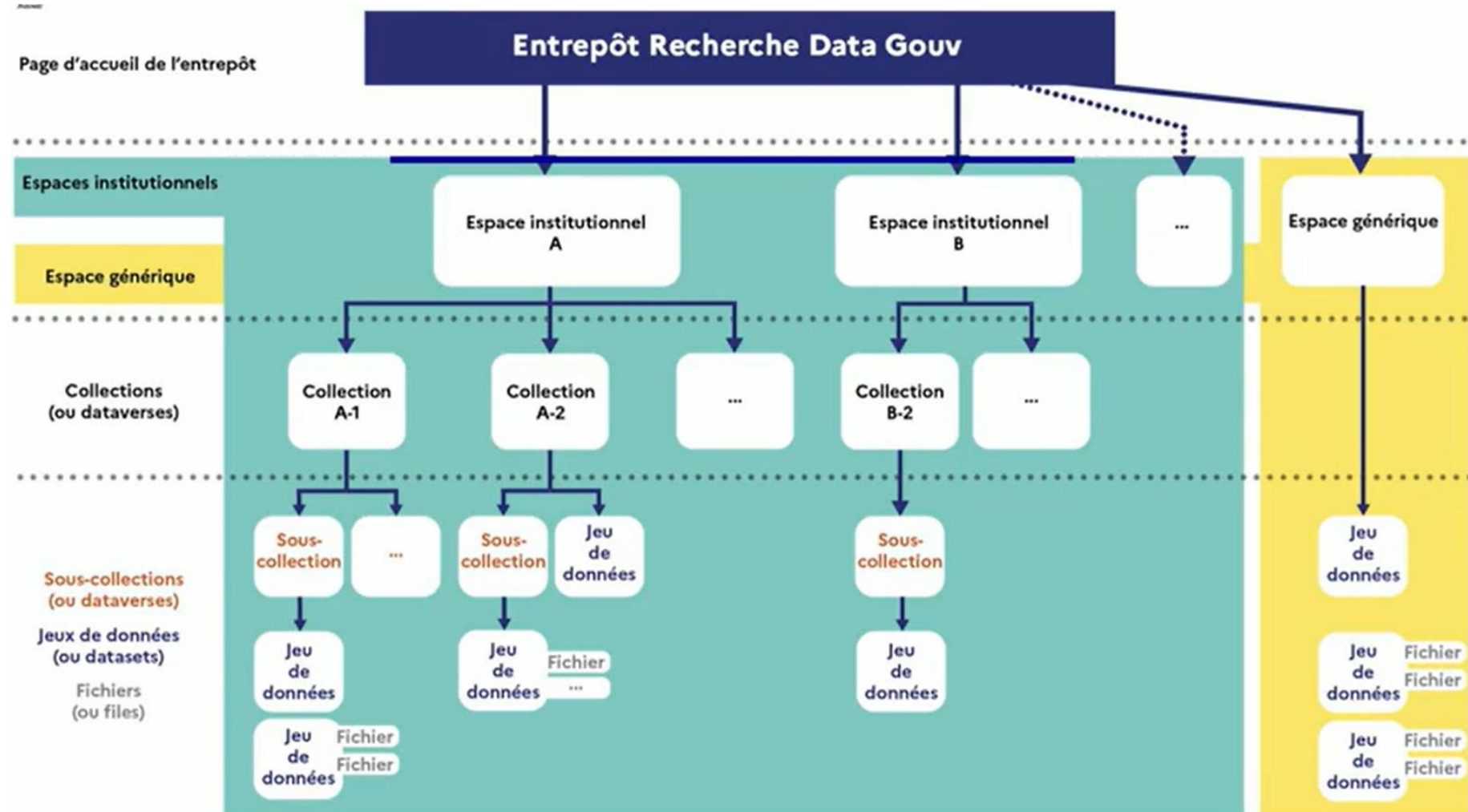
- Les métadonnées sont sous licence CC0
- Attribution de licences ouverte (par défaut Licence Ouverte Etalab 2,0)
- La provenance des données est documentée dans les métadonnées
- Au-delà du schéma de métadonnées DataCite, il est possible de préciser des métadonnées disciplinaires conformes aux standards des communautés

3

L'ENTREPÔT RECHERCHE DATA GOUV

Organisation de l'entrepôt

- Une solution basée sur le logiciel libre DataVerse
- Il n'existe pas de liens hiérarchiques entre l'espace générique et les espaces institutionnels
- Les espace institutionnels, collections et sous-collections sont appelés des dataverses
- Un jeu de données est déposé dans un seul espace. Il peut ensuite être lié à d'autres espaces pour y être affiché



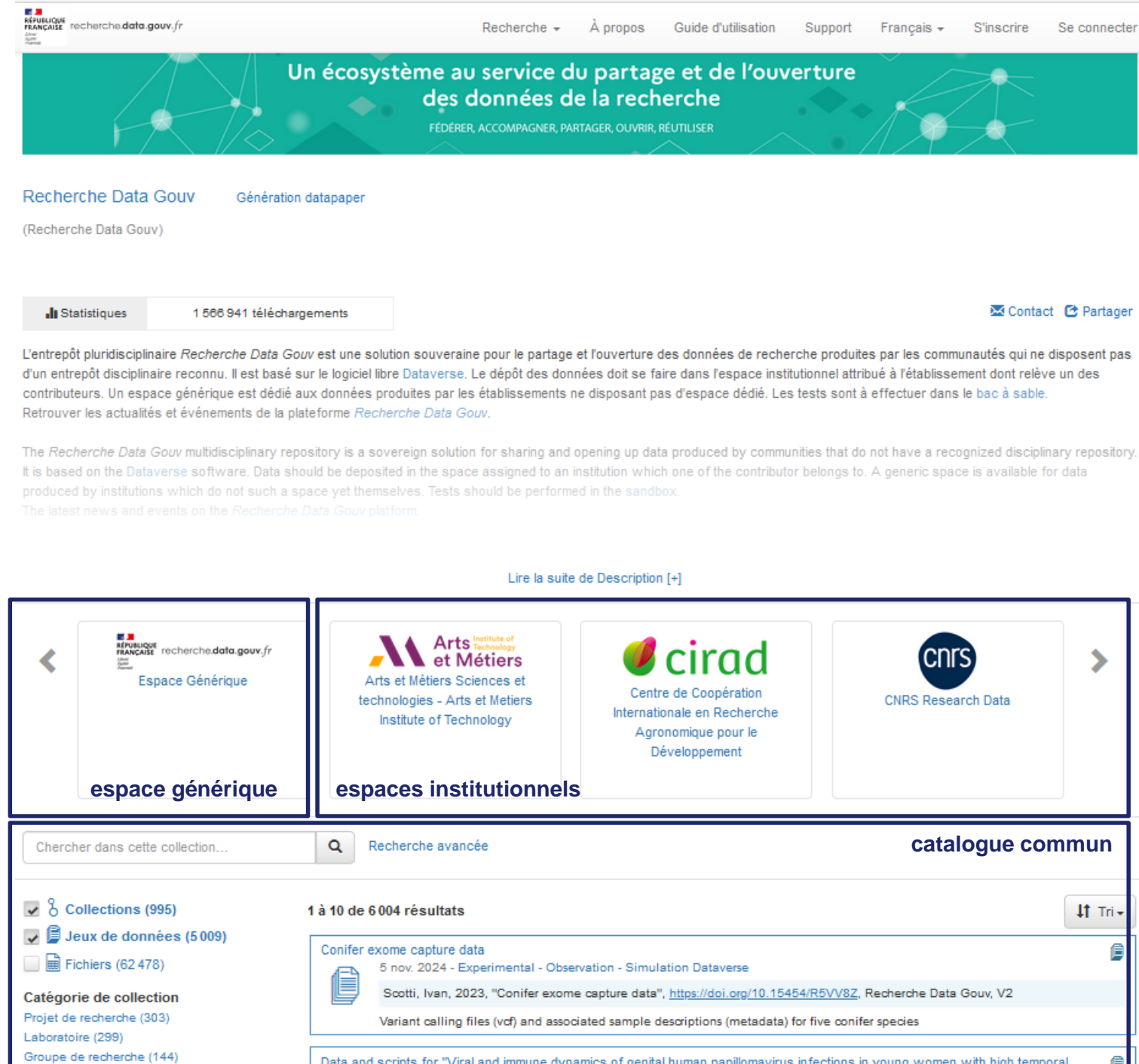
Organisation de l'entrepôt

La page d'accueil de l'entrepôt permet d'accéder aux différents espaces. Elle propose aussi un catalogue qui indexe l'espace générique et des espaces institutionnels.

À terme un véritable catalogue indexant l'entrepôt Recherche Data Gouv et d'autres entrepôts sera déployé.

Accéder à l'entrepôt de données Recherche Data Gouv :

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



The screenshot shows the homepage of the Recherche Data Gouv platform. At the top, there is a navigation bar with links for 'Recherche', 'À propos', 'Guide d'utilisation', 'Support', 'Français', 'S'inscrire', and 'Se connecter'. The main header features a 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 the banner, there are tabs for 'Recherche Data Gouv' and 'Génération datapaper'. A statistics bar shows '1 566 941 téléchargements'. The main content area includes a description of the platform as a multidisciplinary repository based on Dataverse, with instructions on data deposition and a link to the 'bac à sable' (sandbox) for testing. Below this, there are logos for various institutional spaces: 'Espace Générique', 'Arts et Métiers Sciences et technologies - Arts et Metiers Institute of Technology', 'cirad Centre de Coopération Internationale en Recherche Agronomique pour le Développement', and 'CNRS Research Data'. At the bottom, there is a search bar with the text 'Chercher dans cette collection...' and a 'Recherche avancée' button. The search results show '1 à 10 de 6004 résultats' and a list of collections, including 'Conifer exome capture data' with a detailed entry for a 2023 study by Scotti, Ivan.

L'espace institutionnel de Sorbonne Université



BIBLIOTHÈQUE
UNIVERSITAIRE

- Ouvert depuis décembre 2022
- Connexion possible via le CAS de Sorbonne Université
- Administré par la Cellule Données de la Bibliothèque de Sorbonne Université (data-bsu@sorbonne-universite.fr)
- Permet d'échanger avec une équipe de curation de Sorbonne Université qui maîtrise le contexte local
- Notre cellule peut également faire le lien avec ses partenaires locaux sur des problématiques plus générales sur la Science ouverte :
 - **LabRADoR** : pour les questions sur la gestion des données de recherche
 - **Le département Publications et Open Access** : pour les questions liées aux publications et à HAL

recherche.data.gouv.fr

Recherche - À propos - Guide d'utilisation - Support - Français - S'inscrire - Se connecter

SORBONNE UNIVERSITÉ

Sorbonne Université

Recherche Data Gov >

Contact Partager

Sorbonne Université est une université pluridisciplinaire de recherche intensive organisée en trois facultés : Lettres, Santé et Sciences & Ingénierie. Elle propose une recherche transversale et interdisciplinaire mobilisant des disciplines allant des humanités à la médecine en passant par les sciences technologiques, formelles ou naturelles.

Cet entrepôt permet aux chercheurs et chercheuses de Sorbonne Université de partager les données produites dans le cadre de leurs travaux afin de les valoriser et de permettre leur réutilisation.

Pour toute question relative à la gestion ou au partage de vos données de recherche, contactez-nous à l'adresse data-bsu@sorbonne-universite.fr

Chercher dans cette collection... Recherche avancée + Ajouter des données

1 à 10 de 18 résultats

26 août 2024

MARBACH, Sophie; ZHENG, Jeana Aojie; PINE, David J. 2024. "Replication data for Hopping and Crawling DNA Coated Colloids", <https://doi.org/10.57745/OCFXYZ>, Recherche Data Gov, V2

This data set contains particle trajectories for DNA-coated colloids diffusing above a DNA-coated surface, with the surface coated with 5% sticky fraction of DNA. Please see the ReadMe for more details.

6 août 2024 - Experimental - Observation - Simulation Dataverse

Morris, Cindy E.; Lacroix, Christelle; Chandeysson, Charlotte; Guilbaud, Caroline; Montell, Cécile; Piry, Sylvain; Rochelle Newall, Emma; Fiorini, Sarah; Van Gijsegem, Frédérique; Barny, Marie-Anne; Berge, Odile, 2023. "Population densities of the Pseudomonas syringae species complex in the Durance river in 2016-2017", <https://doi.org/10.57745/TB89SC>, Recherche Data Gov, V2

Accéder à l'espace institutionnel de Sorbonne Université sur Recherche Data Gov:

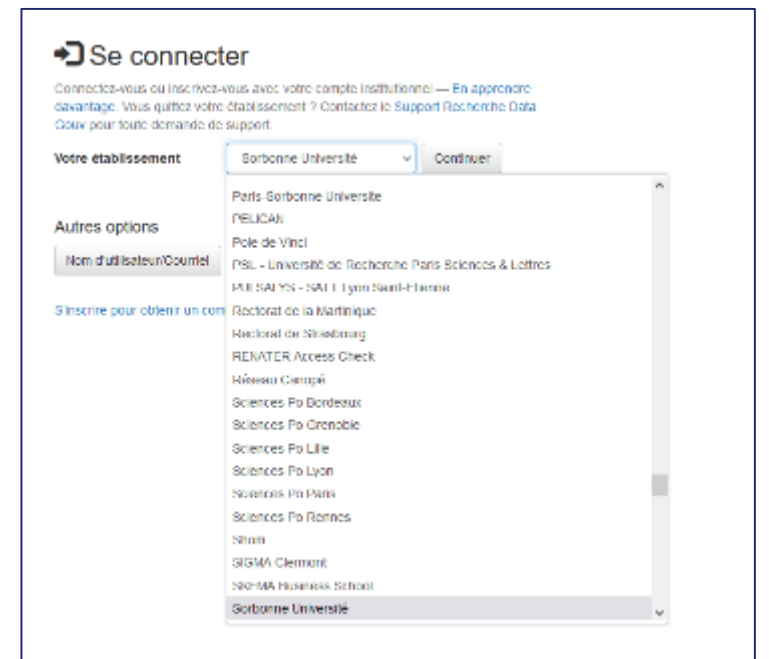
<https://entrepot.recherche.data.gouv.fr/dataverse/sorbonne-univ>

Se connecter avec votre compte Sorbonne Université



Recherche Data Gouv (Recherche Data Gouv) Génération datapaper

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



Se connecter

Connectez-vous ou inscrivez-vous avec votre compte institutionnel — En apprendre davantage. Vous quittez votre établissement ? Contactez le Support Recherche Data Gouv pour toute demande de support.

Votre établissement: Sorbonne Université

Autres options

S'inscrire pour obtenir un compte

- Paris Sorbonne Université
- PELICAN
- Pole de Vinci
- PSL - Université de Recherche Paris Sciences & Lettres
- PSL SAUVY - SAU 1 Lyon Saint-Herme
- Rectorat de la Martinique
- Rectorat de Strasbourg
- REMIAT Access Check
- Réseau Campus
- Sciences Po Bordeaux
- Sciences Po Grenoble
- Sciences Po Lille
- Sciences Po Lyon
- Sciences Po Paris
- Sciences Po Rennes
- Shim
- SIGMA Clermont
- SIRMA Business School
- Sorbonne Université



Bienvenue sur le portail d'authentification

Votre identifiant

Votre mot de passe

Se connecter avec son ORCID



RECHERCHE DATA GOUV

Recherche • À propos • Guide d'utilisation • Support • Français • S'inscrire • **Se connecter**

Un écosystème au service du partage et de l'ouverture des données de la recherche
FÉDÉRER, ACCOMPAGNER, PARTAGER, CLARIFIER, RÉUTILISER

Recherche Data Gouv (Recherche Data Gouv) Génération datapaper

Statistiques 243 257 téléchargements Contact Partager

L'entrepôt pluridisciplinaire Recherche Data Gouv est une solution souveraine pour le partage et l'ouverture des données de recherche produites par les communautés qui ne disposent pas d'un entrepôt disciplinaire reconnu. Il est basé sur le logiciel libre Dataverse. Le dépôt des données doit se faire dans l'espace institutionnel attribué à l'établissement dont relève un des contributeurs. Un espace générique est dédié aux données produites par les établissements ne disposant pas d'espace dédié. Les tests sont à effectuer dans le bac à sable.
Retrouver les actualités et événements de la plateforme Recherche Data Gouv

The Recherche Data Gouv multidisciplinary repository is a repository. It is based on the Dataverse software. Data should be available for data produced by institutions which do not have a dedicated space. The latest news and events on the Recherche Data Gouv platform

←  Espace Générique



Se connecter

Connectez-vous ou inscrivez-vous avec votre compte institutionnel — [En apprendre davantage](#). Vous quittez votre établissement ? Contactez le [Support Recherche Data Gouv](#) pour toute demande de support.

Votre établissement

[Me permettre de saisir le nom de mon établissement](#)

Autres options

[S'inscrire pour obtenir un compte Dataverse.](#)



Se connecter

Connectez-vous ou inscrivez-vous avec votre compte ORCID — [En apprendre davantage](#). Vous éprouvez des problèmes ? Veuillez contacter [Recherche Data Gouv](#) pour obtenir de l'aide.

ORCID est une initiative communautaire, ouverte et sans but lucratif visant à fournir un registre d'identifiants uniques de chercheurs ainsi qu'une méthode transparente pour lier les activités et produits de la recherche à ces identifiants. ORCID est unique dans sa capacité à atteindre les intervenants de toutes disciplines, secteurs de recherche et frontières nationales confondus ainsi que de coopérer avec d'autres systèmes d'identifiants. Pour en apprendre davantage visitez orcid.org/about.

Ce dépôt utilise votre ORCID ID pour l'authentification (vous n'avez donc pas besoin d'une autre combinaison nom d'utilisateur / mot de passe). Le fait que votre compte ORCID soit associé à vos datasets facilite également la recherche des datasets que vous avez publiés.

4

PRÉPARER SON JEU DE DONNÉES EN AMONT DU
DÉPÔT

Est-ce que je peux déposer dans Recherche Data Gouv ?

- **Est-ce que mon jeu de données a déjà été déposé dans un autre entrepôt fournissant un identifiant pérenne ?**
Si oui, je peux simplement le signaler dans Recherche Data Gouv en renseignant les métadonnées sans ajouter de fichiers et en ajoutant l'identifiant dans « Lien vers les données »
- **Est-ce que j'ai déjà déposé une version de ce jeu de données dans Recherche Data Gouv ?**
Si oui, je modifie simplement le jeu de données existant et je crée une nouvelle version du jeu de données
- **Est-ce que mon jeu de données contient des données personnelles, ou de santé ?**
Si oui, je ne peux pas le déposer dans Recherche Data Gouv
- **Est-ce qu'il existe un entrepôt de confiance dans ma discipline dans lequel je pourrais déposer mes données ?**
Si oui, je privilégie cet entrepôt (voir [slide](#))

Les ressources :

La liste des entrepôts de confiance

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

Guide « Où déposer et publier ses données » ?

<https://recherche.data.gouv.fr/fr/categorie/3/guide/ou-deposer-et-publier-ses-donnees>

Les formats

- Privilégiez les formats de fichiers ouverts :
 - ils correspondent à des fichiers encodés de façon transparente
 - leurs spécifications font partie du domaine public
 - ils sont interopérables
 - ils peuvent être créés, lus et modifiés par tous les logiciels destinés à traiter le type du fichier
- Si vous êtes contraints d'utiliser des formats propriétaires
 - Précisez bien les logiciels ou instruments desquels ils sont issus dans le fichier ReadMe ([voir slide](#)) et dans les métadonnées. Mentionnez de manière précise le nom, l'éditeur ou le fabricant et la version utilisée

Les ressources :

DORANum, Format ouvert ou fermé ?

https://doranum.fr/stockage-archivage/quiz-format-ouvert-ou-ferme_10_13143_mcwq-qs64/

FACILE, service de validation de format

<https://facile.cines.fr/>

Au-delà des format ouvert, cet outil identifie les formats compatible avec la conservation à long des données

JHOVE Open source file format
identification, validation & characterisation

<http://jhove.openpreservation.org/>

L'arborescence du jeu de données

- Adoptez une structure logique et cohérente et compréhensible par d'autres
- Documentez-là dans le fichier ReadMe ([voir slide](#))
- S'il existe un standard encadrant l'organisation des données dans votre discipline, utilisez-le (BIDS en neuro-imagerie, NeXus en physique ...)
- Il est souvent préférable de zipper son jeu de données dans une archive (.zip). Cela peut faciliter le dépôt et permettre de conserver votre arborescence de dossiers.
- A éviter : les doublons de fichiers; les arborescence trop profondes...

Les ressources :

Colomb, Julien, Thorsten Arendt, Deepti Mittal, et Keisuke Sehara. 2020. « Folder Structure Template for Research Repositories ». <https://doi.org/10.5281/zenodo.4410128>

Towards a Standardized Research Folder Structure – GenR. 2021. <https://web.archive.org/web/20230425045409/https://genr.eu/wp/towards-a-standardized-research-folder-structure/>

Archifiltre, est un logiciel de gestion documentaire qui permet de visualiser vos arborescences de fichiers et d'identifier les doublons
<https://archifiltre.fabrique.social.gouv.fr/archifiltre-doc/>

```
/project_name/  
├── raw_data/  
│   ├── satellite/  
│   │   ├── 2024_01_01_satellite1.nc  
│   │   └── 2024_01_02_satellite2.nc  
│   └── field_observations/  
│       ├── 2024_01_01_siteA.csv  
│       └── 2024_01_02_siteB.csv  
├── processed_data/  
│   ├── analysis1/  
│   │   └── 2024_01_01_analysis1.nc  
│   └── analysis2/  
│       └── 2024_01_01_analysis2.nc  
├── documentation/  
│   ├── project_description.pdf  
│   └── data_dictionary.xlsx  
├── scripts/  
│   ├── preprocess_data.py  
│   └── analyze_data.R  
└── results/  
    ├── figures/  
    │   ├── figure1.png  
    │   └── figure2.png  
    └── tables/  
        ├── table1.csv  
        └── table2.csv
```

Le nommage des fichiers

- Nommez vos fichiers de manière explicite, structurée et cohérente
- Adoptez une convention de nommage et mentionnez-la dans le fichier ReadMe (voir slide)
- S'il existe un standard encadrant le nommage des fichiers dans votre discipline, utilisez-le

A éviter :

- Les espaces, caractères spéciaux et accentués
- Les noms trop longs
- Les intitulés trop lacunaires

A privilégier :

- Les formats de date sans ambiguïté type YYYYMMDD ou YYYY-MM-DD
- Les numérotations à deux ou trois chiffres selon le volume de fichiers

Exemples :

- * Présentation_étudeDMP@HUG_fév.18.ppt ✓ 201802_DMP_HUGpresentation.ppt
- * ResultatsEnquêteV.corrr ✓ 2018_QualiteVie_EnqueteGE_Results_v02.csv

Voici un exemple de convention de nommage :

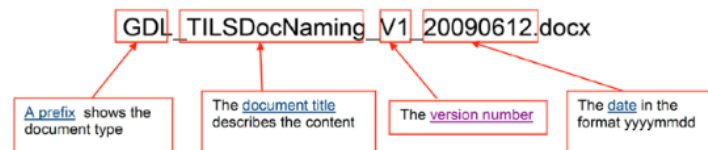


Figure 1 Convention de nommage de la Queensland University of Technology.
<https://www.library.qut.edu.au/about/management/documents/QUTTILSDocNamingConvention.pdf>

Les ressources :

DORANum, Comment bien nommer ses fichiers ?

https://doranum.fr/stockage-archivage/comment-nommer-fichiers_10_13143_wgqw-aa59/

Ant Renamer est un logiciel libre de renommage de fichiers (Windows)

<https://www.antp.be/software/renamer/fr>

Le fichier Readme

- Documente le dépôt
- Présente les **règles de nommage** et d'**organisation** du jeu de données
- Décrit le **contenu**
- Précise les **logiciels, méthodes, équipements** utilisés pour collecter et/ou analyser les données. Et les logiciels nécessaires pour les lire
- Dans les cas où les métadonnées/variables au sein des fichiers ne correspondent à aucun standard et ne sont spécifiées dans aucun fichier annexe, le ReadMe peut également les décrire (description de la variable, type de variable, unités, valeurs autorisées etc..)
- À rédiger au format texte (.txt) ou en markdown (.md)

Les ressources :

- Recherche Data Gouv. Modèle de Readme.
<https://recherche.data.gouv.fr/fr/categorie/33/guide/modele-de-readme>
- 4TU. Guidelines for creating a README file.
https://data.4tu.nl/s/documents/Guidelines_for_creating_a_README_file.pdf
- Make a Readme (Because no one can read your mind)
<https://www.makeareadme.com>

5

LE DÉPOT

Le dépôt

Les métadonnées

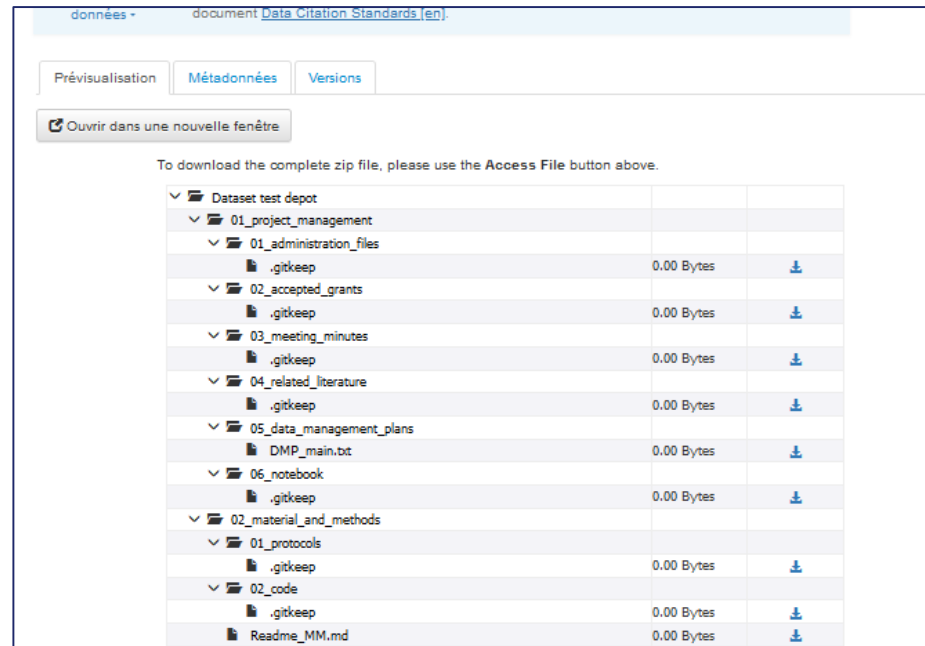
- Lors de la première étape du dépôt, seule une partie des métadonnées est présente. Il sera nécessaire de modifier votre dépôt dans un second temps pour enrichir les métadonnées (voir [slide](#)).
- Vous trouverez toutes les recommandations sur la saisie des métadonnées dans le [Guide de saisie des métadonnées](#).
- Si vous ne souhaitez pas parcourir ce guide, vous pouvez remplir les métadonnées a minima et nous vous indiquerons quelles métadonnées enrichir lors de la curation.
- Quelques éléments essentiels :
 - Pour l'**affiliation** renseignez : Nom du laboratoire, l'ensemble des tutelles, le pays
 - Dans la mesure du possible renseignez les identifiants ORCID des auteurs et autrices
 - Les **producteurs** sont les tutelles des auteurs et autrices
 - Il est important de lier vos **mots-clés** à des référentiels disciplinaires (ex : [MESH](#) , [V.O.INRAE](#), [IUPAC Gold Book](#), [NCI Thesaurus](#) ...)
 - Il est indispensable de renseigner vos **financeurs**

Le dépôt

Les fichiers : l'arborescence

Deux solutions :

Téléverser un fichier .zip de votre jeu de données : possibilité d'ouvrir le zip dans l'application pour visualiser l'arborescence et possibilité de télécharger les fichiers à l'unité. En revanche, impossible de pré-visualiser les fichiers.



données - document [Data Citation Standards \[en\]](#)

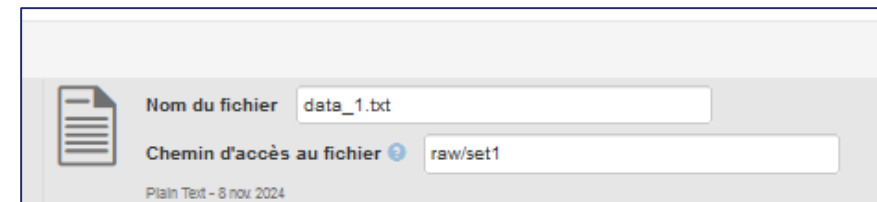
Prévisualisation **Métadonnées** Versions

Ouvrir dans une nouvelle fenêtre

To download the complete zip file, please use the **Access File** button above.

Arborescence	Taille	Action
Dataset test depot		
01_project_management		
01_administration_files		
.gitkeep	0.00 Bytes	↓
02_accepted_grants		
.gitkeep	0.00 Bytes	↓
03_meeting_minutes		
.gitkeep	0.00 Bytes	↓
04_related_literature		
.gitkeep	0.00 Bytes	↓
05_data_management_plans		
DMP_main.txt	0.00 Bytes	↓
06_notebook		
.gitkeep	0.00 Bytes	↓
02_material_and_methods		
01_protocols		
.gitkeep	0.00 Bytes	↓
02_code		
.gitkeep	0.00 Bytes	↓
Readme_MM.md	0.00 Bytes	↓

Reconstituer l'arborescence dans l'entrepôt : vous déposez l'ensemble des fichiers et reconstituez l'arborescence en renseignant le chemin d'accès de chaque fichier. Cela permet de pré-visualiser les fichiers (notamment fichiers texte et fichiers tabulés)



Nom du fichier

Chemin d'accès au fichier

Plain Text - 8 nov 2024



Fichiers **Métadonnées** Conditions Versions

Changer l'affichage Tableau Arborescence

- README.md (0 o)
- raw
 - set1
 - data_1.txt (0 o)
 - set2
 - simu

Le dépôt

Les fichiers : restreindre l'accès

L'accès restreint

Possibilité de restreindre l'accès à un fichier. Dans ce cas vous devez renseigner les conditions d'accès. Pour permettre à des personnes de le consulter deux solutions :

- Autoriser une liste de personnes ([contactez-nous](#))
- Autoriser les demandes d'accès

The screenshot shows a file management interface with a list of files: 'Dataset test depot.zip' and 'README.md'. A context menu is open over the files, with the 'Restreindre' option highlighted in red. A large dialog box titled 'Restrictions d'accès au fichier' is overlaid on the interface. The dialog contains the following text: 'Les fichiers publiés sont en accès restreint. Vous pouvez ajouter ou modifier les conditions d'accès pour le jeu de données et autoriser les personnes à demander l'accès aux fichiers restreints. Si vous désactivez le bouton Demander l'accès, vous devez ajouter de l'information dans le champ Conditions d'accès.' Below this, there is a section 'Demander l'accès' with a checked checkbox 'Autoriser la demande d'accès'. Under 'Conditions d'accès', there is a text input field containing 'Fichier accessible uniquement à la communauté de l'ESR français'. At the bottom of the dialog are buttons for 'Enregistrer les modifications' and 'Annuler'.

L'embargo

L'embargo permet de définir une date de publication plus tardive. Si vous attendez la parution d'une publication par exemple. La durée maximale est de 18 mois.

The screenshot shows the same file management interface as the previous one. The context menu is open, and the 'Embargo' option is highlighted in red. A dialog box titled 'Modifier l'embargo' is overlaid on the interface. The dialog contains the following text: 'Modifier l'embargo prévu pour le ou les fichiers sélectionnés. Une fois le jeu de données publié, vous devez communiquer avec un administrateur pour modifier la date ou la raison de l'embargo du ou des fichiers.' Below this, there is a section 'Ajouter ou modifier' with a label 'Choisir une date de fin de l'embargo'. A date input field shows '2026-02-09'. Below that is a text input field for 'Entrez une courte raison justifiant l'embargo' with the placeholder text 'Ajouter une raison...'. At the bottom of the dialog are buttons for 'Enregistrer les modifications' and 'Annuler'.

Partage de données nos publiées aux relecteurs de mon article

Si vous avez besoin de partager un jeu de données dans le cadre de la relecture d'un article voici la marche à suivre :

- Déposez votre jeu de données :
 - Sur [l'espace Sorbonne Université](#) si c'est une relecture « classique »
 - Sur l'espace générique, dans la collection « [revues anonymes](#) » si vous devez rester anonyme (votre jeu pourra être rattaché à l'espace Sorbonne Université a posteriori)
- Lors de la curation, demandez la non-publication du jeu de données et la génération d'une URL privée. Précisez également si vous avez besoin d'une URL privée anonyme.
- Dans votre article, vous pouvez déjà utiliser le DOI de votre jeu de données, même s'il n'est pas encore publié (Data availability, citations, figures..).
- L'équipe de curation vous transmettra une URL privée que vous pourrez transmettre à vos relecteurs et relectrices
- Une fois votre article publié, n'oubliez pas de demander la publication de votre jeu de données afin que le DOI ne pointe pas vers une page inaccessible !

La licence (conditions)

Par défaut c'est la [licence ouverte Etalab 2.0](#) qui est attribuée à votre dépôt. Elle permet la réutilisation gratuite et sans restriction de votre jeu de données à la seule condition que l'on cite les auteurs et autrices.



BIBLIOTHÈQUE
UNIVERSITAIRE

Pour modifier la licence par défaut, cliquez sur « Modifier les conditions »

Puis dans « condition d'utilisation, renseignez le code html correspondant à la licence souhaitées. Par exemple pour une licence CC-BY 4.0 :

```
<a rel="license"
href="http://creativecommons.org/licenses/by/4.0/"></a><br
/>Ce(tte) œuvre est mise à disposition selon les termes de la <a
rel="license"
href="http://creativecommons.org/licenses/by/4.0/">Licence
Creative Commons Attribution 4.0 International</a>
```

Notez que vous pouvez spécifier des conditions d'utilisation différentes pour les fichiers en accès restreint.

The screenshot shows a navigation bar with tabs for 'Fichiers', 'Métadonnées', 'Conditions', and 'Versions'. A 'Modifier les conditions' button is visible. The main content area is titled 'Conditions d'utilisation' and contains the text: 'Licence/Conditions d'utilisation des données' and 'Les [normes de la communauté](#) de même que les bonnes pratiques scientifiques prônent l'attribution du crédit au moyen d'une citation. Veuillez utiliser la référence bibliographique indiquée sur la page du jeu de données.' Below this text is the 'Etalab 2.0' logo.













The screenshot shows the same 'Conditions d'utilisation' page but with the 'Conditions d'utilisation' tab selected. The text now reads: 'Licence/Conditions d'utilisation des données' and 'Les [normes de la communauté](#) de même que les bonnes pratiques scientifiques prônent l'attribution du crédit au moyen d'une citation. Veuillez utiliser la référence bibliographique indiquée sur la page du jeu de données.' Below this is the text: 'Conditions personnalisées du jeu de données — les conditions d'utilisation personnalisées suivantes ont été définies pour le jeu de données.' At the bottom, there is a 'Conditions d'utilisation' section with the 'Creative Commons Attribution 4.0 International' logo and the text: 'Ce(tte) œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International'.






La licence (conditions)

Quelques licences ouvertes

Les Creatives commons (international)

La Licence ouverte Etalab (France)

CREATIVE COMMONS LICENSES		 COPY & PUBLISH	 ATTRIBUTION REQUIRED	 COMMERCIAL USE	 MODIFY & ADAPT	 CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✗
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✗

 You can redistribute (copy, publish, display, communicate, etc.)
 You have to attribute the original work
 You can use the work commercially
 You can modify and adapt the original work
 You can choose license type for your adaptations of the work.



LICENCE OUVERTE
OPEN LICENCE

L'enrichissement des métadonnées

Une fois votre dépôt enregistré, retournez sur les métadonnées afin de les enrichir.



The screenshot shows a data deposit page for a dataset titled "Test". At the top, there are status tabs: "Version provisoire" (selected), "Sous embargo", and "Non publiée". Below this, the dataset title "Test" is displayed, followed by the author "Moisan, Aurelien, 2024, 'Test'", a DOI link, and the source "Recherche Data Gov, VERSION PROVISoire". There are buttons for "Citer le jeu de données" and a link to "Data Citation Standards [en]". A dropdown menu is open on the right, showing options: "Modalités d'accès au jeu de données", "Publier le jeu de données", "Modifier le jeu de données", "Fichiers (développé)", "Métadonnées" (highlighted with a red box), "Conditions d'utilisation", and "Autorisations". Below the main content, there are fields for "Description" (containing "Test") and "Sujet" (containing "Earth and Environmental Sciences").

Dans ce formulaire augmenté vous pourrez notamment renseigner :

- Des **contributeurs et contributrices** : par exemple pour mentionner les plateformes de Sorbonne Université impliquées dans l'acquisition ou l'analyse de vos données
- Une **date** et un **lieu de production** : pour une meilleure traçabilité
- Les **logiciels utilisés** pour collecter ou générer vos données
- Des **jeux de données associés**

Vous trouverez toutes les recommandations sur la saisie de ces métadonnées dans le [Guide de saisie des métadonnées](#).

6

LE PROCESSUS DE DÉPÔT À SORBONNE UNIVERSITÉ

Le processus de publication d'un jeu de données à Sorbonne Université



BIBLIOTHÈQUE
UNIVERSITAIRE

L'équipe de curation

La curation

Lors de curation nous vérifions la qualité documentaire du jeu de données et nous vous transmettons des demandes de corrections

La publication

Le jeu est rendu public

Les corrections

Les corrections peuvent être réalisées conjointement par les déposant-es et l'équipe de curation

Le dépôt

Vous déposez vos fichiers en renseignant les métadonnées minimales

L'enrichissement

Vous modifiez votre dépôt pour accéder à la liste complète des métadonnées et les enrichir

La soumission

Vous soumettez votre dépôt à la curation

Les mises à jour

Les déposant-es



7

QUELQUES BONNES PRATIQUES

Mentionner mon jeu de données dans ma publication

A minima, mentionnez votre jeu de données dans un encart « Data Availability ». Consultez les consignes de votre éditeurs sur la mention des jeux de données

Data Availability Statement

A part of the data supporting this article is available within the article and **Supplementary Materials (Tables S1–S4 and more detailed descriptions of the program and examples of SQL queries adapted to our database)**. The main part of the data is publicly accessible in the 'Recherche Data Gouv' repository (space of Sorbonne Université) at <https://doi.org/10.57745/CBUPP3> (accessed on 22 April 2023) [30]. This data repository contains the database in various formats (CSV, SQL (preferred to import the database [43]), JSON and XML), 13361 models in XYZ format, 28 original Spartan files, the S8 DFT geometry in XYZ format, the files created by analysis of the models by the C program, the C program itself and the PHP program. See the README.txt file for details [44]. The PHP program was also referenced on Software Heritage via Hal [31].

Encart « Data Availability », de la publication [10.3390/ijms241512288](https://doi.org/10.3390/ijms241512288)

Il est également recommandé de transmettre le DOI du jeu de données à l'éditeur dès la soumission quand c'est possible afin qu'il puisse l'ajouter directement dans les métadonnées de la publication (« basedOnData ») bien que cette pratique soit encore peu répandue

Pour en savoir plus, consultez le guide [*Partager les données liées aux publications scientifiques*](#), du Comité pour la Science Ouverte.

Associer mon jeu de données à ma publication sur HAL

Il est également recommandé de mentionner votre jeu de données lors du partage de votre publication dans **HAL**, dans le champ « Données associées »

Pour toute question relative à HAL contactez l'équipe hal@sorbonne-universite.fr

Identifiants

Ajoutez les identifiants de votre dépôt dans d'autres entrepôts comme arXiv, PubMed, ADS, ...

DOI	10.3390/ijms241512288	Supprimer
PubMed	37569665	Supprimer
PubMed Central	PMC10419020	Supprimer
ArXiv		+ Ajouter

Données associées

Ajoutez les identifiants DOI fournis par l'entrepôt où vos données sont archivées.

10.57745/CBUPP3	Supprimer
	+ Ajouter

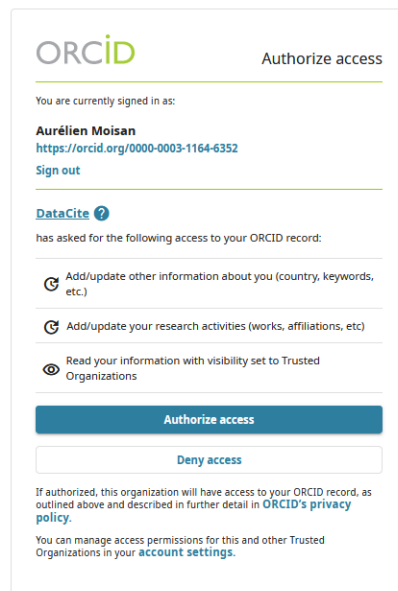
Logiciels associés

Ajoutez les identifiants SWHID des logiciels associés à votre dépôt

	+ Ajouter
--	-----------

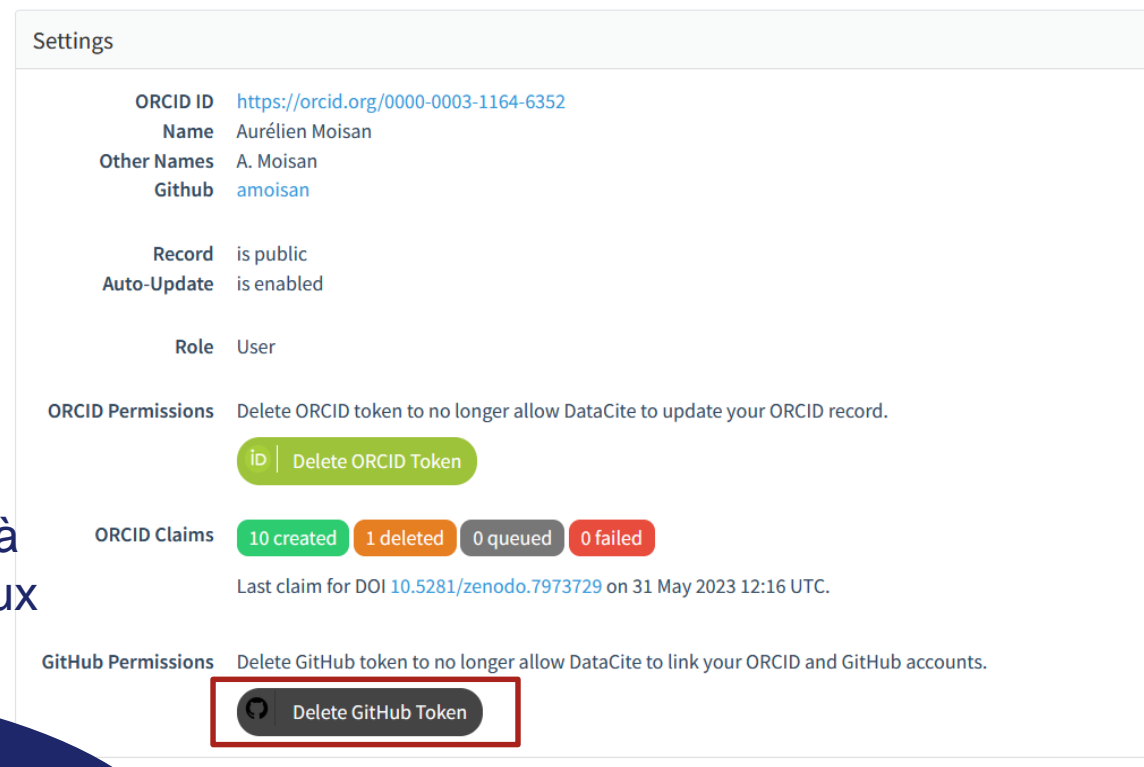
Afficher mes jeux de données sur mon profil ORCID

1- Connectez-vous sur votre profil ORCID, dans la rubrique « Travaux », faites [+ Ajouter] > [Chercher & Lier] et sélectionnez « DataCite »



2 - Autorisez Datacite à mettre à jour vos travaux

DataCite Profiles



Support Aurélien Moisan ▾

- Settings
- Commons Page**
- ORCID Record
- Sign Out

Grace à cette autorisation, à chaque fois que votre ORCID sera lié à un produit de la recherche (hors publication) identifié par un DOI, il sera automatiquement ajouté à votre profil

3 - Générez un Token GitHub si vous avez un compte GitHub et allez sur « Common Page »

Afficher mes jeux de données sur mon profil ORCID



3 - Recherchez vos jeux de données dans « Works »

4 - Consultez la notice complète et cliquez sur « Add to ORCID Record »



ORCID

Pages - Support Aurélien Moisan -

[Works](#) [People](#) [Organizations](#) [Repositories](#)

30,407 Works

[Identity Awareness: Toward an Invisible e-Infrastructure for Identifying Data and Authors; eResearch Australasia 2012](#)

Amir Aryani & Adrian Burton

Journal Contribution published 2013 in [figshare Academic Research System](#)

In this paper, we describe identity awareness in the context of e-Infrastructure and looks at the interoperability of identifiers for datasets, researchers and research projects. In addition, we discuss the role of Australian National Data Service (ANDS) in the European Union (EU) funded ORCID and DataCite Interoperability Network (ODIN) project. **Citation information in bibtex format:**
@CONFERENCE{aryani2012id,



Type to search...

Pages - Support Aurélien Moisan -

[Works](#) [People](#) [Organizations](#) [Repositories](#)

[Identity Awareness: Toward an Invisible e-Infrastructure for Identifying Data and Authors; eResearch Australasia 2012](#)

<https://doi.org/10.6084/m9.figshare.155650>

[Description](#) [Creators](#) [Registration](#)

In this paper, we describe identity awareness in the context of e-Infrastructure and looks at the interoperability of identifiers for datasets, researchers and research projects. In addition, we discuss the role of Australian National Data Service (ANDS) in the European Union (EU) funded ORCID and DataCite Interoperability Network (ODIN) project. **Citation information in bibtex format:**
@CONFERENCE{aryani2012id,
author = {Amir Aryani and Adrian Burton},
title = {Identity Awareness: Toward an Invisible e-Infrastructure for Identifying Data and Authors},
booktitle = {eResearch Australasia},
year = {2012},
address = {Sydney, Australia},
month = {October},
doi = {dx.doi.org/10.6084/m9.figshare.155650}}

Cite as


Aryani, A., & Burton, A. (2013). *Identity Awareness: Toward an Invisible e-Infrastructure for Identifying Data and Authors; eResearch Australasia 2012*. <https://doi.org/10.6084/m9.FIGSHARE.155650>

Share

Email


Publier un data paper

- Article décrivant un jeu de données, notamment les méthodes de collecte
- Détaille les potentiels de réutilisation de données
- Le DMP peut servir de trame
- Peut être publié dans une revue académique classique, plus souvent publié dans un data journal
- Les articles font l'objet d'un peer-reviewing

SCIENTIFIC DATA 

Data Descriptor | OPEN | Published: 10 April 2018

Long-term dataset on aquatic responses to concurrent climate change and recovery from acidification

Taylor H. Leach , Luke A. Winslow, Frank W. Acker, Jay A. Bloomfield, Charles W. Boylen, Paul A. Bukaveckas, Donald F. Charles, Robert A. Daniels, Charles T. Driscoll, Lawrence W. Eichler, Jeremy L. Farrell, Clara S. Funk, Christine A. Goodrich, Toby M. Michelena, Sandra A. Nierzwicki-Bauer, Karen M. Roy, William H. Shaw, James W. Sutherland, Mark W. Swinton, David A. Winkler & Kevin C. Rose

Scientific Data 5, Article number: 180059 (2018) | [Download Citation](#) ↓

Abstract

Concurrent regional and global environmental changes are affecting freshwater ecosystems. Decadal-scale data on lake ecosystems that can describe processes affected by these changes are important as multiple stressors often interact to alter the trajectory of key ecological phenomena in complex ways. Due to the practical challenges associated with long-term data collections, the majority of existing long-term data sets focus on only a small number of lakes or few response variables. Here we present physical, chemical, and biological data from 28 lakes in the Adirondack Mountains of northern New York State. These data span the period from 1994–2012 and harmonize multiple open and as-yet unpublished data sources. The dataset creation is reproducible and transparent; R code and all original files used to create the dataset are provided in an appendix. This dataset will be useful for examining ecological change in lakes undergoing multiple stressors.

Metadata summary

Design Type(s)	observation design • time series design • data integration objective
Measurement Type(s)	geographic location • Physical Phenomenon or Property • Inorganic Chemistry • planktonic material • weather
Technology Type(s)	digital curation • data acquisition system • Analytical Chemistry • light microscopy
Factor Type(s)	
Sample Characteristic(s)	Adirondack Park • freshwater lake biome

[Download metadata file](#) ↓

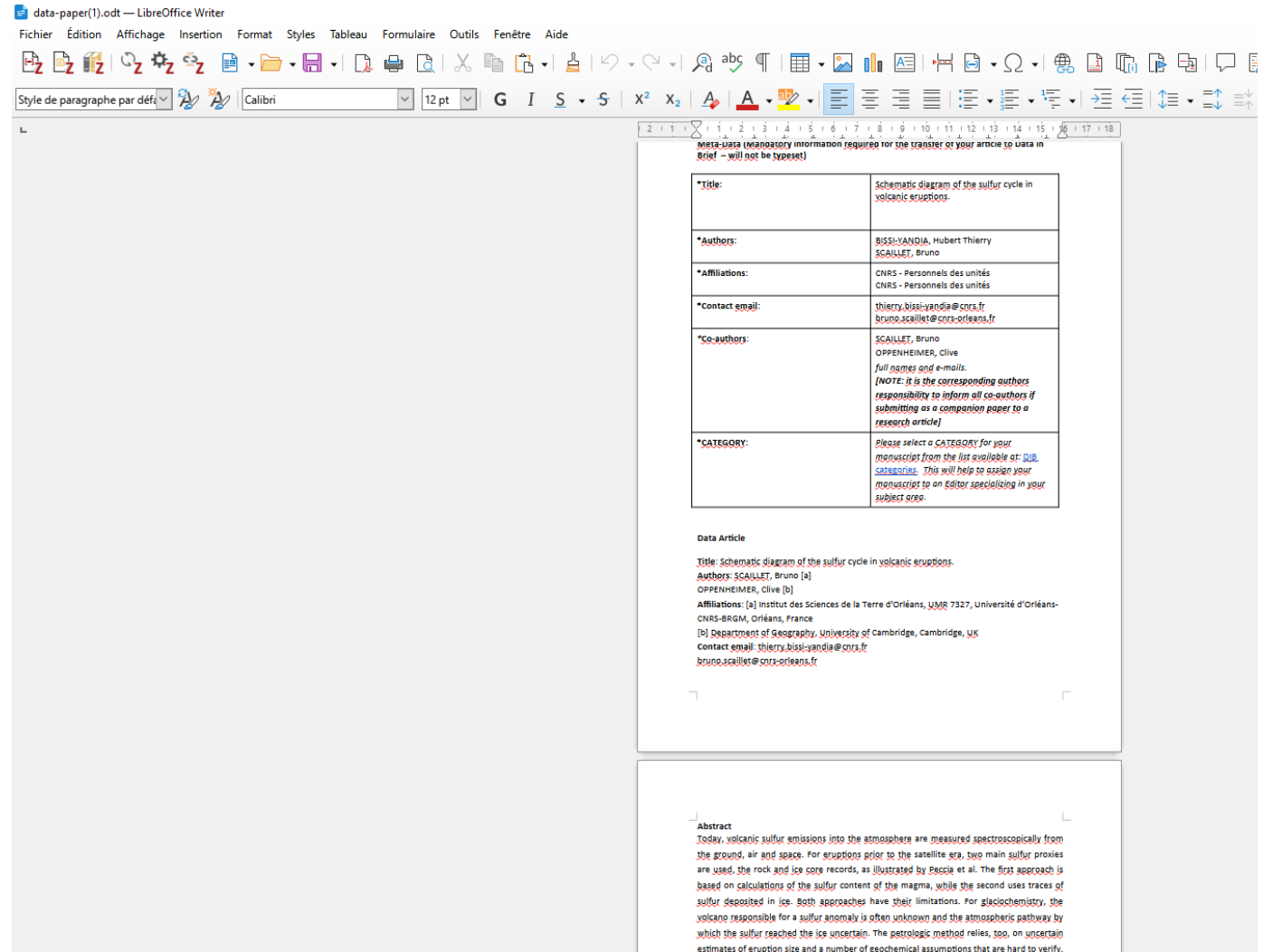
Leach TH, Winslow LA, Acker FW, Bloomfield JA, Boylen CW, Bukaveckas PA, et al. Long-term dataset on aquatic responses to concurrent climate change and recovery from acidification. Scientific Data [Internet]. 2018 [cité 28 mars 2019];5. Disponible sur: <https://www.nature.com/articles/sdata201859>

Publier un data paper

- Recherche Data Gouv vous permet de générer une trame de data paper à partir du DOI de votre jeu de données. Si les métadonnées de votre jeu sont complètes, un grand nombre d'informations seront déjà renseignées et cela facilitera votre rédaction.

Accéder au générateur :

<https://entrepot.recherche.data.gouv.fr/datapartage-datapapers-web/>



The screenshot shows the LibreOffice Writer interface with a document titled "data-paper(1).odt". The document contains a form for generating a data paper. The form includes fields for Title, Authors, Affiliations, Contact email, Co-authors, and Category. Below the form, there is a section for "Data Article" with a title, authors, affiliations, and contact email. At the bottom, there is an "Abstract" section with a paragraph of text.

data-paper(1).odt — LibreOffice Writer

Fichier Édition Affichage Insertion Format Styles Tableau Formulaire Outils Fenêtre Aide

Style de paragraphe par défaut Calibri 12 pt G I S - S x² x₂ A A

meta-data (mandatory information required for the transfer of your article to Data in Brief - will not be typeset)

*Title:	Schematic diagram of the sulfur cycle in volcanic eruptions.
*Authors:	BISSI-YANQIA, Hubert Thierry SCAILLET, Bruno
*Affiliations:	CNRS - Personnels des unités CNRS - Personnels des unités
*Contact email:	thierry.bissi-yanqia@cnrs.fr bruno.scaillet@cnrs-orleans.fr
*Co-authors:	SCAILLET, Bruno OPPENHEIMER, Clive Full names and e-mails. <i>[NOTE: it is the corresponding authors responsibility to inform all co-authors if submitting as a companion paper to a research article]</i>
*CATEGORY:	Please select a CATEGORY for your manuscript from the list available at: DIB categories . This will help to assign your manuscript to an Editor specializing in your subject area.

Data Article

Title: Schematic diagram of the sulfur cycle in volcanic eruptions.
Authors: SCAILLET, Bruno [a]
OPPENHEIMER, Clive [b]
Affiliations: [a] Institut des Sciences de la Terre d'Orléans, UMR 7327, Université d'Orléans-CNRS-BRGM, Orléans, France
[b] Department of Geography, University of Cambridge, Cambridge, UK
Contact email: thierry.bissi-yanqia@cnrs.fr
bruno.scaillet@cnrs-orleans.fr

Abstract

Today, volcanic sulfur emissions into the atmosphere are measured spectroscopically from the ground, air and space. For eruptions prior to the satellite era, two main sulfur proxies are used, the rock and ice core records, as illustrated by Peccia et al. The first approach is based on calculations of the sulfur content of the magma, while the second uses traces of sulfur deposited in ice. Both approaches have their limitations. For glaciochemistry, the volcano responsible for a sulfur anomaly is often unknown and the atmospheric pathway by which the sulfur reached the ice uncertain. The petrologic method relies, too, on uncertain estimates of eruption size and a number of geochemical assumptions that are hard to verify.

Pour aller plus loin

- Consultez [DORANum](#), la plateforme d'apprentissage sur les données de la recherche

Les ressources de Recherche Data Gouv :

- Suivez les [classes virtuelles](#)
- La [boîte à outils](#)
- Le [guide du déposant](#)
- Le [guide de saisie des métadonnées](#)
- La bac à sable : <https://demo.recherche.data.gouv.fr>
- Pour toute question sur Recherche Data Gouv ou les données de la recherche : labrador@sorbonne-universite.fr

Accompagnement sur les données à Sorbonne Université



<https://labrador.sorbonne-universite.fr/>

Pour toute question sur les données :
labrador@sorbonne-universite.fr

- Accompagnement personnalisé concernant les données des projets
- Accompagnement à la rédaction de plan de gestion de données et relecture
- Aide au dépôt de vos jeux de données sur Recherche Data Gouv, Nakala ou un autre entrepôt
- Accompagnement sur le stockage, le partage et l'archivage de vos données

Formations

<https://labrador.sorbonne-universite.fr/besoin-daccompagnement/se-former-sur-les-donnees-sorbonne-universite>

Accompagnement Science Ouverte

HAL

hal@sorbonne-universite.fr

Permanences :

<https://us02web.zoom.us/j/87056231909>

- mardi 14h-15h
- vendredi 11h-12h

Formation Science Ouverte

<https://www.sorbonne-universite.fr/engagement-pour-la-science-ouverte>

Pour toute question

publications@sorbonne-universite.fr

MERCI



BIBLIOTHÈQUE
UNIVERSITAIRE



Sauf mention contraire, cette présentation est mise
à disposition selon les termes de la Licence
Creative Commons Attribution 2.0 France.
Icônes : freepik