CELLULE DATA
GRENOBLE ALPES

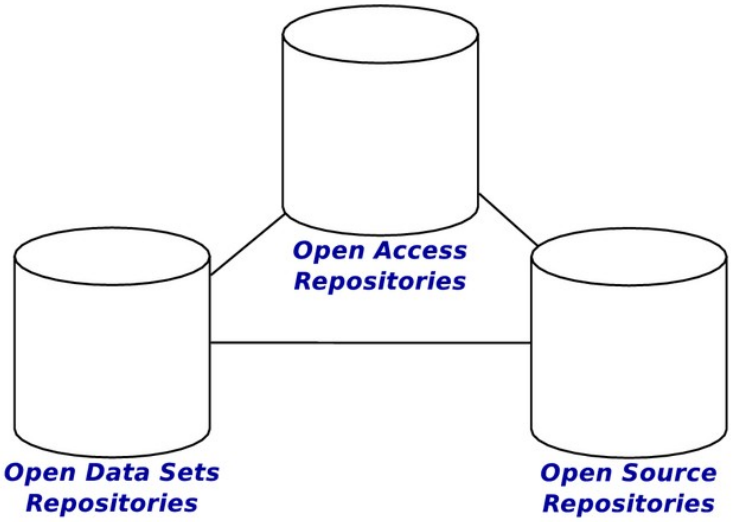
Comment bien diffuser ses données à l'issue de sa thèse

Formation doctorale 11/06/2026

Licence CC-BY

Contexte

- **Processus de publication :**
- 1 - Article scientifique : principaux résultats d'un travail scientifique - citabilité
- 2 - Diffusion des données : reproductibilité – réutilisabilité- citabilité
- 3 - Diffusion des codes : reproductibilité – réutilisabilité - citabilité
- 4 - Diffusion thèses : sites institutionnels dédiés - plus de détails scientifiques et techniques - citabilité



Pourquoi diffuser ?

- **Pourquoi ne pas partager ?**

  - Données **personnelles** et données **sensibles**
    - Par ex, environnementales, santé,
      - Procédures spécifiques (anonymisation…)

  - Données présentant des risques
    - défense nationale
    - sécurité publique, des Etats, établissements

  - Données liées à une zone à régime restrictif (ZRR) >>> voir plus loin

  - Données liées à des secrets professionnels

Pourquoi diffuser ?

- **Pourquoi ne pas partager ?**
  - Données comportant une **valeur économique**
    - Brevet
    - Convention/contrat avec des entreprises
      - En vérifier les termes
      - Discussions nécessaires
  - **Publication en cours** de préparation
    - Possibilité de déposer ses données associées et de mettre un embargo avant la diffusion

    A noter : ces points de vigilance sont valables aussi pour les publications

Pourquoi diffuser ?

- **Intégrité et ré-utilisabilité**
  - **Éthique** et intégrité scientifique (voir code de la recherche, art D21 1-2 et l'art de l'OFIS)
  - Retour vers les citoyens (financement public)
  - Garantir la **reproductibilité** des résultats
    - Fiabilité
    - Transparence
  - Assurer la **re-utilisabilité** des données
    - Intérêt pour d'autres projets scientifiques
    - Favoriser le progrès de la recherche et l'émergence de nouvelles recherches
    - Mettre en oeuvre les principes **FAIR** pour les codes et les données

FAIR = Findable, Accessible, Interoperable, Reusable

Pourquoi diffuser ?

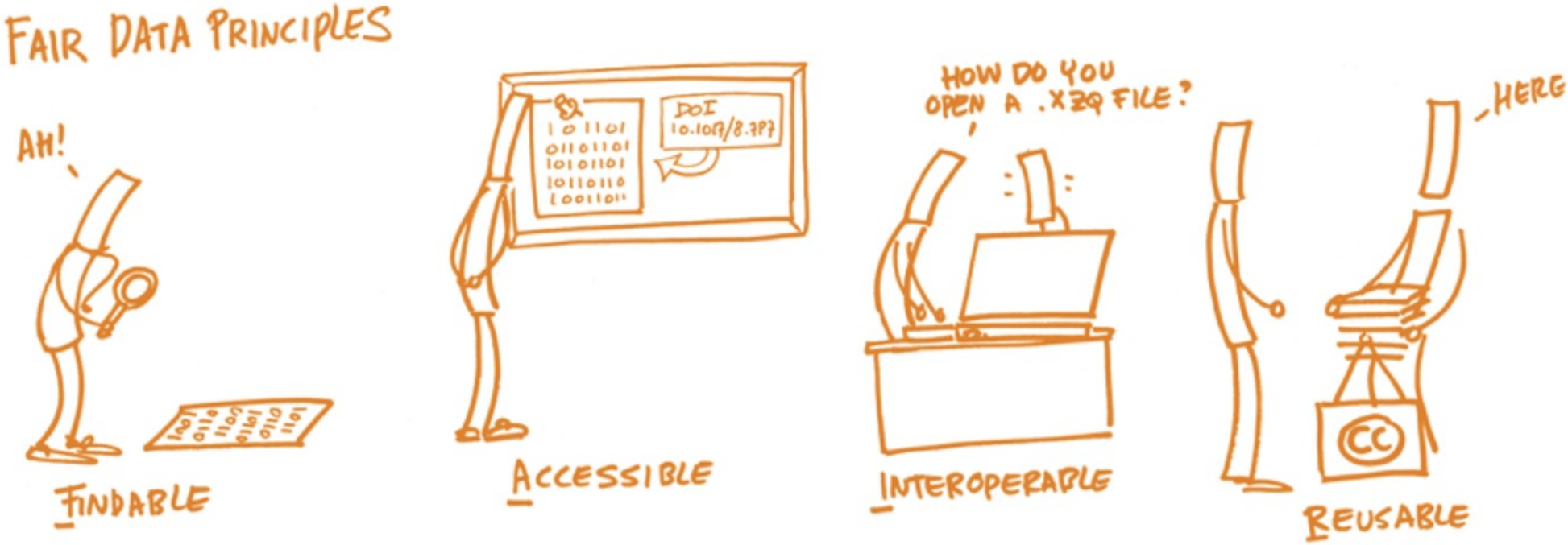


Image: https://book.fosteropenscience.eu/

Voir aussi la fiche : Le minimum à FAIR

Pourquoi diffuser ?

FAIR = Findable, Accessible, Interoperable, Reusable



Source : The Cambridge Crystallographic Data Centre

Voir aussi la fiche : Le minimum à FAIR

Pourquoi diffuser ?

- **Pourquoi partager ? En réalité**

  Certaines données ne sont produites ou traitées que par le biais de coopérations nationales et internationales

  - Des données **précieuses ou uniques**

    - Dont le coût de production ou de traitement
      - mobilise des moyens très importants tout au long du cycle de vie au niveau des Etats
      - et nécessite une exploitation scientifique commune ou partagée

        - Ex : données astronomiques, physique des hautes particules, données génomiques, données de sismologie…
        - Ex: boson de Higgs

Pourquoi diffuser ?

- **Pourquoi partager ? En réalité**

Augmenter la visibilité

Données accessibles et **citables** indépendamment de l'article

Lier les données à ses publications

Augmenter la visibilité de ses recherches,

Donc la citation de sa recherche !

Voir :

- Institut de l'Information Scientifique et Technique, Ognibene, M., & Tandar, S. (2023). DoRANum-Dépôts et entrepôts : Ouverture des données, entre injonctions et intérêts. DoRANum. **https://doi.org/10.13143/MKR2-S776**
- Les données de la recherche : un sujet stratégique. Collège Données de la recherche du Comité pour la science ouverte. [En ligne]. https://www.ouvrirlascience.fr/les-donnees-de-la-recherche-un-sujet-strategique/
- Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. **https://doi.org/10.1371/journal.pone.0230416**

Pourquoi diffuser ?

Pourquoi partager ? En réalité

Répondre aux exigences des éditeurs**:**

Préconisation des éditeurs sur l'accès aux données liées aux publications (data sharing)

Quelques exemples
EDP Sciences
Plos One
Elsevier
Springer / Nature
Taylor and Francis

Pourquoi diffuser ?

Pourquoi partager ? En réalité

Répondre aux exigences des financeurs

- Horizon Europe 2021-2027 :

La «science ouverte» deviendra le mode opératoire d'Horizon Europe. Il exigera donc un accès ouvert aux publications et aux données.»

- ANR (contribuer à l'ouverture des données quand c'est possible)
- NIH Data Management and Sharing Policy

Répondre aux exigences des Etats et établissements

- Loi pour une République Numérique (art 30)
- France : 2e Plan National pour la Science Ouverte (2021-2024)

  Axe 2 : Structurer, partager et ouvrir les données de la recherche

  Axe 3 : Ouvrir et promouvoir les codes sources produits par la recherche
- CNRS : Plan données de la recherche (nov 2020)
- Charte du CEA pour la science ouverte (2021)
- Charte UGA

Pourquoi diffuser ?

Focus Loi pour une République Numérique

- Article 30 alinea 2, repris dans le code de la recherche, chapitre III du titre III du livre V
  « II.-Dès lors que les données issues d'une activité de recherche financée au moins pour moitié par des dotations de l'Etat, des collectivités territoriales, des établissements publics, des subventions d'agences de financement nationales ou par des fonds de l'Union européenne ne sont pas protégées par un droit spécifique ou une réglementation particulière et qu'elles ont été rendues publiques par le chercheur, l'établissement ou l'organisme de recherche, leur réutilisation est libre.
  « III.-L'éditeur d'un écrit scientifique mentionné au I ne peut limiter la réutilisation des données de la recherche rendues publiques dans le cadre de sa publication.

- Sur les données : Cécile Arènes, Lionel Maurel, Stephanie Rennes. Guide d'application de la Loi pour un e République numérique pour les données de la recherche. Comité pour la science ouverte. 2022. ⟨hal-03968218⟩

Pourquoi diffuser ?

Focus Loi pour une République Numérique

- **Rappel pour les publications alinéa 1 : droit de déposer le postprint dans une archive ouverte !**
  Lorsqu'un écrit scientifique issu d'une activité de recherche financée au moins pour moitié par des dotations de l'Etat, des collectivités territoriales ou des établissements publics, par des subventions d'agences de financement nationales ou par des fonds de l'Union européenne **est publié dans un périodique paraissant au moins une fois par an,** son auteur dispose, même après avoir accordé des droits exclusifs à un éditeur, du **droit de mettre à disposition gratuitement dans un format ouvert**, par voie numérique, **sous réserve de l'accord des éventuels coauteurs, la version finale de son manuscrit acceptée pour publication**, dès lors que l'éditeur met lui-même celle-ci gratuitement à disposition par voie numérique ou, à défaut, à l'expiration d'un délai courant à compter de la date de la première publication. Ce délai est au maximum de six mois pour une publication dans le domaine des sciences, de la technique et de la médecine et de douze mois dans celui des sciences humaines et sociales.

  Voir :
- la FAQ sur la Loi pour une République Numérique
- Je publie, quels sont mes droits ? Comité pour la science ouverte – Collège Compétences – GT Juridique, 2ème édition – Octobre 2020

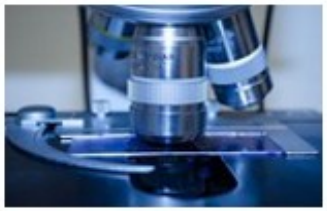
-

Types de données de la recherche

Il existe différents types de données de la recherche qui diffèrent selon la manière dont les données sont produites et selon leur valeur supposée.

Données d'observation

- capturées en temps réel ;
- habituellement uniques et donc impossibles à reproduire ;
  - Exemples : neuroimagerie, photographie astronomique, données d'enquête.



publicdomainpictures.net



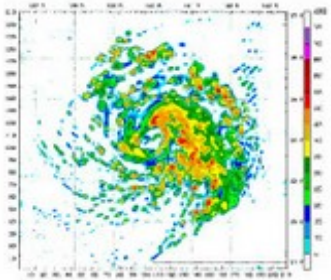
publicdomainpictures.net

Données expérimentales

- obtenues à partir d'équipements de laboratoire ;
- souvent reproductibles mais parfois coûteuses ;
  - Exemples : chromatogrammes, puces à ADN.

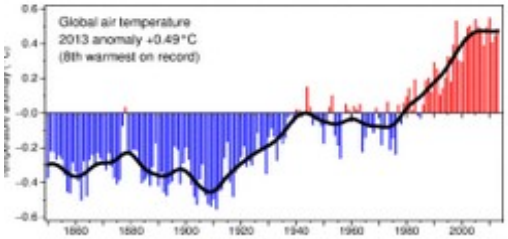
Données computationnelles ou de simulation

- générées par des modèles informatiques ou de simulation ;
- souvent reproductibles si le modèle est correctement documenté ;
  - Exemples : modèle météorologique, modèle de simulations sismiques, modèle économique.



Typhoon_Mawar_2005_computer_simulation.gif: Atmoz

Données de la recherche et publications scientifiques dans *Une introduction à la gestion et au partage des données de la recherche*, INIST



cru.uea.ac.uk/

Données dérivées ou compilées

- issues du traitement ou de la combinaison de données "brutes" ;
- souvent reproductibles mais coûteuses ;
  - Exemples : fouille de texte, bases de données compilées.

Données de référence

- collection ou accumulation de petits jeux de données qui ont été revus par les pairs, annotés et mis à disposition ;
  - Exemples : GenBank, base de données de cristallographie, collection de lettres ou archive d'images historiques.



publicdomainpictures.net

Que diffuser ?

Des exemples en Science, Technologie et Médecine

- **Est-ce que les données brutes suffisent à assurer la reproductibilité ou la réutilisation**
  - **Est-ce que les données ont nécessité un pré-traitement / analyses?**
    - coûteux en temps ou en ressources
    - Si oui, diffusion du processus de pré-traitement / analyses
  - **Est-ce que des outils/codes/logiciels sont nécessaires pour exploiter les données (pour pouvoir utiliser les analyses par exemple)**
    - indiquer les outils nécessaires pour l'exploitation des données
    - Si oui, diffuser si possible les outils/codes/logiciels liés

Impact environnemental : limiter le volume de données

Que diffuser ?

Des exemples en Sciences humaines et sociales

- **Observations (par ex, observations urbaines)**
  - Captation unique

- **Enquêtes - questionnaires**
  - Administration unique
  - Données initiales (réponses brutes)
  - Données issues du dépouillement et des croisements
    >>> diffuser les données initiales permettra de nouveaux croisements

- **Sources historiques, littéraires, juridiques, linguistiques**
  - Sont-elles accessibles en ligne (par ex sur Gallica)
    - Si oui, pas la peine de les déposer
  - La constitution du corpus est-elle originale ?
    - Si oui, le diffuser

    - Point de vigilance :
      - Respect du RGPD :
        - consulter la dpo (dpo@univ-grenoble-alpes.fr) !!!!!
      - Impact environnemental : limiter le volume de données

Que diffuser ?

Des exemples en Sciences humaines et sociales

- **Captation et numérisation**
  - **Numérisation**
    - Conserver les images brutes
    - Y-a-t-il des questions de droits d'exploitation ?

  - **Entretiens et transcriptions**
    - Y-a-t-il des questions de respect du RGPD
    - Diffuser les données initiales (enregistrement audio) ?
    - Diffuser uniquement la transcription ?

  - **Captation vidéo et transcription/analyse**
    - Y-a-t-il des questions de respect du RGPD ?

- **Bases de dépouillement – Exploitation - Visualisation**
- Y a-t-il production d'une base de données originale (ex : base de données relationnelle) ?
- Y-a-t-il des visualisations (par ex, graphes) originales ?

Impact environnemental : limiter le volume de données

Que diffuser ?

Focus : Impact environnemental

- **Quelques ressources**
  - EcoInfo, Groupement de recherche et de service
  - GreenData pour un impact environnemental maitrisé : livrables

- **A noter** : appliquer les principes FAIR à ses données permet de rendre leur partage efficace et leur réutilisation possible.

Que diffuser ?

- **Critère d'utilisabilité : le principal critère**
  - Quelles données répondent aux objectifs de votre thèse ?
    - Pensez à la soutenance !
  - Quelles données peuvent représenter un intérêt pour la communauté?
- **Statut des données**
  - Est-ce que ces données existent déjà ?
    - Si oui, vérifier la pérennité de la source des données
  - Est-ce que ces données sont uniques ?
    - Si oui, les diffuser
    - Si non, déterminer s'il vaut mieux diffuser les données ou le processus les ayant fournies (par exemple simulation, code source)

Que diffuser ?

Diffuser des données « négatives » ou « non concluantes » ?

- Problématique dans toutes les disciplines
- Enjeu de reproductibilité
- Biais des résultats

- Voir Plan National Science Ouverte 2021-2024 : « Réduire le biais de publication, qui est la tendance à ne publier que les études ayant obtenu un résultat positif, au détriment des résultats peu concluants ou négatifs. »

Comment diffuser ? focus licences

Définir l'usage des jeux de données que vous diffusez

- **Condition indispensable** pour la réutilisation de vos données
- Nécessité de choisir une licence d'usage

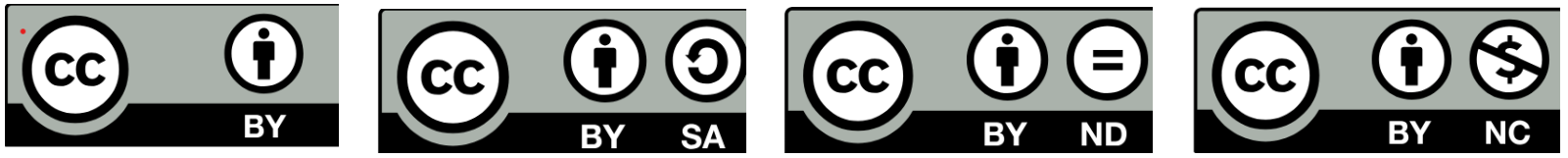
- **Licence** : contrat par lequel un titulaire d'un droit de propriété intellectuelle concède en tout ou partie la jouissance de ce droit (droits de reproduction, de représentation, droit d'autoriser les œuvres dérivées…)

A noter :

- Voir le module asynchrone « Découvrir les données de la recherche », leçon 5 « focus licences » sur E-formation
- Voir le support de la formation « Respecter le(s) droit(s) dans sa thèse »

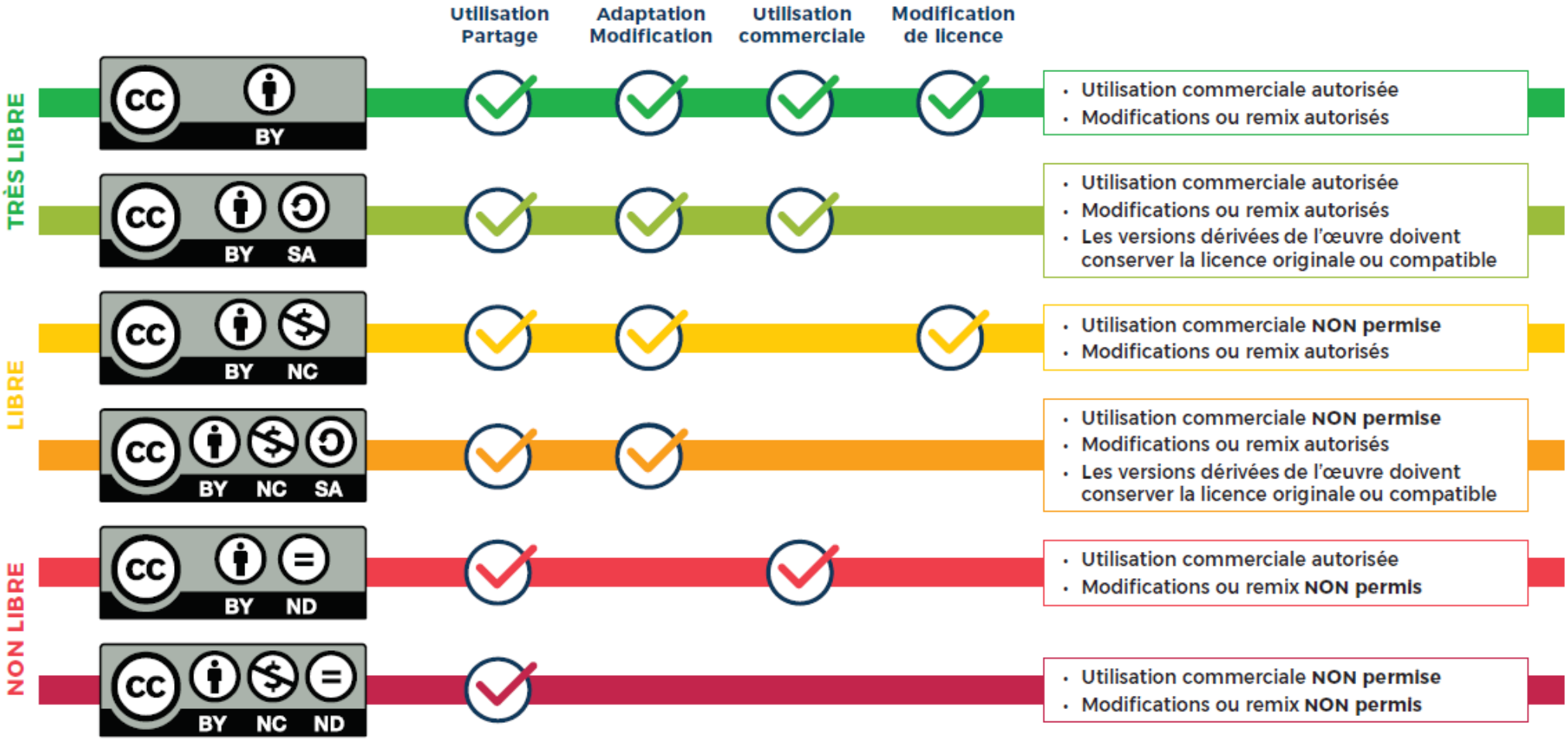
Comment diffuser ? focus licences

- Licences **Creative Commons** (les plus répandues pour les jeux de données)
  - BY : Attribution
  - SA : Share Alike (partage à l'identique)
  - NC : Non Commercial
  - ND : Non Derivative (pas de modification)
  - **Il suffit d'ajouter les logos de chaque famille souhaitée au contenu auquel on veut appliquer la licence**
  - **On combine les différentes licences (ex : cc by-NC, cc-by-SA, cc-by NC-SA …)**



- Licence ouverte Etalab (open licence, compatible cc-by)

Les licences Creative Commons



BY

ATTRIBUTION

Vous pouvez retenir, réutiliser, réviser, remixer et redistribuer.

L'auteur doit être cité

SA

PARTAGE DANS LES MÊMES CONDITIONS

Vous pouvez retenir, réutiliser, réviser, remixer et redistribuer.

Partage sous licence compatible

NC

POUR USAGE NON COMMERCIAL

Vous pouvez retenir, réutiliser, réviser, remixer et redistribuer.

Pour usage non commercial

ND

PAS DE MODIFICATION

Vous pouvez retenir, réutiliser et redistribuer.

Aucune modification permise

fabrique REL
RESSOURCES ÉDUCATIVES LIBRES

L'auteur de l'œuvre doit être cité en toutes circonstances pour l'ensemble de ces licences.

Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International.

Comment diffuser ? focus licences

Définir l'usage des jeux de données que vous diffusez

Choix de licences

Des **outils de sélection de licences** pour les dépôts de données ou de codes :

- License Chooser (Creative Commons)
- License Selector (codes et données)

Comment diffuser ? focus licences

Définir l'usage des jeux de données que vous diffusez

- **Des ressources pour vous aider :**

  - JURISDOR, Ressources juridiques données de la recherche
  - Arbre Aide à la décision sur la diffusion des données de recherche (Cirad)

- **Ressources DoraNum**

  - Licences : Institut de l'Information Scientifique et Technique. (2025). DoRANum-Aspects juridiques, éthiques, intégrité scientifique : Guide des licences ouvertes pour les productions scientifiques. DoRANum. https://doi.org/10.13143/TV6F-SV31
  - Aspects juridiques et éthiques (RGPD, éthique, doit et open data), DoraNum

Comment diffuser ?

Définir l'usage des jeux de données que vous diffusez /
Pour aller plus loin, des guides :

- Nicolas Becard, Céline Castets-Renard, Gauthier Chassang, Martin Dantant, Laurence Freyt-Caffin, et al.. Ouverture des données de la recherche. Guide d'analyse du cadre juridique en France. [Rapport de recherche] Comité pour la science ouverte. 2017, 45 p. ⟨hal-02791224⟩

- INSHS, Les sciences humaines et sociales et la protection des données à caractère personnel dans le contexte de la science ouverte – Guide pour la r echerche, v2, 2021.

- Véronique Ginouvès, Isabelle Gras. La diffusion numérique des données en SHS – Guide des bonnes pratiques éthiques et juridiques, Presses universitaires de Provence, 2018, Digitales, hal-⟨01903040⟩

Comment diffuser ?

Quel accès ? Plusieurs choix possibles : **Libre accès :** Immédiat ? Différé ? Sur demande ?

Focus sur les Zones à Régime Restrictif (ZRR) et dispositifs de protection du potentiel scientifique et technique de la nation (PPST)

Enjeu : Protection des « savoirs et savoir-faire stratégiques » et technologies sensibles (Protection juridique et administrative ; Contrôle des accès physiques et numériques)

Pour la diffusion des données, c'est le chercheur qui décide

La diffusion de données doit au préalable avoir été expressément autorisée par le responsable de la ZRR
Le fonctionnaire de sécurité défense ou d'autres services peuvent être sollicités.

Voir

- Institut de l'Information Scientifique et Technique, & ADOC Lorraine – Université de Lorraine. (2023). DoRANum-Aspects juridiques, éthiques, intégrité scientifique : Les données de la recherche produites en Zone à Régime Restrictif. DoRANum. https://doi.org/10.13143/6F2R-8263
- Bassinet, A., Bouchet Moneret, F., Bracco, L., & Jouneau, T. (2025, September 25). Les données de la recherche produites en zone à régime restrictif. Atelier de la donnée ADOC Lorraine. https://doi.org/10.5281/zenodo.17201551

Comment diffuser ?

Rappel : respect du RGPD à toutes les étapes de la recherche, dès la collecte et le stockage mais aussi pour le traitement, la diffusion, etc

Pour la diffusion :
- Consulter toujours la dpo le plus en amont possible pour la diffusion
- Envisager les modalités de diffusion (lieu, type de diffusion, accès sur demande, .. .

Il n'est pas toujours possible/souhaitable de partager des données personnelles et a fortiori sensibles
Pour diffuser des données personnelles ou sensibles, plusieurs choix possibles :
• **anonymisation des données.** Attention : l'anonymisation est un processus complexe et irreversible. Elle peut conduire à une perte en terme d'utilisabilité des données.
• **accès sur demande (pour les données pseudonymisées)**

Formation : Savoir intégrer des pratiques respectueuses du RGPD à toutes les étapes de la thèse- support 2026
Pour aller plus loin, voir sur le site de la CNIL:
• Les règles concernant la recherche scientifique hors santé
• Quelles formalités pour les traitements de données de santé ?

Comment diffuser ?

Quel accès ? Focus respect du RGPD

Le principe de l'anonymisation et de la pseudonymisation

PROCESSUS	PSEUDONYMISATION	ANONYMISATION
STATUT DES DONNÉES	Personnelles (restent indirectement identifiantes et donc soumises au RGPD et à la loi Informatique et Libertés	Anonymes
RÉUTILISATION DES DONNÉES	Sous conditions	Sans restriction
UTILITÉ DES DONNÉES	Préservée car pas d'altération du niveau de détail des données	Plus ou moins altérée en fonction des objectifs poursuivis et des méthodes appliquées
MÉTHODES À METTRE EN OEUVRE	Compteur, générateur de nombres aléatoires, fonction de **hachage**, chiffrement à clé secrète, etc.	Dépend des objectifs poursuivis : **confidentialité différentielle**, randomisation, k-anonymat, l-diversité, t-proximité, etc.
COMPLEXITÉ DE LA MISE EN OEUVRE	Simple à moyenne	Dépend des objectifs poursuivis : simple dans certains cas comme l'agrégation ou le comptage et complexe dans d'autres

Exemples d'outils : amnesia, ARX

Voir fiche pratique CNIL : enjeux et avantages de l'anonymisation et de la pseudonymisation (2022)

Quand diffuser ?

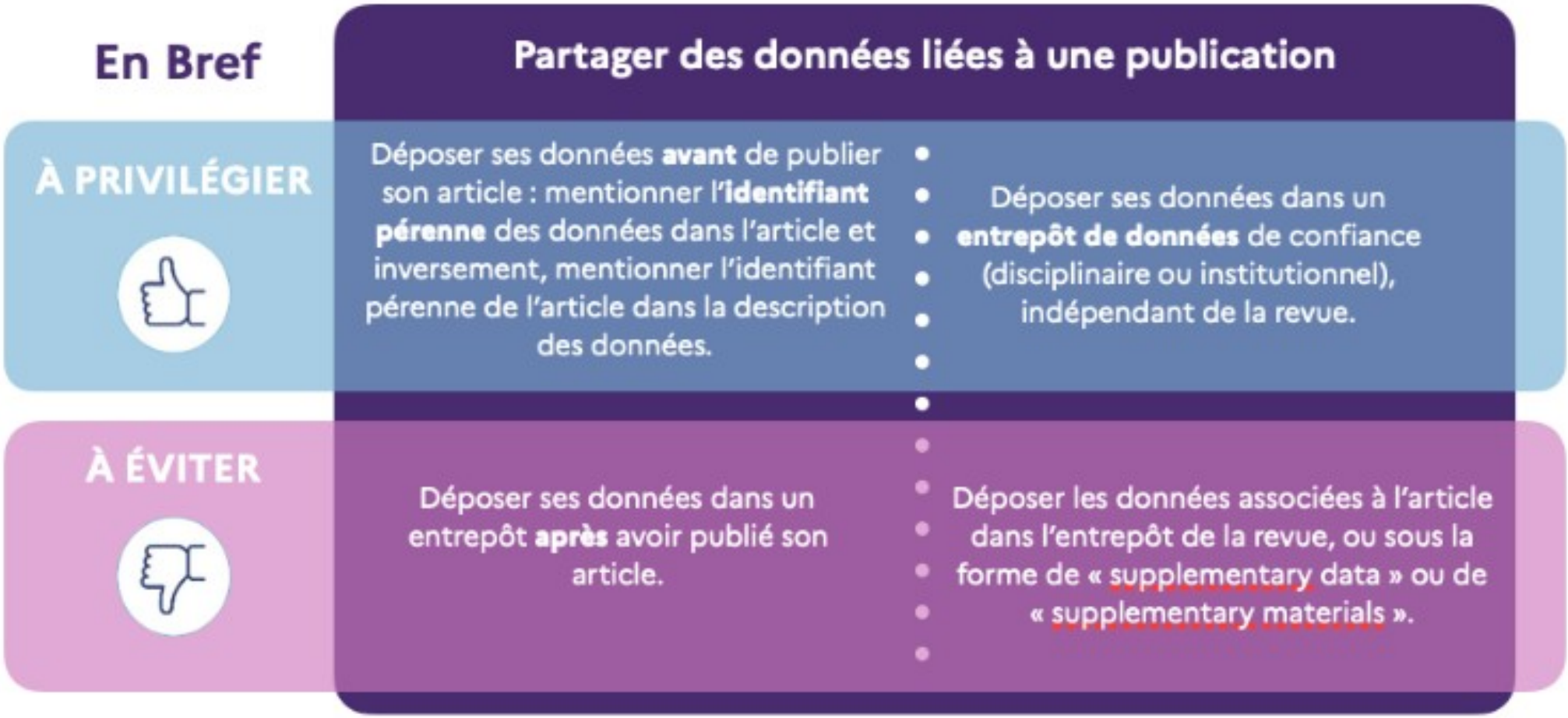
Quand décide-t-on de rendre ses données publiques ?

Il n'y a pas de règles, le mieux est d'ouvrir les données le plus tôt possible

- **Avantages** : vous êtes le premier à produire de nouvelles données
- **Inconvénient** : de nouvelles expériences peuvent confirmer ou non la qualité de vos données

- Exigence des éditeurs : rendre accessible, au moins aux reviewers, les données liées à une publication
- Un embargo peut aussi être appliqué afin de permettre un délai d'exploitation des données
- Les données sont souvent publiées au moment de la publication des résultats
- Diffuser ses données peut être une justification pour les financeurs

Quand diffuser ?



Partager les données liées aux publications scientifiques. Guide pour les chercheurs. Comité pour la science ouverte. Collège des Données de la recherche. Décembre 2025.

Où diffuser ?

Quel mode de diffusion ?

- Dépôt dans un entrepôt de données
- Publication données intégrées dans un article classique
- Publication de supplementary materials
- Publication d'un data paper

Où diffuser ? via un entrepôt

- Service en ligne permettant le dépôt, la description, 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.

Intérêt : permet

- La réutilisation d'un jeu de données (si il est bien décrit et respecte les principes FAIR)
- la citabilité d'un jeu de données (how to cite)

- Il existe des milliers d'entrepôts !

- Différents types d'entrepôts de données

voir la minute « entrepôts » (DoRaNum)

Où diffuser ? via un entrepôt

Différents types d'entrepôts :

Généralistes : Zenodo (CERN)

Nationaux : Recherche Data Gouv

Institutionnel : ESRF Data Portal

Thématiques (voir plus loin) - **Recommandation : Privilégier un entrepôt thématique** reconnu par sa communauté

Où diffuser ? via un entrepôt

Où diffuser ses données à l'UGA s'il n'existe pas d'entrepôt thématique pertinent ?

La plateforme nationale Recherche Data Gouv

Objectif :
Proposer une solution de dépôt dans un entrepôt national de confiance

Une plateforme nationale des données de recherche

- Entrepôt pour déposer et publier des données
- Catalogue pour signaler des données déposées dans des entrepôts externes (à venir)

Une fédération nationale des centres de compétences

- Géographique :centres de compétences géographiques - ateliers de la donnée
- Thématiques : centres de compétences disciplinaires
- Global : centres de ressources (DoraNum)

Où diffuser ? via un entrepôt

Liés à des éditeurs : Mendeley

voir la note : AKDOGAN, M., Bracco, L., DU PASQUIER, D., Gauvrit, G., Heude, C., Laillier, B., León y Barella, A., Rousselot, C., & Sadowska, J. (2024). Les entrepôts de données d'éditeurs commerciaux : quelle stratégie adopter ?. GTSO Données de Couperin. https://doi.org/10.5281/zenodo.14382823

« les éditeurs d'écrits scientifiques ne peuvent valablement obtenir de cessions exclusives de droits sur des données de recherche liées à la publication, qu'elles soient déposées sur un entrepôt de données ou qu'elles figurent comme supplementary materials de l'article ».

Cécile Arènes, Lionel Maurel, Stephanie Rennes. Guide d'application de la Loi pour une République numérique pour les données de la recherche. Comité pour la science ouverte. 2022. ⟨hal-03968218⟩.

Où diffuser ? via un entrepôt

Thématiques : quelques exemples …

PANGAEA (Sciences de l'environnement)

ORTOLANG (Linguistique)

Protein Data Bank in Europe (Biologie)

Nakala (Sciences humaines et sociales)

PublicNeuro (Neurobiologie)

Chemotion (Chimie)

HEPData (Physique)

Où diffuser ? via un entrepôt

Quel entrepôt choisir ?

- En premier lieu s'il en existe, dans un entrepôt disciplinaire
  - Pratique communautaire
    - Vérifier les entrepôts cités dans les publications
    - En parler avec ses encadrants
- Recommandations du financeur, de l'éditeur (listes d'entrepôts)
- Recommandations du projet de recherche, des partenaires
- Recommandations de l'établissement ou l'organisme de rattachement

Où diffuser ? via un entrepôt

Liste de critères d'exclusion d'un entrepôt
(cf note méthodologiquedu Comité Science Ouverte « Sélectionner un entrepôt thématique de confiance pour diffuser ses données »)

- Absence de modération des dépôts
- Absence d'identifiant pérenne
- Absence de garantie sur la pérennité de l'infrastructure
- Entrepôts pratiquant la cession de droits
- Politique tarifaire excessive
- **Localisation des données hors Union européenne** pour certains types de données
- Dépôt restreint par l'affiliation institutionnelle

Où diffuser ? via un entrepôt

Statut et politique de l'entrepôt

- Le statut public / privé de l'entrepôt?
- Garantie sur la pérennité de l'infrastructure
  Institution responsable de l'entrepôt ? Identifiant pérenne ?
- Lieu d'hébergement du serveur?
  Localisation des données hors Union européenne (cf données personnelles)
- Qualité de la description (qualité des métadonnées) >>>> standards ?
  Modération des dépôts
- Certification ? Reconnu par votre communauté?
- La préservation sur le long terme des données ?

Fonctionnalités

- Types de données, type de formats, volume accepté
- Gestion des versions
- Type d'accès possibles aux données (embargo / restriction de l'accès)
  Licences
- Lien avec la publication

En tout cas, **c'est à vous de décider ...**

Où diffuser ? Diffuser via un entrepôt

Des outils pour vous aider à choisir un entrepôt :

Le logigramme sur Recherche Data Gouv

Une liste (non exhaustive) mise à jour régulièrement (juin 2026) d’entrepôts thématiques de confiance

Sur Recherche Data Gouv, accès et description aux entrepôts thématiques

Pour aller plus loin :
Sélectionner un entrepôt thématique de confiance pour le dépôt de do nnées : méthodologie et analyse de l'offre existanteComité pour la Science Ouverte. 2024.

Où diffuser ? via un entrepôt

Des outils pour vous aider à choisir un entrepôt :

CAT OPIDOR – catalogue des services et entrepôts pour les données de la recherche. (Inist-CNRS)

DATACC’ – dispositif d'accompagnement à la gestion des données de la recherche en physique et en chimie. (UGA et Lyon 1)

Repositories for Sharing Scientific Data (National Institutes of Health)

Où diffuser ? via un entrepôt

Utilisation d'un **annuaire pour identifier un entrepôt dans sa discipline :**

- re3data (Registry of Research Data Repositories)
- Repository Finder(DataCite)
- FAIRsharing

(attention, ces annuaires recensent plus de 3000 entrepôts. Il peut être difficile de s'y retrouver)

Un service :
Faites appel à sos-data@univ-grenoble-alpes.fr pour vous aider à identifier un entrepôt pertinent !!

Où diffuser ? via un entrepôt

Utilisation de moteurs de recherche spécialisés pour trouver des jeux de données et donc des entrepôts :

- Datacite Commons
- Dataset Search
- OpenAire

Plus d’infos ici

Identifier des entrepôts mentionnés dans des publications

Où diffuser ? Je prépare mes données

Je prépare mes données

- 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
  - En veillant à mettre les éléments nécessaires pour les utiliser (codes, logiciels, etc.)
- Choix des formats (ouverts)
- Choix des éléments de description : métadonnées générales et disciplinaires, Readme ...
- Organisation et nommage des fichiers (arborescence)
- Choix de la licence et des modalités d'accès (ouvert, restreint, avec embargo)

Où diffuser ? Je prépare mes données

Je prépare mes données (suite)

S'être créé un compte et posséder des droits sur l'entrepôt de son choix (en faire la demande via l'adresse « support » en cas de besoin)

Avoir un ensemble de fichiers prêts

- Arborescence : organisation logique, nommage clair et sans accents,
- Présence d'un README (fichier texte permettant d'expliquer le contenu du jeu de données et d'éventuelles consignes d'utilisation)

Optionnel : Avoir à portée de main la référence de la publication à laquelle lier le jeu de données (identifiant hal, doi ..)

Pour vous aider, consulter le chapitre Comment bien gérer ses données pendant sa thèse

Où diffuser ? Je prépare mes données

Se créer un ORCID (Open Researcher and Contributer IDentifier) pour :

- désambiguïser les noms et résoudre les problèmes d'homonymie, de translittération

- permettre son identification fiable (selon les principes FAIR).

- lier ses productions

Il est unique (sur toute sa carrière)

Recommandé au niveau national, par les éditeurs, etc.

Pour créer son ORCID, voir :

- le tutoriel d'ORCID

- les recommandations du groupe national sur les consignes ORID/ROR (notamment sur la section Emploi).

Je prépare mes données (suite)

Respecter les principes FAIR (Findable, Accessible, Interoperable, Reusable)

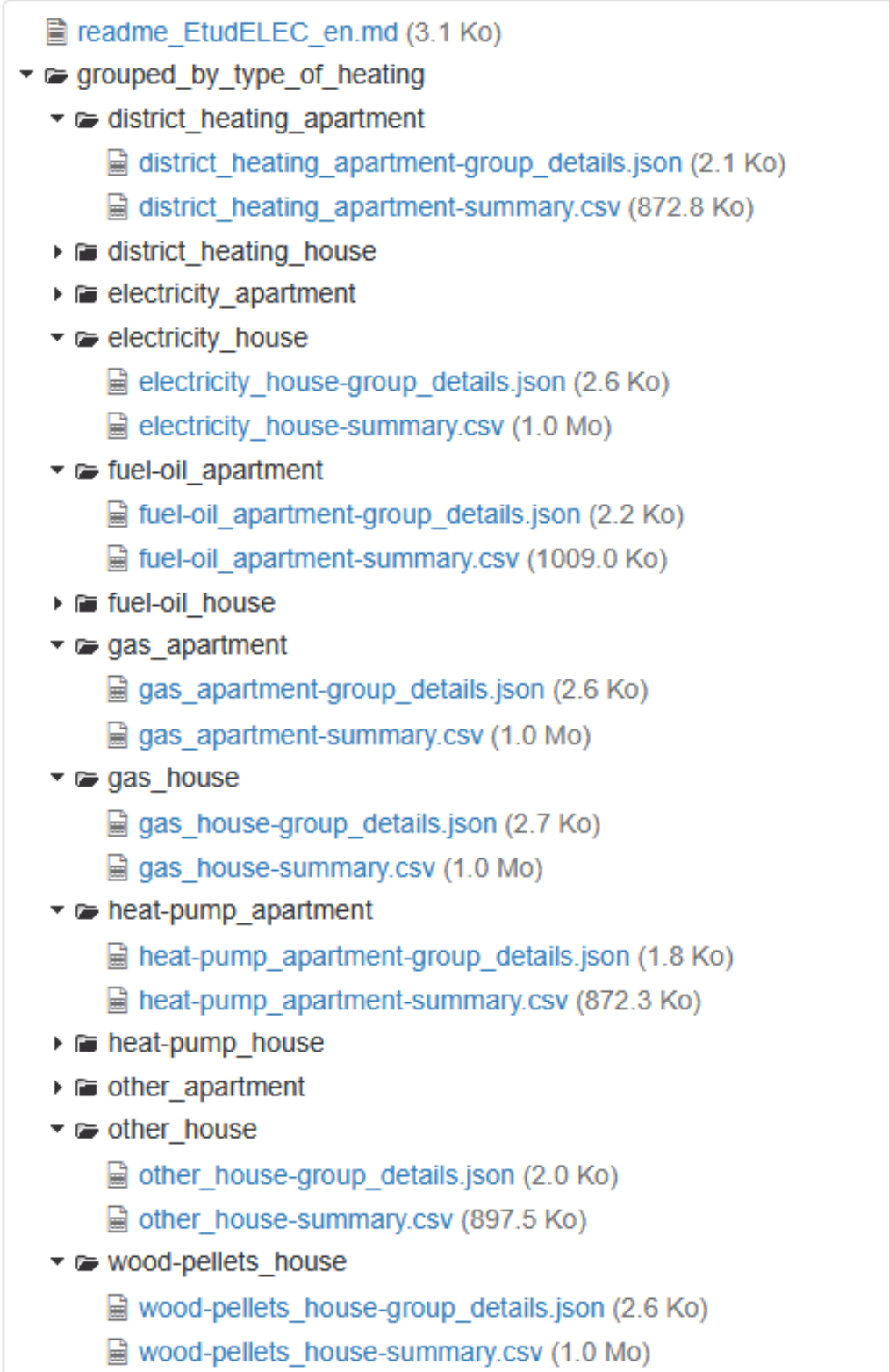
Attention à la charge de travail que représente tous ces aspects, prenez-la en compte !

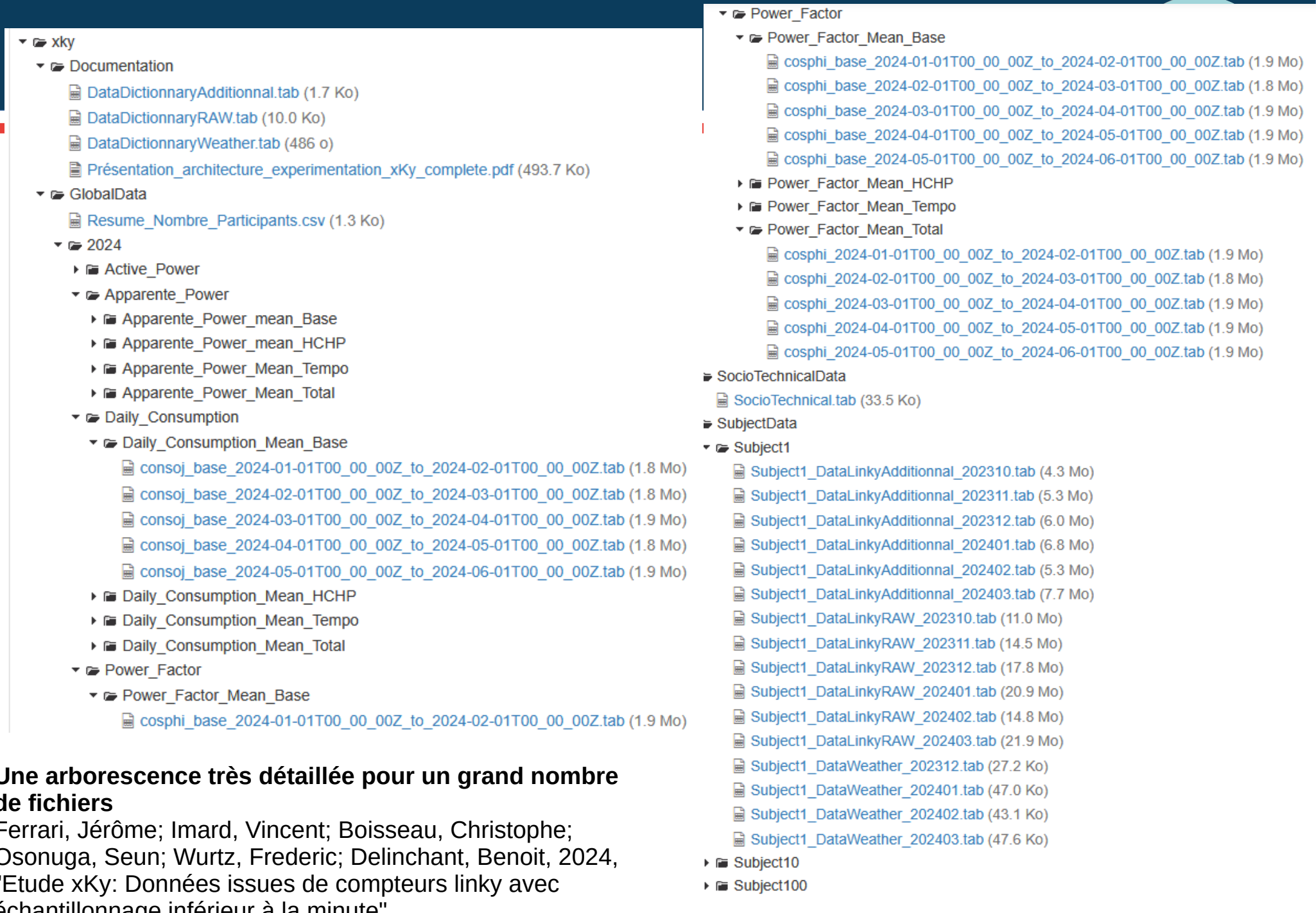
- Anticiper (description, organisation des fichiers, etc..)
- Si nécessaire, préparer une arborescence
- Rédiger un Read me

Un exemple d’arborescence

Avec un read me

Osonuga,Seun; Imard, Vincent; Shahid, Muhammad-Salman; Ferrari, Jerome; Boisseau, Christophe; Wurtz, Frederic; Delinchant, Benoit; Llerena, Daniel; Roussillon, Béatrice; Fadhuile, Adélaïde, 2024, "EtudELEC Data, aggregated electricity consumption data from 400+ residential customers in France", https://doi.org/10.57745/WIWMMK, Recherche Data Gouv, V1





Une arborescence très détaillée pour un grand nombre de fichiers

Ferrari, Jérôme; Imard, Vincent; Boisseau, Christophe; Osonuga, Seun; Wurtz, Frederic; Delinchant, Benoit, 2024, "Etude xKy: Données issues de compteurs linky avec échantillonnage inférieur à la minute", https://doi.org/10.57745/SHDDTW, Recherche Data Gouv, V2, UNF:6:MrdU3BHudurwidX1/uYyKw== [fileUNF]

Où diffuser ? Je prépare mes données

Un exemple avec une répartition selon la durée du phénomène

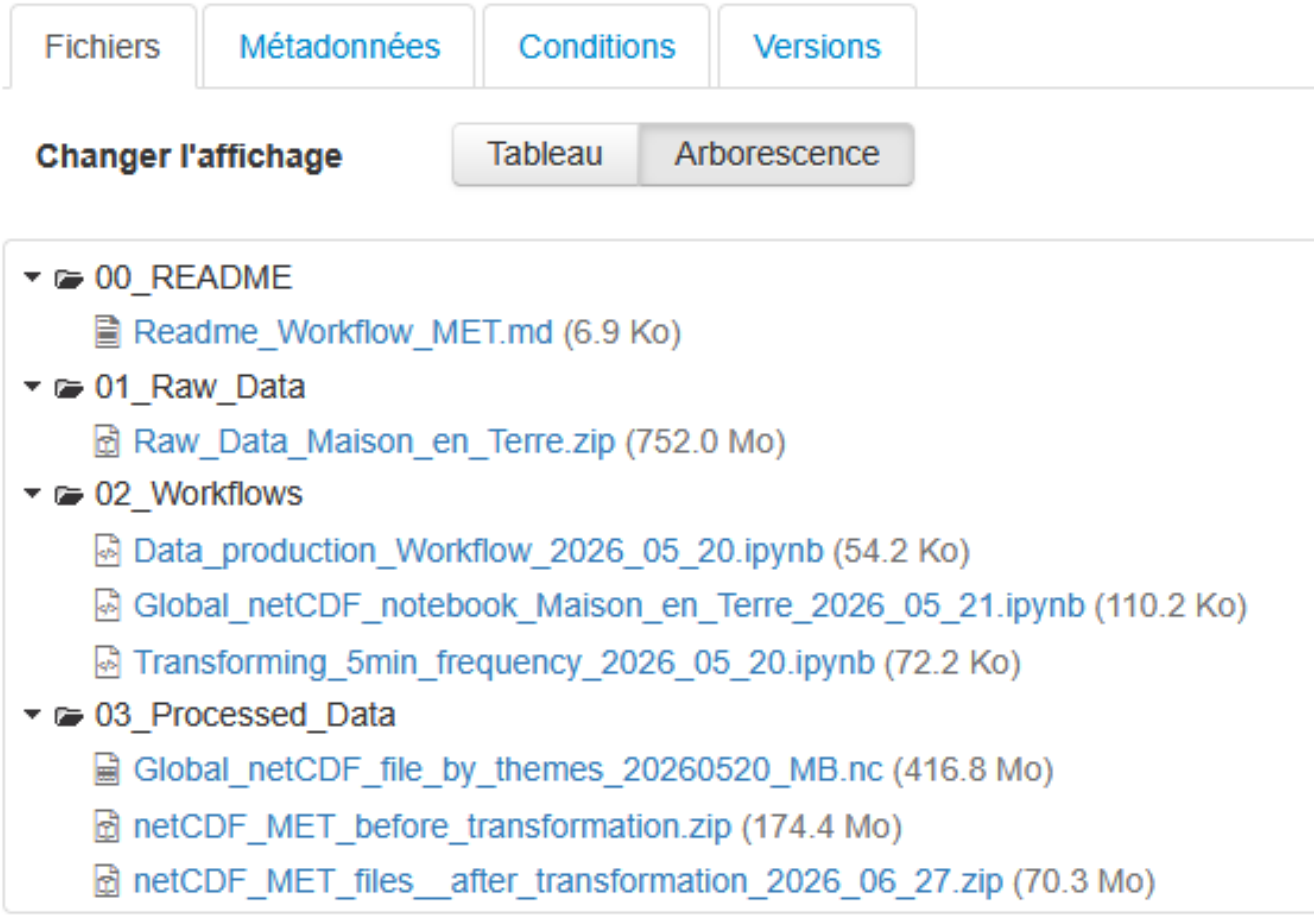
Fagour, Clément; Mignot, Emmanuel; Proust, Sébastien, 2025, "Laboratory experiment on the pollutant dynamics inside a porous city block during urban flooding under steady flow regime", https://doi..org/10.57745/XOEQ2Z, Recherche Data Gouv, V1, UNF:6:ZoxLekxCy9KI4NXOYr4Kzg== [fileUNF]



Où diffuser ? Je prépare mes données

Un exemple avec une distinction entre données brutes et données traitées

Daudon, Dominique; Blaise, Marine, 2025, "Raw and processed data obtained from Maison en Terre, alongside workflows.", https://doi.org/10.57745/OSEHHA, Recherche Data Gouv, V2



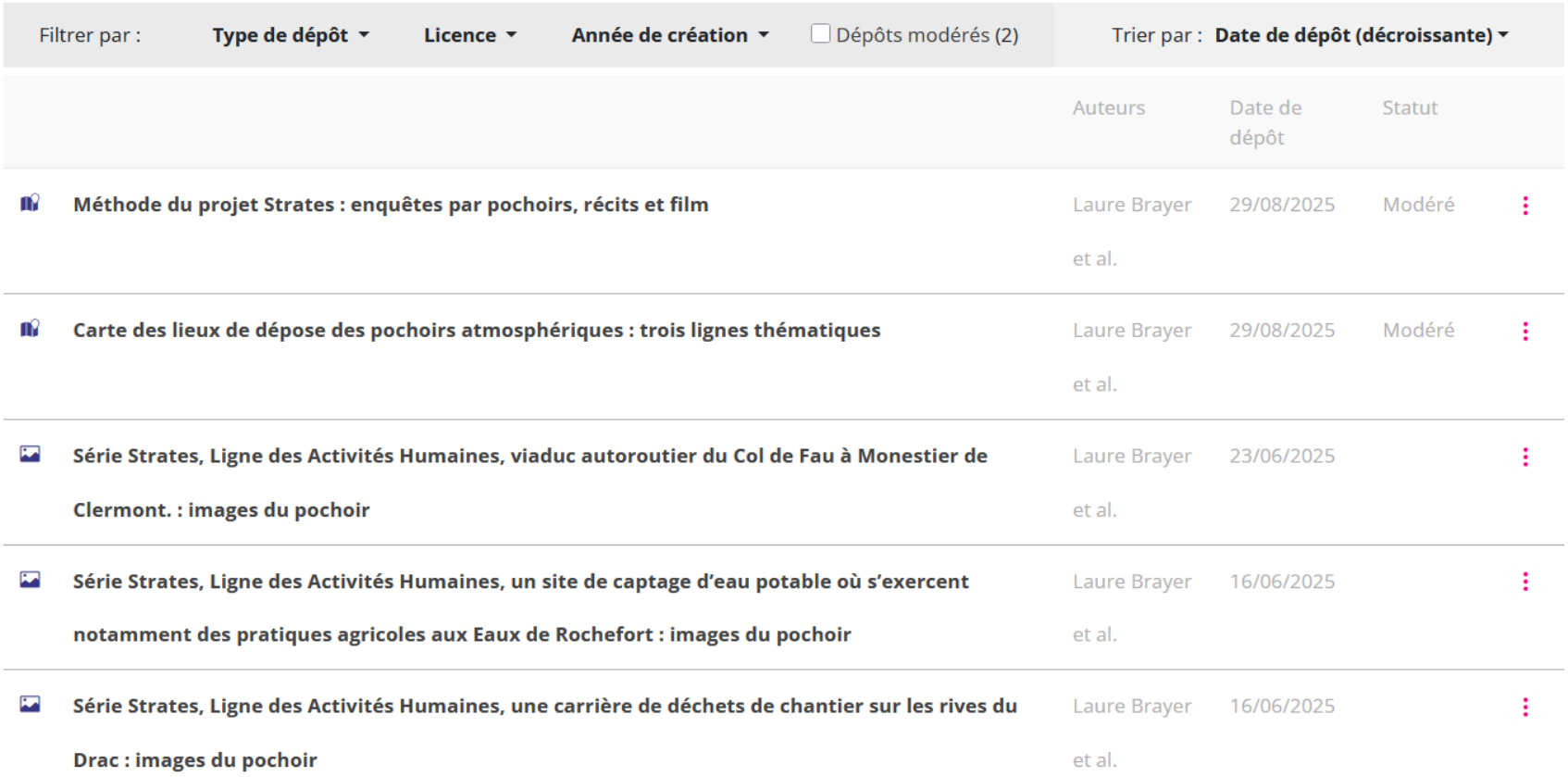
Où diffuser ? Je prépare mes données



Équipe du projet IThAC, 2025, "Données XML-TEI de "L'invention du théâtre antique dans le corpus des paratextes savants du XVIe siècle. Analyse, traduction, exploration numérique"", https://doi.org/10.57745/TYD3GW, Recherche Data Gouv, V2

Où diffuser ? Je prépare mes données

SENSIBILIA. Chantier "STRATES". Ligne des activités humaines
ID : 10.34847/nkl.5aa7tv3j
Publique
Créée le 20/06/2025



Dépôt sur Recherche Data Gouv

Espace UGA sur Recherche Data Gouv : Data Repository Grenoble Alpes

Organisation en collections par laboratoires, projets de recherche, etc

Tous les types de données sont acceptés

Taille max des jeux de données : 50 Go par fichier.

Accès restreint possible pour certains ou tous les fichiers d'un jeu de données.

Possible de créer une url privée pour un jeu de données non publié, par exemple pour de la relecture par les pairs

Modération assurée

Rappel : **privilégier un entrepôt thématique** reconnu par sa communauté

Dépôt sur Recherche Data Gouv

Des exemples de jeux de données :

- ExpeSmartHouse - Données d'une maison connectée
- Cross-referenced data on electoral disputes and French legislative election results
- Long term monitoring of near surface soil temperature in the Ecrins National Park
- Performance d'élèves du CE2 à la 3e concernant des savoirs grammaticaux et orthographiques fondamentaux, autour du verbe et de l’adjectif",
- Raw and processed data obtained from Maison en Terre, alongside workflows.

Une collection : le chantier scientifique Notre Dame de Paris

..déclinée en 10 sous collections dont

- GT données numériques
- GT métal
- GT décors monumentaux

Où diffuser ? via un entrepôt

Dépôt dans Recherche Data Gouv (suite)

Pour vous aider :
- classes virtuelles
- tutoriels
- guides
- FAQ
- Une fiche chrono-dépot

Un modèle de Readme

Où diffuser ? via un entrepôt

Le bac à sable de Recherche Data Gouv : https://demo.recherche.data.gouv.fr/

Déposons un jeu de données

Où diffuser ? via un entrepôt



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

Des centres de compétences pour accompagner les équipes de recherche sur toute question relative aux données.

Une plateforme de données
pour déposer, publier et signaler des données de recherche.

En savoir plus sur Recherche Data Gouv

Rechercher, déposer, publier un jeu de données

Je ne sais pas où déposer et publier mes données

Bienvenue ! Pour déposer des données, rendez-vous dans votre espace institutionnel ou dans l'espace générique, puis connectez-vous via le menu en haut à droite. (Vous ne savez pas où déposer ? Cliquez ici.)



Recherche Data Gouv

Génération datapaper

(Recherche Data Gouv)

Statistiques | 5 185 698 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 sovereign solution for sharing and opening up data produced by communities that do not have a recognized disciplinary repository. It is based on the Dataverse software. Data should be deposited in the space assigned to an institution which one of the contributor belongs to. A generic space is available for data produced by institutions which do not such a space yet themselves. Tests should be performed in the sandbox.
The latest news and events on the *Recherche Data Gouv* platform.

Lire la suite de Description [+]

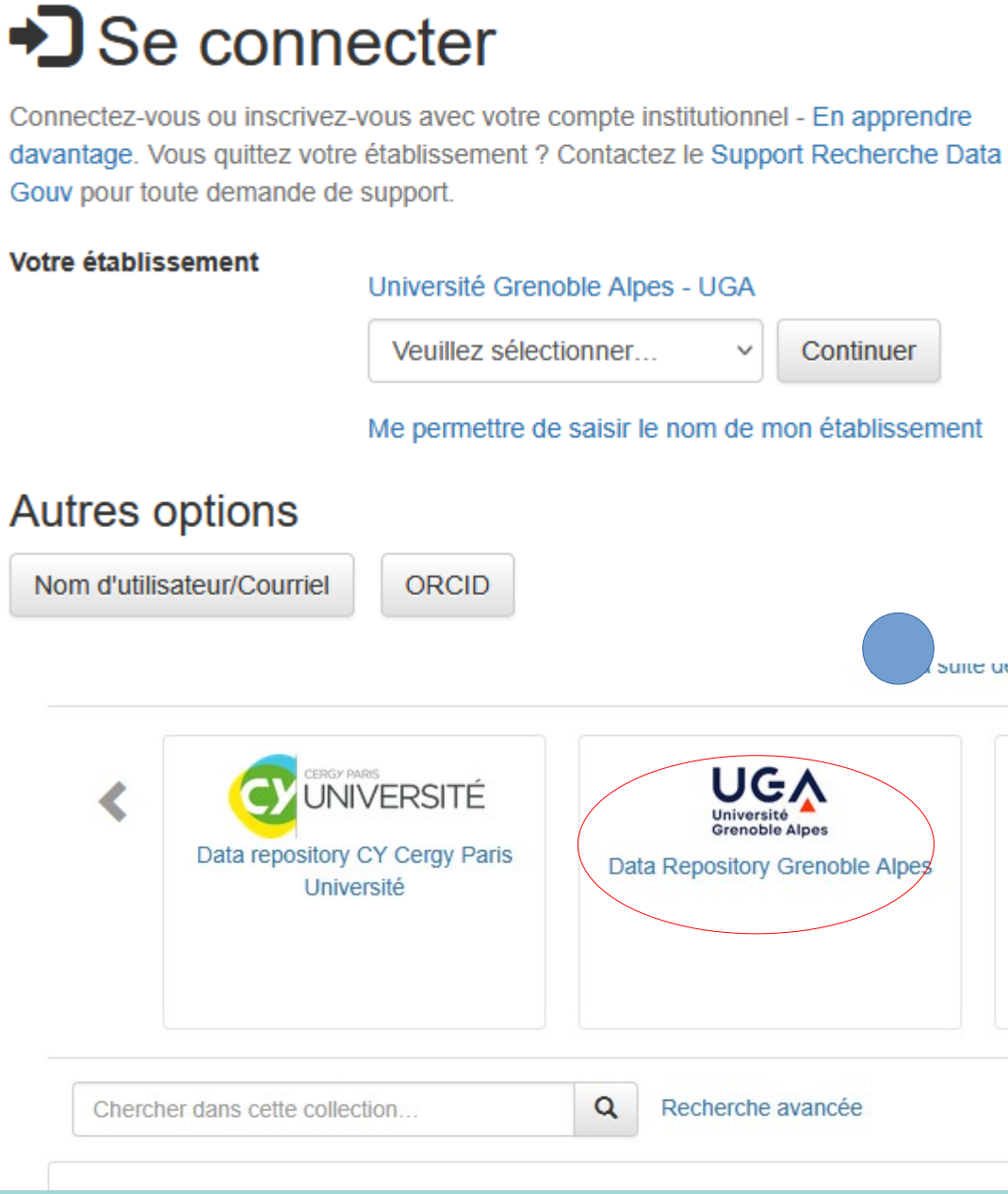
recherche.data.gouv.fr
Espace Générique

Arts et Métiers Institute of Technology
Arts et Métiers Sciences et technologies - Arts et Metiers Institute of Technology

ASNR
Autorité de Sûreté Nucléaire et de Radioprotection

cea
CEA Scientific and Technical Data

Où diffuser ? via un entrepôt



Où diffuser ? via un entrepôt

Lire la suite de Description [+]

Chercher dans cette collection...

Recherche avancée

+ Ajouter des données

Collections (18)

Jeux de données (129)

Fichiers (8 337)

Catégorie de collection

Projet de recherche (8)
Laboratoire (7)
Groupe de recherche (2)
Organisation ou établissement (1)

Année de publication

2026 (5)
2025 (47)
2024 (26)
2023 (28)
2022 (12)

Statut de publication

Publié (98)
Version provisoire (43)
Non publié (31)
Retiré (4)
En révision (1)

Origine des données

observational data (33)
experimental data (20)
analysis data (19)
computer code (19)
simulation data (7)

1 à 10 de 147 résultats

Tri

Altimpact dispositif citoyen pour la sobriété numérique

3 juin 2026

Jeu de donnée

Berthoud, Françoise; Delmas Julie; Marie, Camille, 2026, "Altimpact dispositif citoyen pour la sobriété numérique", https://doi.org/10.57745/KDYY0D, Recherche Data Gouv, V1

Ce dépôt inclut l'ensemble des documents associés au dispositif citoyen sur la sobriété numérique qui a été déployé en France entre l'été 2024 et l'été 2026 : les rapports de proposition de chacune des six conventions délibératives locales à destination de : la Métropole de Greno...

SPER01 Plant Metabarcoding: DNA Sequencing Data from Saint-Martin-d'Hères Campus Version provisoire Non publiée

29 mai 2026 - CAMPUS-SOL

Blaise, Marine; Boyer, Frédéric; Pompanon, François, 2026, "SPER01 Plant Metabarcoding: DNA Sequencing Data from Saint-Martin-d'Hères Campus", https://doi.org/10.57745/LNZOAK, Recherche Data Gouv, DRAFT VERSION, UNF:6:I55MAqVcUeM8+St5lxJwlg== [fileUNF]

Raw sequencing data, processed sequencing and physico-chemical data, taxonomy and OTU files, plus metadata. SPER01 primer targets Spermatophyta via P6 region in trnL chloroplastic intron. For sequencing data and replication, you can either start from raws or demultiplexed files....

Raw and processed data obtained from Maison en Terre, alongside workflows.

28 mai 2026 - E-colos project - An ecological innovation to monitor locally sustainable buildings

Daudon, Dominique; Blaise, Marine, 2025, "Raw and processed data obtained from Maison en Terre, alongside workflows.", https://doi.org/10.57745/OSEHHA, Recherche Data Gouv, V2

Raw and processed data files for Maison en Terre,. Raws are in *.CSV format, and processed data are in *.nc format. Contains all 61 distincts measurments, with outliers removed. After processing, data is fully organized by themes, for better replication and understanding of every...

CAMPUS-SOL (Univ. Grenoble Alpes, CNRS, IRD, INRAE, Météo France, OSUG, 38000 Grenoble, France) Non publiée

6 mai 2026 Observatoire Junior

Collection

Où diffuser ? via un entrepôt

Modèle de jeu de données

Modifier le modèle effacera tous les champs dans lesquels vous auriez entré des données.

Recherche Data Gouv V2

Langue des métadonnées du jeu de données

Sélectionner la langue que vous souhaitez utiliser pour entrer les métadonnées.

English

*Les astérisques indiquent les champs obligatoires.

Métadonnées générales

Titre *

Ajouter « Données de réplication pour » au titre

Lien vers les données

Adresse URL, commençant par https://

Point de contact *

Nom

Albaret, Lucie

Affiliation

Université Grenoble Alpes - UGA

Courriel *

lucie.albaret@univ-grenoble-alpes.fr

Auteur *

Pour chaque coauteur renseigner l'affiliation complète, contenant unité et tutelle(s), par exemple sous la forme : Unité ; Tutelle1, Tutelle2, etc. ; PAYS

Nom *

Albaret, Lucie

Affiliation

Université Grenoble Alpes - UGA

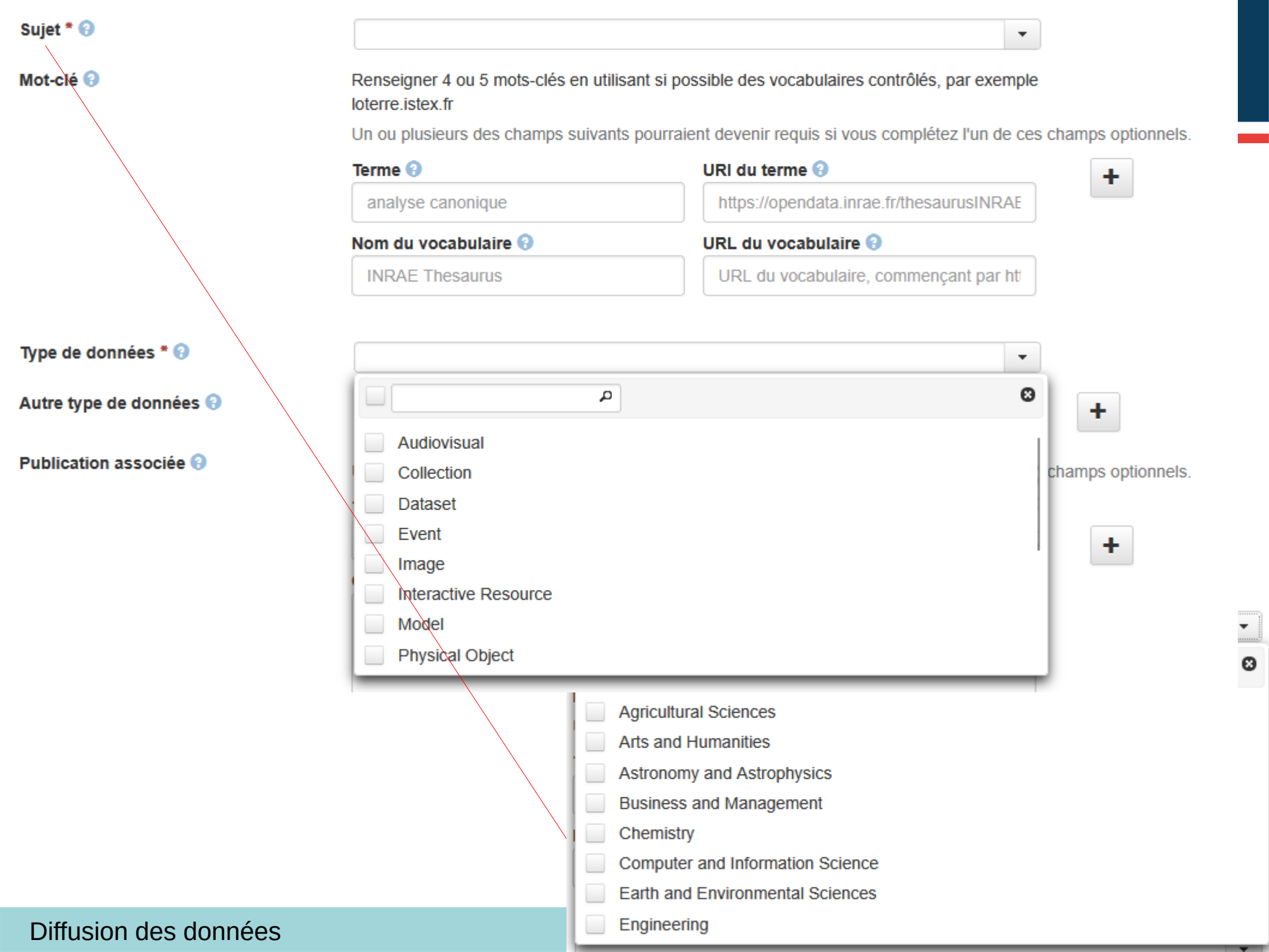
Type d'identifiant

Sélectionner...

Identifiant

ex. pour ORCID : "0000-0000-0000-000(

Ouverture du formulaire : champs généraux


Sujet *
Mot-clé
Renseigner 4 ou 5 mots-clés en utilisant si possible des vocabulaires contrôlés, par exemple loterre.istex.fr
Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.
Terme
analyse canonique
URI du terme
https://opendata.inrae.fr/thesaurusINRAE
Nom du vocabulaire
INRAE Thesaurus
URL du vocabulaire
URL du vocabulaire, commençant par ht
Type de données *
Autre type de données
Publication associée
champs optionnels.
Audiovisual
Collection
Dataset
Event
Image
Interactive Resource
Model
Physical Object
Agricultural Sciences
Arts and Humanities
Astronomy and Astrophysics
Business and Management
Chemistry
Computer and Information Science
Earth and Environmental Sciences
Engineering
Diffusion des données

Producteur

Tout organisme détenant une responsabilité sur les données, généralement il s'agit des tutelles.

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Nom: Université de Groningen

Affiliation: Organization XYZ

Nom abrégé: RUG

URL: Adresse URL, commençant par https://

URL du logo: URL de l'image, commençant par https://

Distributeur

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Nom: Entrepôt Recherche Data Gouv

Affiliation: Organization XYZ

Nom abrégé:

URL: https://entrepot.recherche.data.gouv.fr

URL du logo: https://s3.fr-par.scw.cloud/rdg-portal/logo

Description *

Si la description est en français, ajouter une seconde description en anglais.

Ce champ prend en charge uniquement certaines balises HTML.

Texte *

Langue: Sélectionner...

Date: AAAA-MM-JJ

Publication associée

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Type de relation

Sélectionner...

- Sélectionner...
- Est cité par
- Cite
- Est un supplément de
- A pour supplément

Type d'identifiant

doi

Identifiant

ex. pour DOI : "10.15454/AEIOUY"

URL

Adresse URL, commençant par https://

Informations sur la subvention

Agence

Identifiant

Ex. ANR-XX-XXXX-XXXX

Informations sur le projet

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Identifiant

Ex. ANR-XX-XXXX-XXXX

Acronyme

EOSC-Pillar

Titre

Coordination and Harmonisation of Natio

Lien vers le site web

Adresse URL, commençant par https://

Work Package

Tâche

Déposant

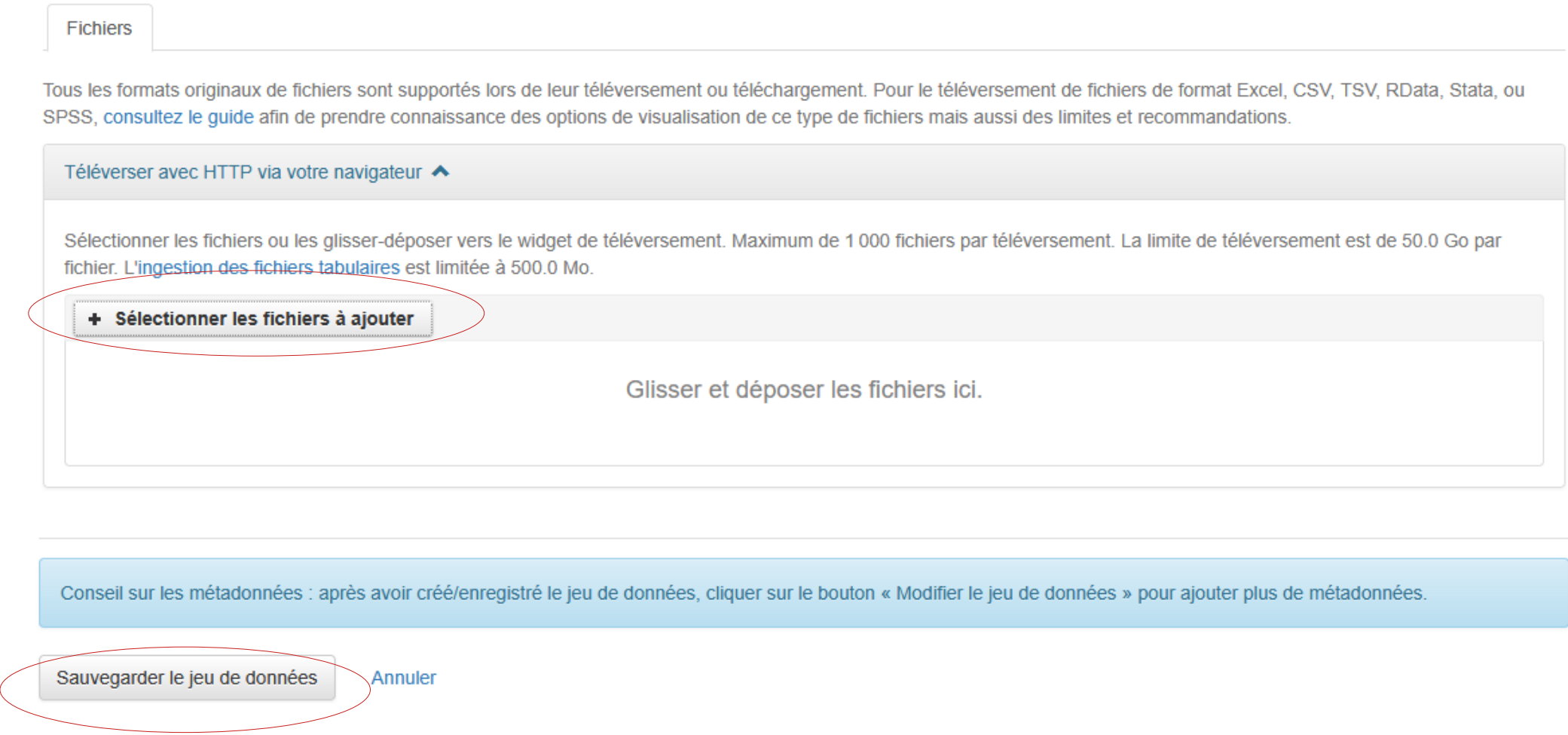
Albaret, Lucie

Date de dépôt

2026-06-08

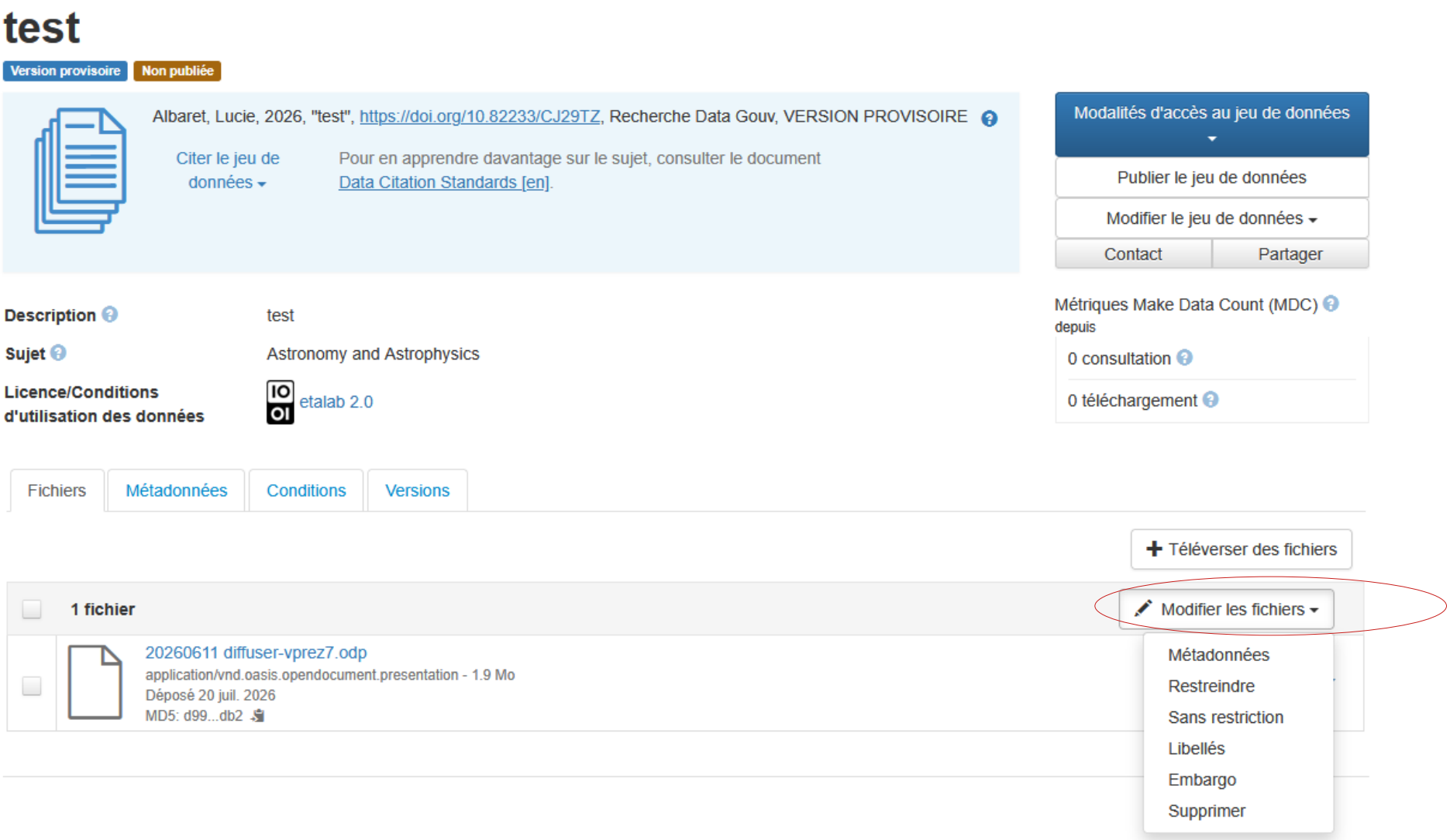
Où diffuser ? via un entrepôt

Ajout des fichiers et sauvegarde du jeu de données



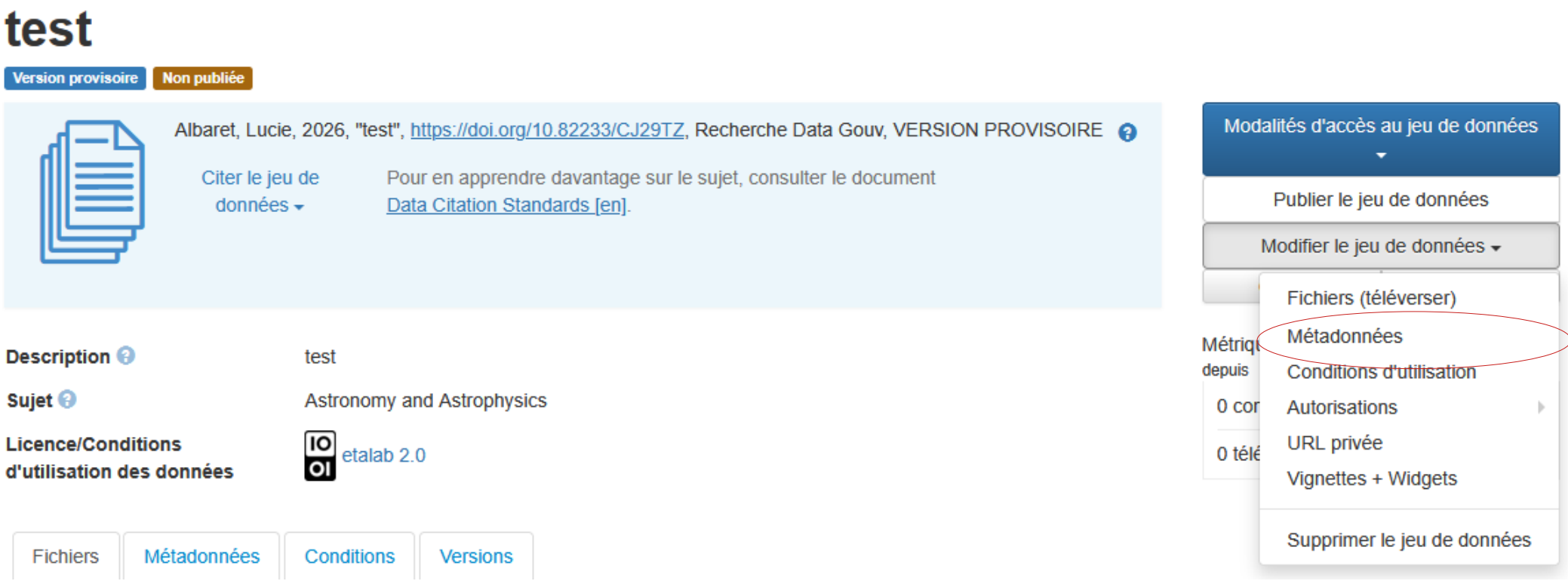
Où diffuser ? via un entrepôt

Après la sauvegarde, possibilité d'ajouter/modifier les fichiers (modalités d'accès)



Où diffuser ? via un entrepôt

Après la sauvegarde, possibilité (et recommandation !) d'ajouter/compléter les metadonnées



Où diffuser ? via un entrepôt

Modifier le jeu de données >>> métadonnées

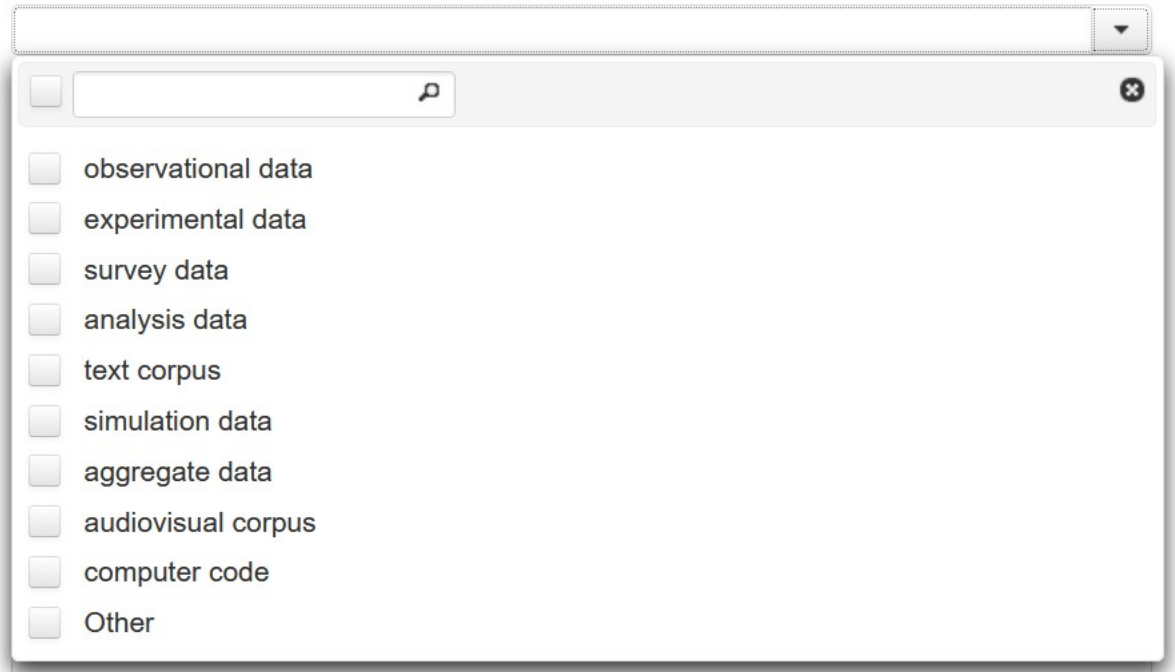
De nouveaux champs « génériques » apparaissent
- origine des données (menu déroulant : par ex, données expérimentales)
- source des données (identifiant pérenne ou citation)
- origine des sources historiques
- caractéristiques des sources

- logiciels
- série (par ex, série temporelle de données climatique)
- étape du cycle de vie (menu déroulant, par ex data processing)
- notes

Ressources associées :
- publications
- autres jeux de données

- période couverte
- date de collecte

Où diffuser ? via un entrepôt



Où diffuser ? via un entrepôt

Logiciel

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces cha

Nom R

Version 4.1.0

Séries

Nom

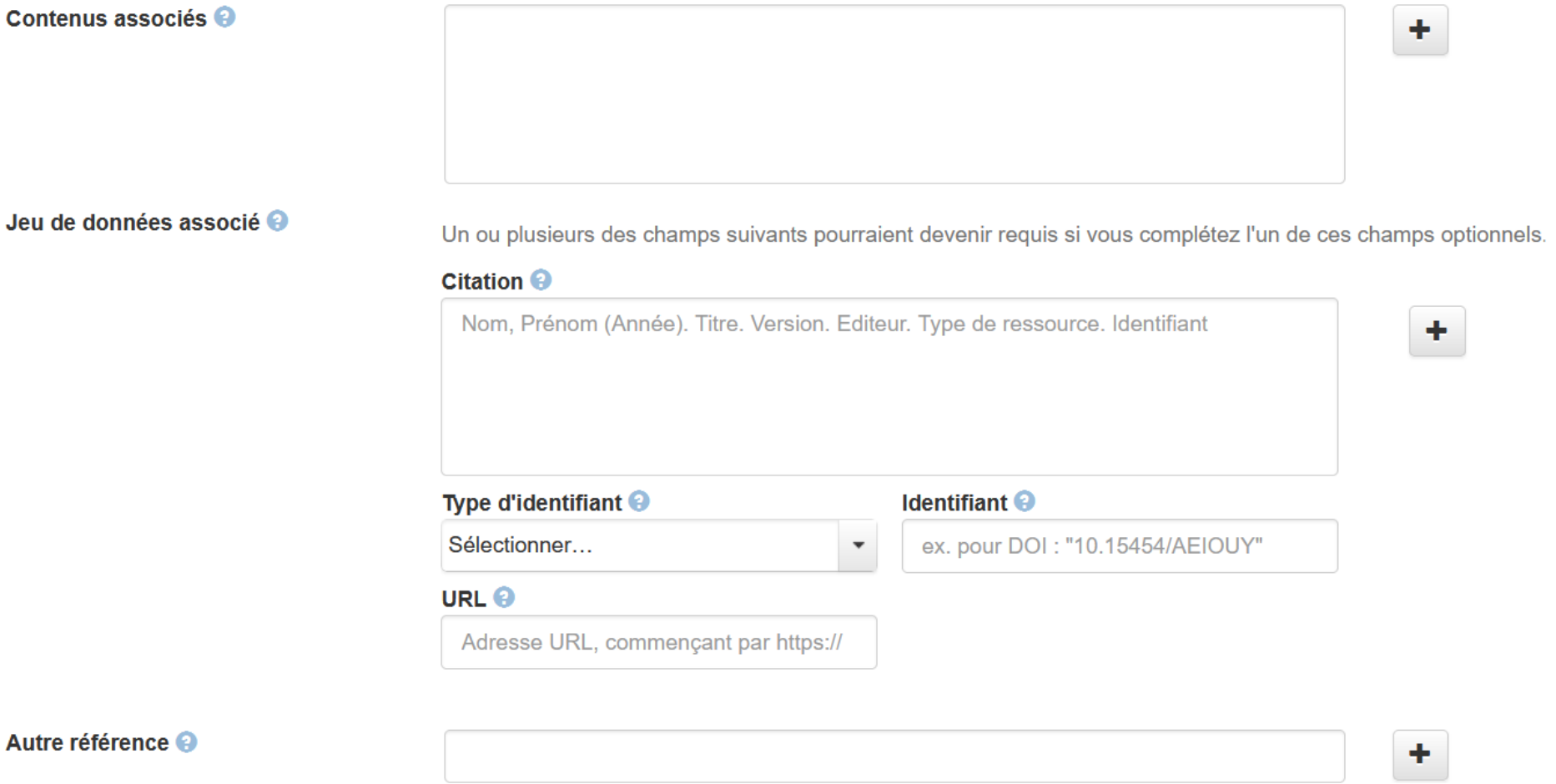
Informations

Étape du cycle de vie

Notes

Où diffuser ? via un entrepôt

En plus des publications, il est possible d'associer d'autres contenus :



Où diffuser ? via un entrepôt

Période couverte

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Début AAAA-MM-JJ **Fin** AAAA-MM-JJ +

Date de collecte

Un ou plusieurs des champs suivants pourraient devenir requis si vous complétez l'un de ces champs optionnels.

Début AAAA-MM-JJ **Fin** AAAA-MM-JJ +

Déposant Albaret, Lucie

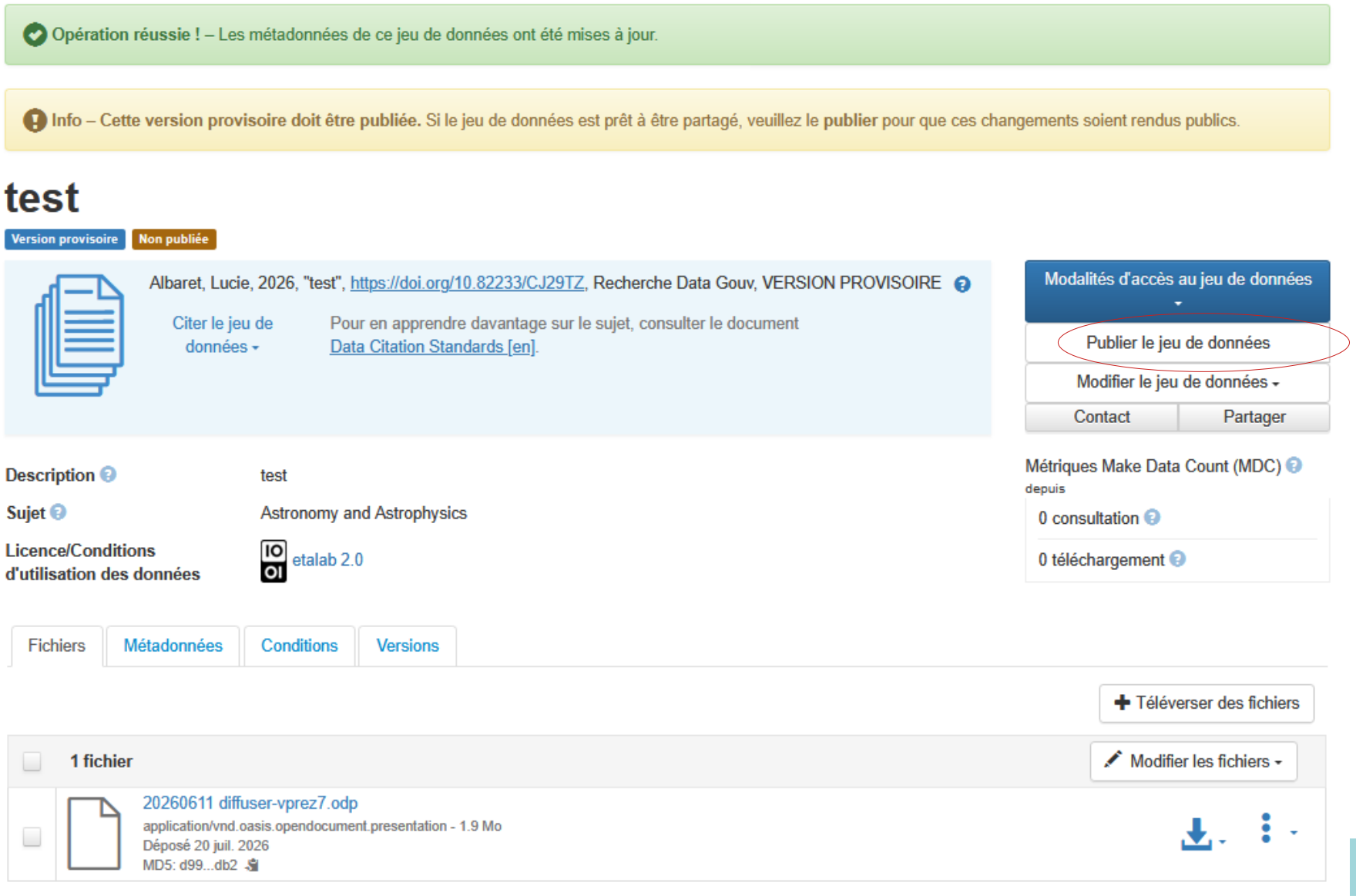
Où diffuser ? via un entrepôt

Modifier le jeu de données >>> métadonnées (suite) : Des métadonnées spécifiques/thématiques à dérouler



Où diffuser ? via un entrepôt

Après avoir complété les métadonnées, on peut publier le jeu de données en vue de sa modération



Où diffuser ? via un entrepôt



Un doi est alors attribué.
Le jeu de données est alors soumis à la modération (qui sera effectuée par la cellule Codes Données Grenoble Alpes)

Focus Codes

Si vous développez-vous des codes, des scripts ….

A noter : la formation Codes et logiciels : bonnes pratiques depuis leur création pour la gestion, la diffusion, l'archivage : support 2026 disponible sur E-formation

Focus Codes

- Archivage (automatique) dans Softwar e Heritage
- Signalement dans HAL

Fiche pratique
Webinaire (support et vidéo)

Documentation HAL : déposer le code source d'un logiciel

INIST (et al.) (2025). DoRANum-Accès et visualisation : Préserver et diffuser les codes sources des logiciels de recherche. DoRANum. https://doi.org/10.13143/657A-ZK37

Pensez au ReadMe et aux licences

J'ouvre mon code en 4 étapes

1. **J'intègre 4 fichiers**

README.md	→	Description et liens vers les documentations
AUTHORS.md	→	Les auteurs et les contributeurs du logiciel
LICENSE.txt	→	Je veux un **copyleft = GNU GPL** Je veux une **licence permissive = MIT licens**
CODEMETA.md	→	À générer automatiquement avec le **CodeMetagenerator de github**

2. **Quand je code**
   - Je **documente et commente mon cod** pour moi et pour les autres
   - J'utilise une **forge** comme le Gitlab de Gricad
3. **J'archive mon code sur Software Heritage**
4. **Je signale mon code sur HAL grâce au SWHID**

Mon code est ouvert

Infographie : Nicole Lambert /Gricad/CNRS

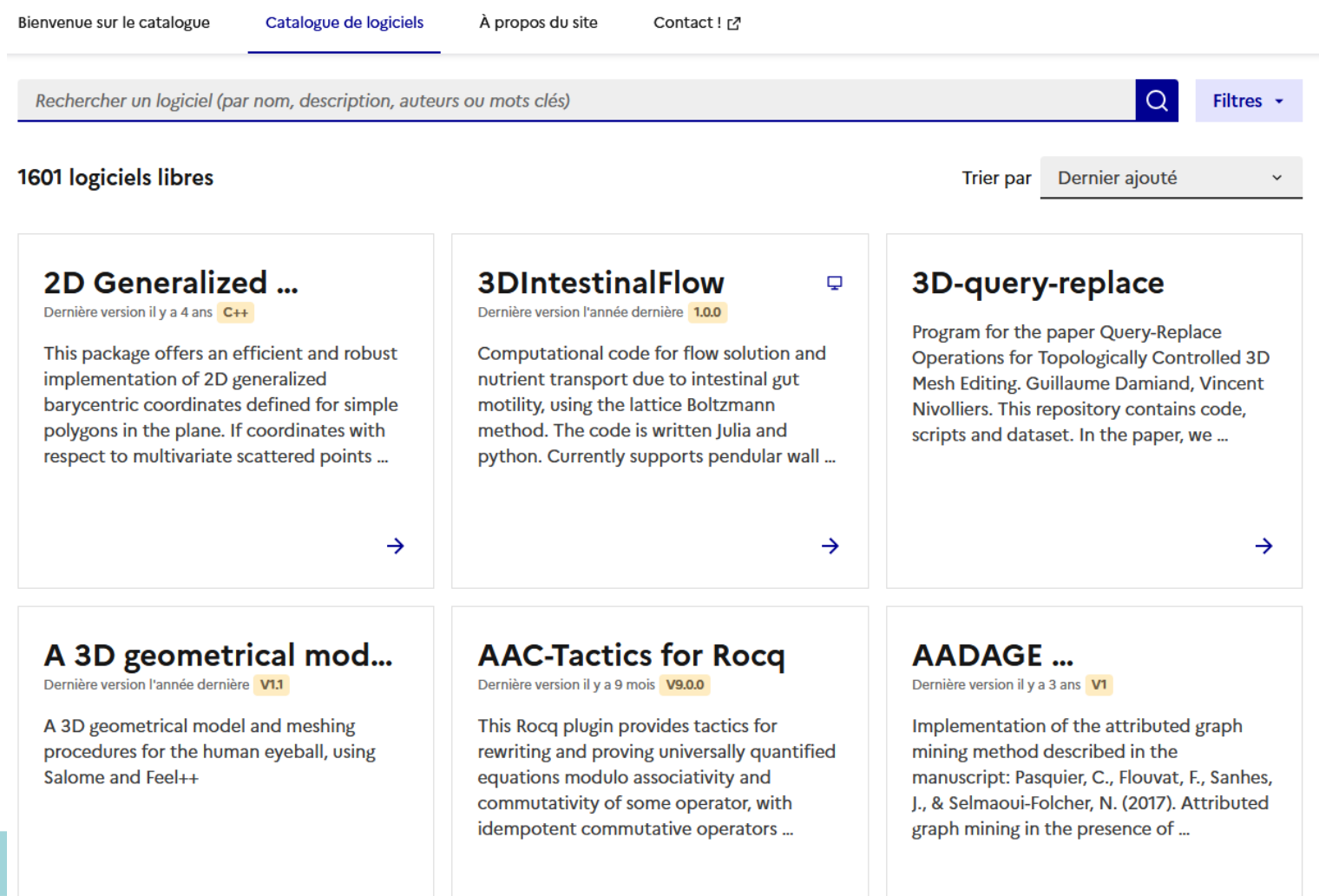
Contacts

Pour plus d'informations sur la diffusion des codes, contacter la Cellule Data Université Grenoble Alpes :

sos-codes-recherche@univ-grenoble-alpes.fr

Focus Codes

L’archivage (automatique) dans Software Heritage et le signalement dans HAL permet d’alimenter le catalogue des logiciels libres de l’ESR (version beta)



Où diffuser ? via une publication

3 manières de publier des données

- Les inclure dans un article (données intégrées ou embedded data)
- Les assembler en annexe dans un matériel supplémentaire (« supplementary materials »)
- Les publier dans un data paper (data article, data descriptor).

Où diffuser ? via une publication

Focus sur les datapapers

Définition :
Les data papers (data articles / data descriptors) sont des articles qui ont pour but de décrire un ou plusieurs jeux de données, plutôt que des résultats d'analyse.
Les data papers peuvent paraître dans des revues classiques ou dans des revues spécifiques « data journals ».

Pratique :
Les données peuvent être déposées dans un entrepôt, recommandé par l'éditeur ou au choix de l'auteur

Voir la minute « Publier un data paper », DoRaNum

Où diffuser ? via une publication

Focus sur les datapapers

Enjeu :

Le datapaper valorise les données en exposant leur potentiel pour de nouveaux usages

Il facilite la réutilisation des données en mettant en évidence la qualité des données et des procédures, ainsi que la rigueur scientifique de l'étude.

Il apporte de la visibilité aux données, les rend plus facilement repérables et citables par d'autres études.

Le datapaper est examiné par les pairs

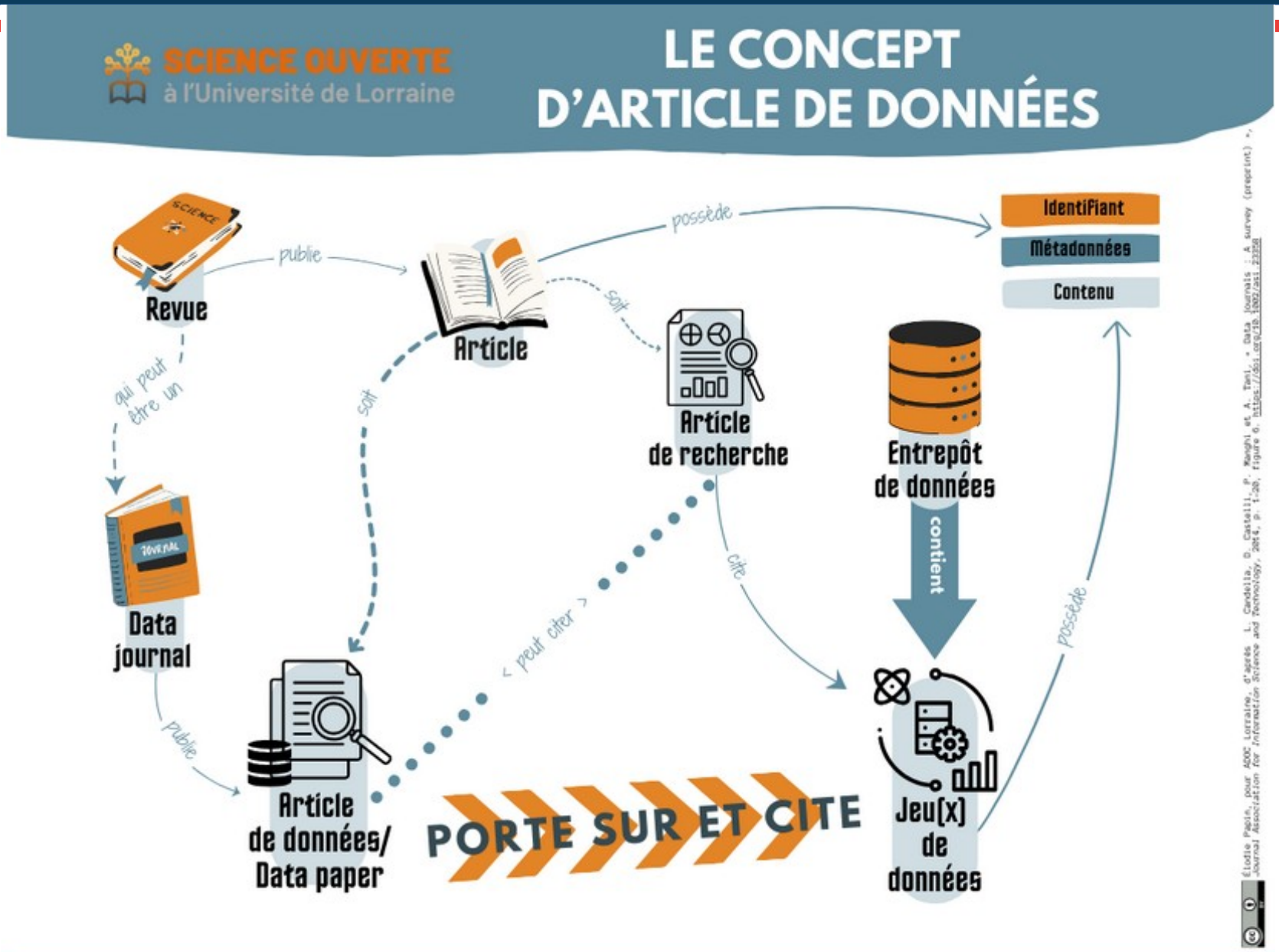
Voir : Urfist Méditerranée & INRAE (2024). **DoRANum-Data papers et data journals : Les critères d'évaluation des data papers**. DoRANum. https://doi.org/10.13143/DK71-W757

Recommandation :

Déposez dans HAL le post-print de la publication

(voir Dedieu, L. 2014. Rédiger et publier un data paper dans une revue scientifique, en 5 points. Montpellier (FRA) : CIRAD, 7 p. https://doi.org/10.18167/coopist/0057)

Où diffuser ? via une publication



Papin, E. (2023). Le concept d'article de données. Atelier de la donnée ADOC Lorraine. https://doi.org/10.5281/zenodo.10254897

Où diffuser ? via une publication

Focus sur les datapapers

Quelques exemples de datajournals :
- Data Science Journal – codata
- Scientific Data (Nature)
- Journal of Open Humanities Data(JOHD)
- Data In Brief (Elsevier)
- Data & Corpus (Univ Lorraine) - SHS

Quelques listes :
La base « Où publier » du Cirad (sélectionner data papers)
liste « Revues publiant des Data papers et des Software papers » (CIRAD)
Datacc (Physique-Chimie)
Possible dans de nombreuses revues « classiques » (PlosOne, CyberGeo …)

Comment choisir ? S'informer sur :
Son modèle économique (accès ouvert ? APC?)
Ses modalités de diffusion (dépôt dans un entrepôt possible?)
Ses contraintes juridiques (cession de droit exclusive? Possibilité de mettre des licences CC ?)

Où diffuser ? via une publication

« Le data paper est généralement structuré en 3 points :

- description du jeu de données et de son contexte,
- description des méthodes d'obtention,
- explication du potentiel de réutilisation des données. »

(voir Dedieu, L. 2022. *Publier un Data paper, en 5 points*. Montpellier (FRA) : CIRAD, 5 p. https://doi.org/10.18167/coopist/0057)

Exemple :

(Scientific Data)	*(JOHD)*
Methods	1- overview
Data Records	2 Method
Data Overview (optional)	3 Dataset Description
Technical Validation	4 Reuse Potential
Usage Notes (optional)	
Data Availability Code Availability	

Où diffuser ? via une publication

Focus sur les datapapers

Dans tous les cas, se référer :
- aux guides/instructions des éditeurs à destination des auteurs
- aux templates mis à disposition par les éditeurs

Exemples :
- Template for data descriptor pour Scientific Data (Nature Publishing Group – Overleaf)
- Le template de Data in Brief
- Journal of Open Humanities Data (JOHD) data paper template (open humanities data)
- L'outil arpha Writing Tool.
- L'outil de génération de datapaper à partir du doi d'un jeu de donnée sur Recherche Data Gouv

Voir Urfist Méditerranée. (2018). **DoRANum-Data papers et data journals : Le contenu d'un data paper. DoRANum**. https://doi.org/10.13143/8R3D-K505

Où diffuser ? via une publication

5 exemples de datapapers :

- Nardy, A., Rousset, I., & Gaubil, R. (2025). Recueil non supervisé de données audio massives à l'école maternelle : de l'enregistrement au développement d'un logiciel d'enrichissement et d'analyse (Version 2, Vol. 1042, Issue 166). Data & Corpus. https://doi.org/10.46298/dc.15849
- Culos M, Ouvrier A, Lerigoleur E, Bitsch S, Dewost M, Guédon A, Guignet J, Le Guével A, Metz A, Vilbert O, Vinette C, Vimal R (2024) Camera-trapping: wild and domestic species occurrences in three Pyrenean pastures. Biodiversity Data Journal 12: e126097. https://doi.org/10.3897/BDJ.12.e126097
- Rodríguez-Pérez, Q. and Zúñiga, F. R.: An earthquake focal mechanism catalog for source and tectonic studies in Mexico from February 1928 to July 2022, Earth Syst. Sci. Data, 15, 4781–4801, http://doi.org/10.5194/essd-15-4781-2023, 2023.
- Testolini, V. (2021). Data from "Ceramic Technology and Cultural Change in Sicily from the 6th to the 11th Century AD." PhD Thesis. Journal of Open Archaeology Data, 9, 11. DOI: http://doi.org/10.5334/joad.77

Où diffuser ? via une publication

Ressources

CNRS, 2021. Guide de bonnes pratiques sur la gestion des données de recherche. Publier un Datapaper pour valoriser et expliciter les données.

Publier un software paper

Institut de l'Information Scientifique et Technique. (2024). **DoRANum-Data papers et data journals : Les software papers.** DoRANum. https://doi.org/10.13143/HGWJ-8S90

Software Sustainability Institute, 2021. In which journals should I publish my software?

DoRANum, Data papers et Data journals
L'essentiel et « Pour aller plus loin »

Où diffuser ? via une publication

Diffusion en libre accès de ses articles :

- Dans tous les cas :
  - déposer son post print dans une archive ouverte (HAL, arXiv)
  cf Loi pour une République Numérique, art 30 : Droit de déposer un article accepté pour publication dans une archive ouverte
  voir Guide d'application et la FAQ

  - si paiement de frais de publication (APC, Article Processing Charges), déposer la version éditeur immédiatement dans une archive ouverte

- Si possible, privilégier les revues :
  - des revues sans frais pour les lecteurs et les auteurs
  - avec frais de publication raisonnable (< 2000 €)
  - en libre accès intégral

Voir liste des revues sans APC du Directory of Open Access Journals (DOAJ) et le Journal Checker Tool

Dernière recommandation !

Lier publications, données, et codes

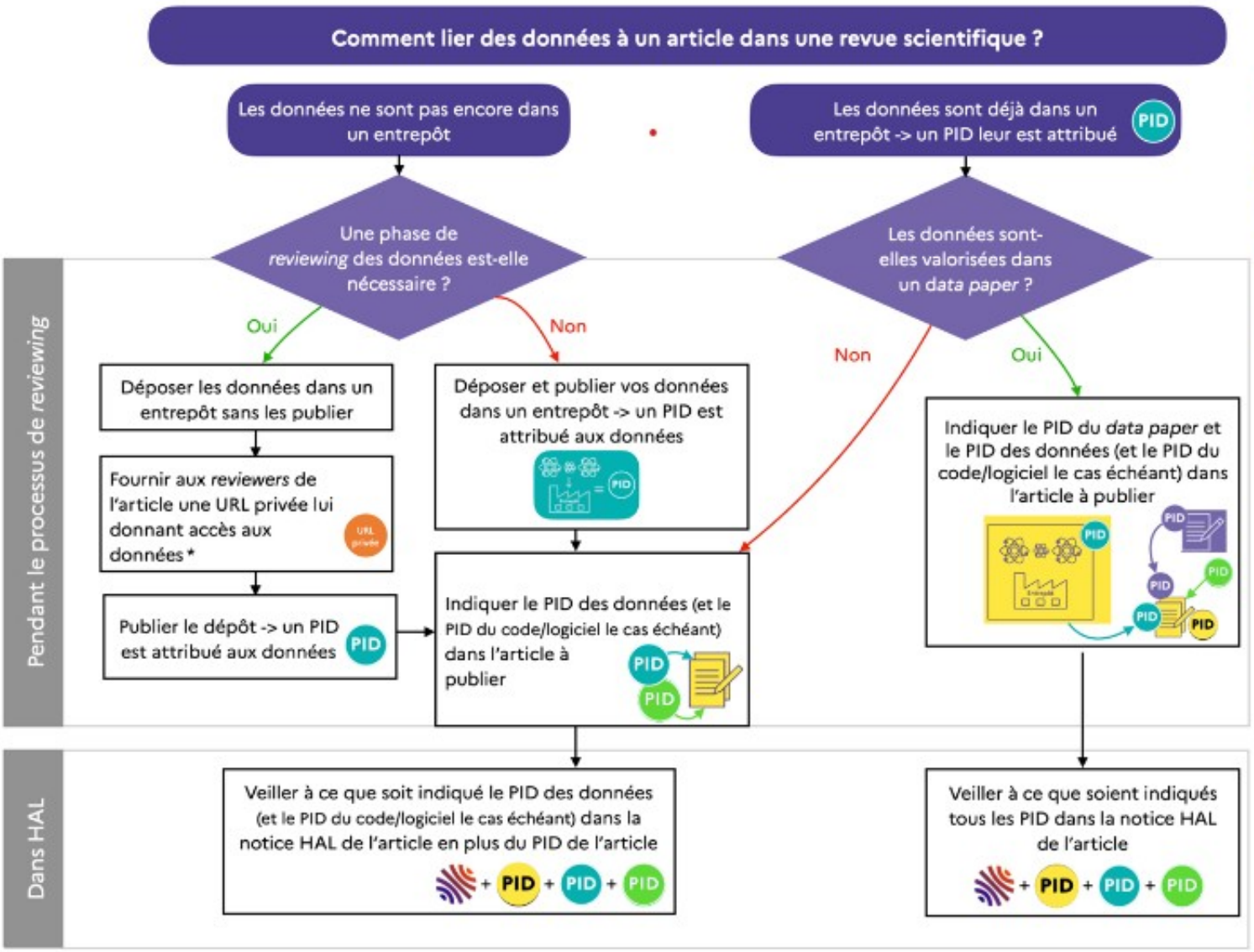
Référencement

- Des publications sur HAL (et dépôt du texte intégral)
- Des données sur des entrepôts thématiques / Recherche Data Gouv /
- Des codes sur Software Heritage et HAL

- Lien entre toutes ces productions via leurs identifiants (doi)

A noter : vous pouvez, depuis un jeu de données, associer également votre thèse déposée dans HAL

Comment lier des données à une publication lors du processus de *reviewing* de l'article (révision par les évaluateurs de la revue) et comment les référencer dans la notice HAL de l'article.



PID des données
PID du data paper
PID de l'article
PID du code ou logiciel ayant servi pour la production des données
Identifiant HAL

PID ? Persistent IDentifier
Les identifiants persistants garantissent un lien stable vers la ressource en ligne (ex : DOI)

* Dans le cas où l'entrepôt ne fournit pas d'URL privée, vous pouvez communiquer le PID des données publiées dans l'entrepôt ou transmettre les données au *reviewer* par un autre moyen

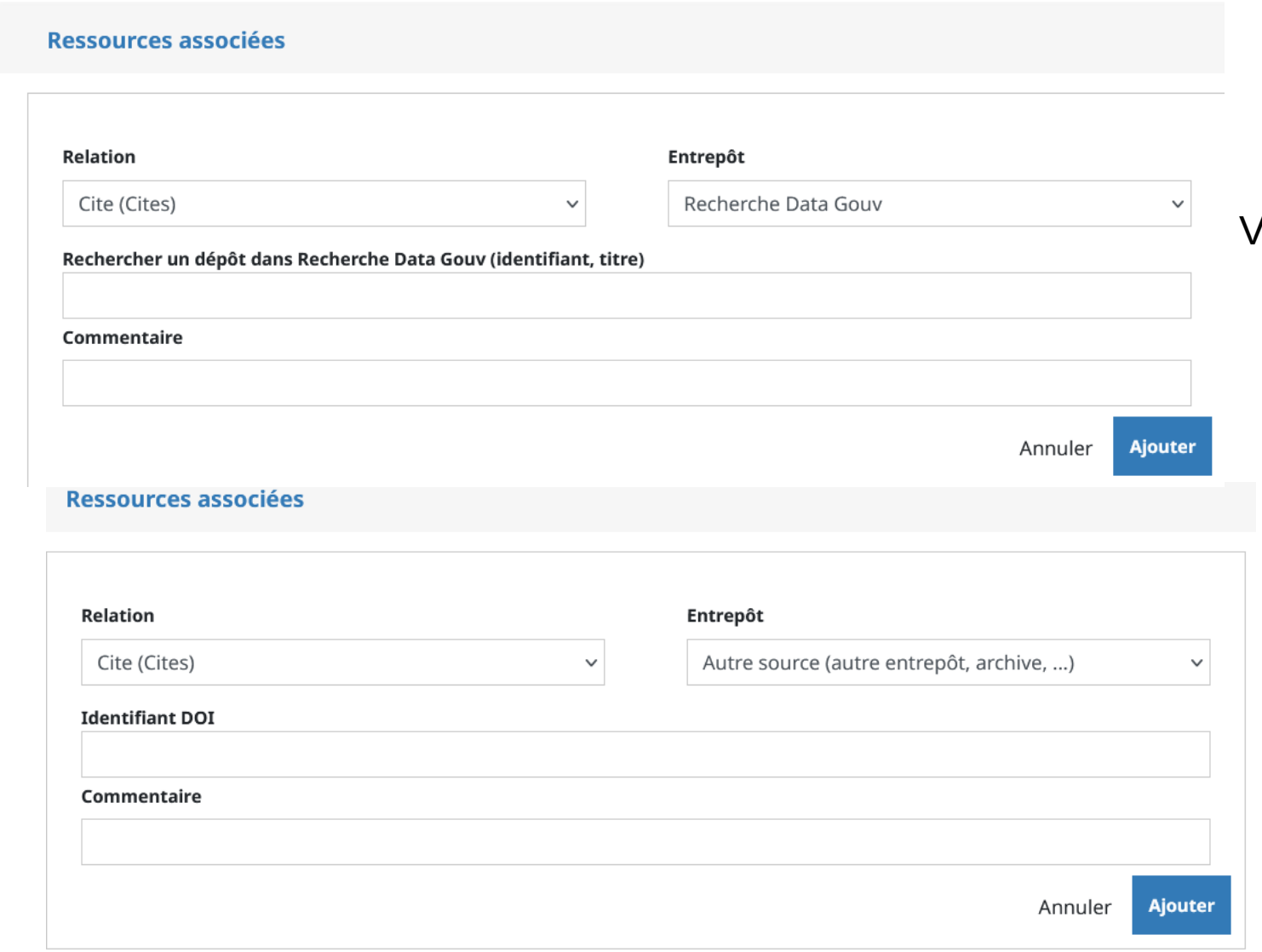
Partager les données liées a ux publications scientifiques. Guide pour les chercheurs. Comité pour la science ouverte. Collège des Données de la recherche. Décembre 2025.

NB : PID = identifiant pérenne, par exemple, un doi, un orcid ...

Dernière recommandation !

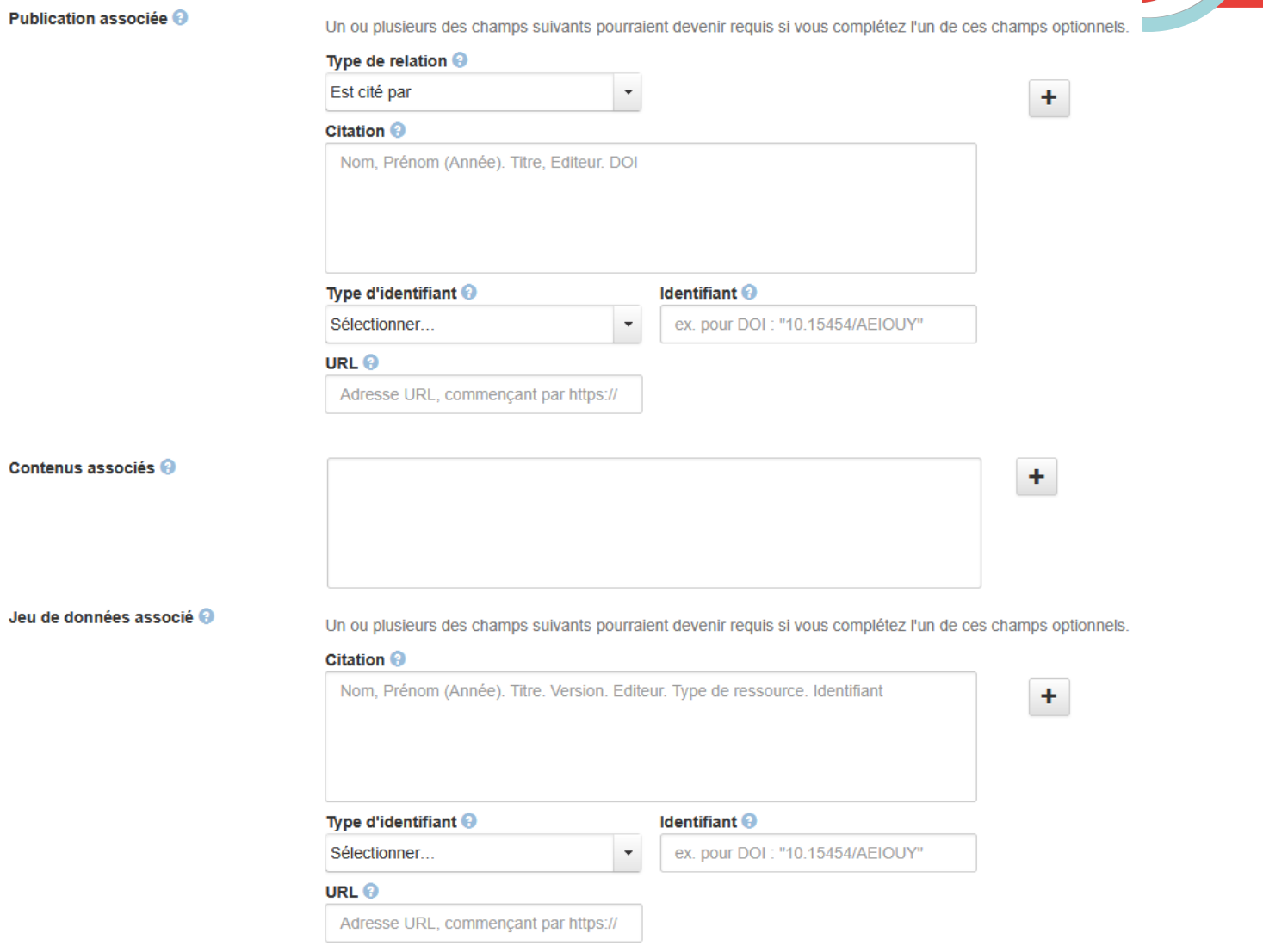
Sur HAL

Voir la documentation sur HAL



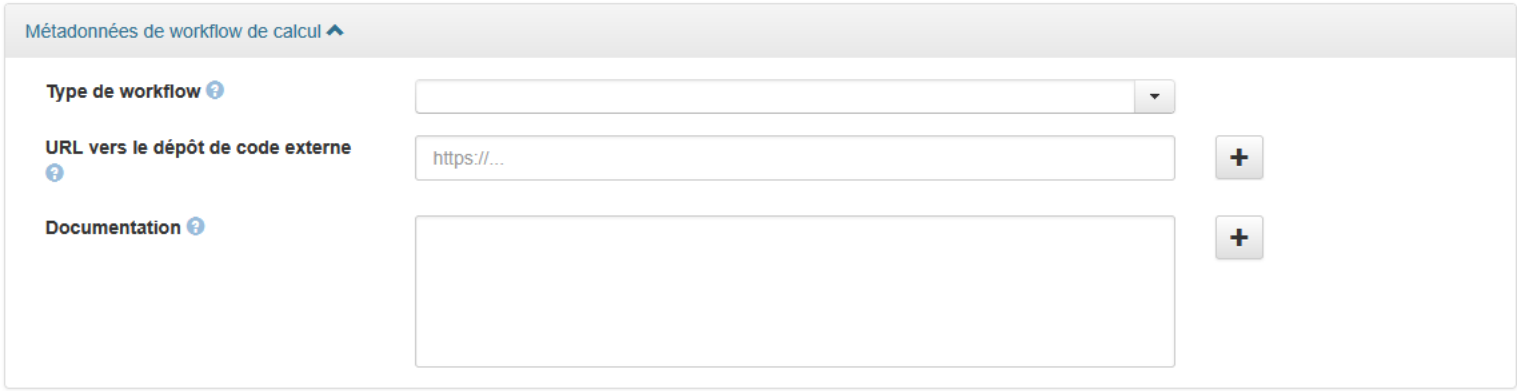
Dernière recommandation !

Sur
Recherche
Data Gouv



Dernière recommandation !

Sur Recherche Data Gouv



Pour aller plus loin :
Partager les données liées aux publications scientifiques. Guide pour les chercheurs. Comité pour la science ouverte. Collège des Données de la recherche. Décembre 2025.

Aide et ressources

Une adresse mail :

sos-data[at]univ-grenoble-alpes.fr

La cellule data Grenoble Alpes répond concrètement à toutes les demandes des communautés scientifiques de Grenoble sur les données.

- Aide à la diffusion des données et des codes
- Aide à la description des données
- Lien publications/données/codes
- Aide juridique
- Diffusion des bonnes pratiques

Et pour les codes : sos-codes-recherche[at]univ-grenoble-alpes.fr

Aide et ressources

Equipe de la la BU : traitement des thèses et accompagnement des doctorants

Rôle :

- Signalement et diffusion de la thèse sur thèses.fr, via HAL thèses

Quelques exemples de services :

- Accompagnement individuel au dépôt de la thèse
- Aide juridique
- Aide pour lier les productions scientifiques

Une adresse : bu-theses@univ-grenoble-alpes.fr

Aide et ressources

- Le site science ouverte de l'UGA

- DoraNum
  - Dépôt et entrepôts
  - Data papers et data journals
- Site du CIRAD, Coopist, Gérer des données
- Passeport science ouverte
  - Codes
  - Données de la recherche