Data.
InDoRES
Entrepôt de données
Inventaire des Données de la Recherche
en Environnement et Sociétés

Version juillet 2024

Guide utilisateurs data.InDoRES

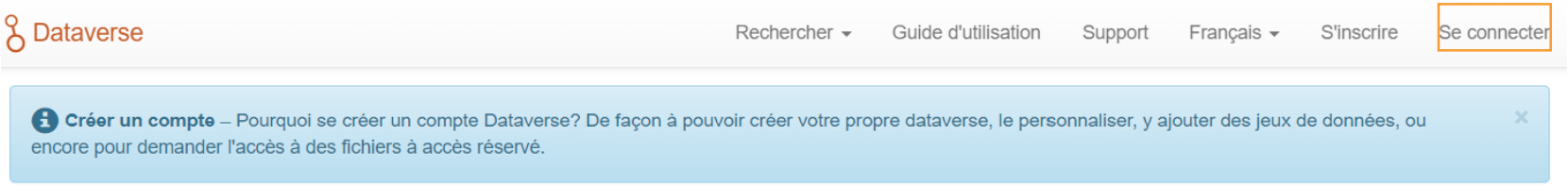
Amandine Hénon
Cécile Callou

Présentation générale

Quelques définitions pour commencer

- **Dataverse** désigne à la fois la technologie utilisée pour l’entrepôt data.InDoRES mais aussi les collections (unités, projets) qui peuvent elles-mêmes comporter des sous-collections.
- **Jeu de données** désigne un ensemble d’informations et de fichiers, unique et cohérent, qui est décrit dans une fiche de métadonnées.
- **Fichier** est l’unité la plus petite. Un jeu de données peut contenir plusieurs fichiers.

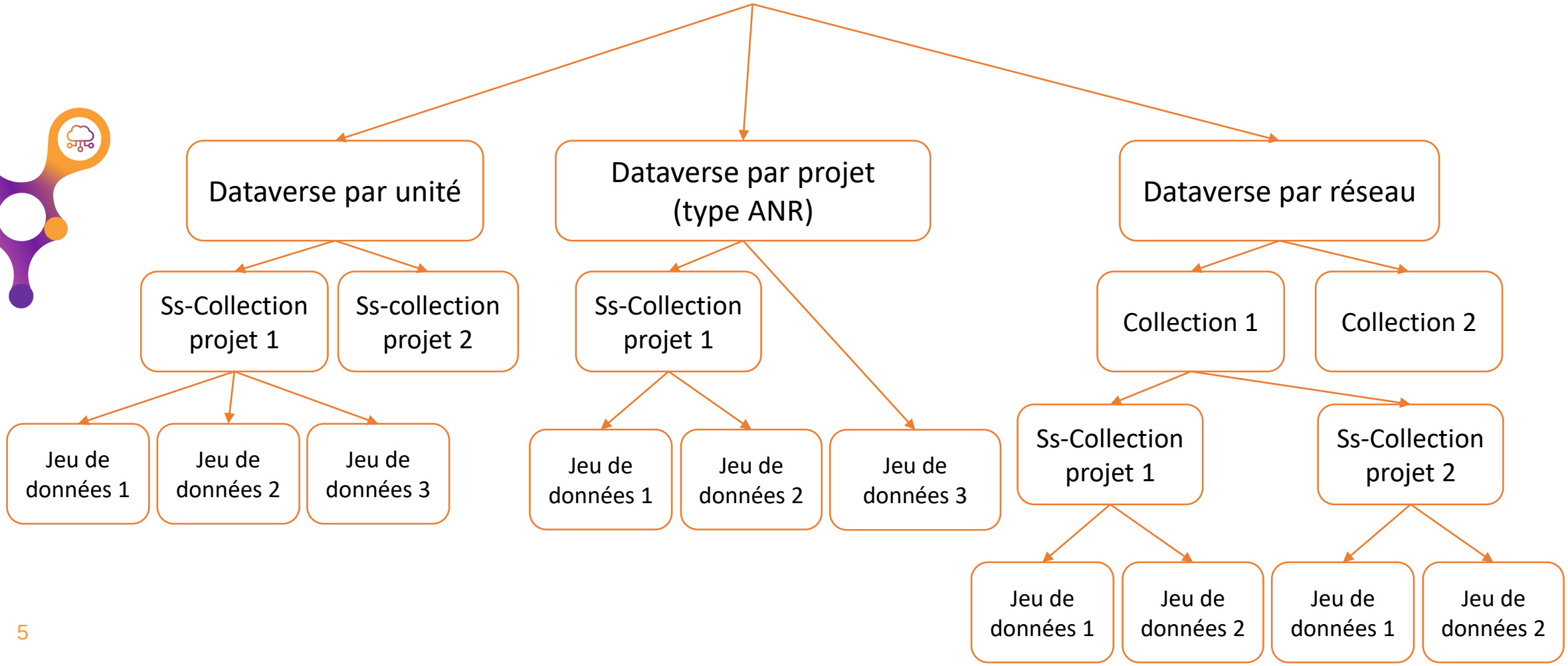
Créer votre compte…



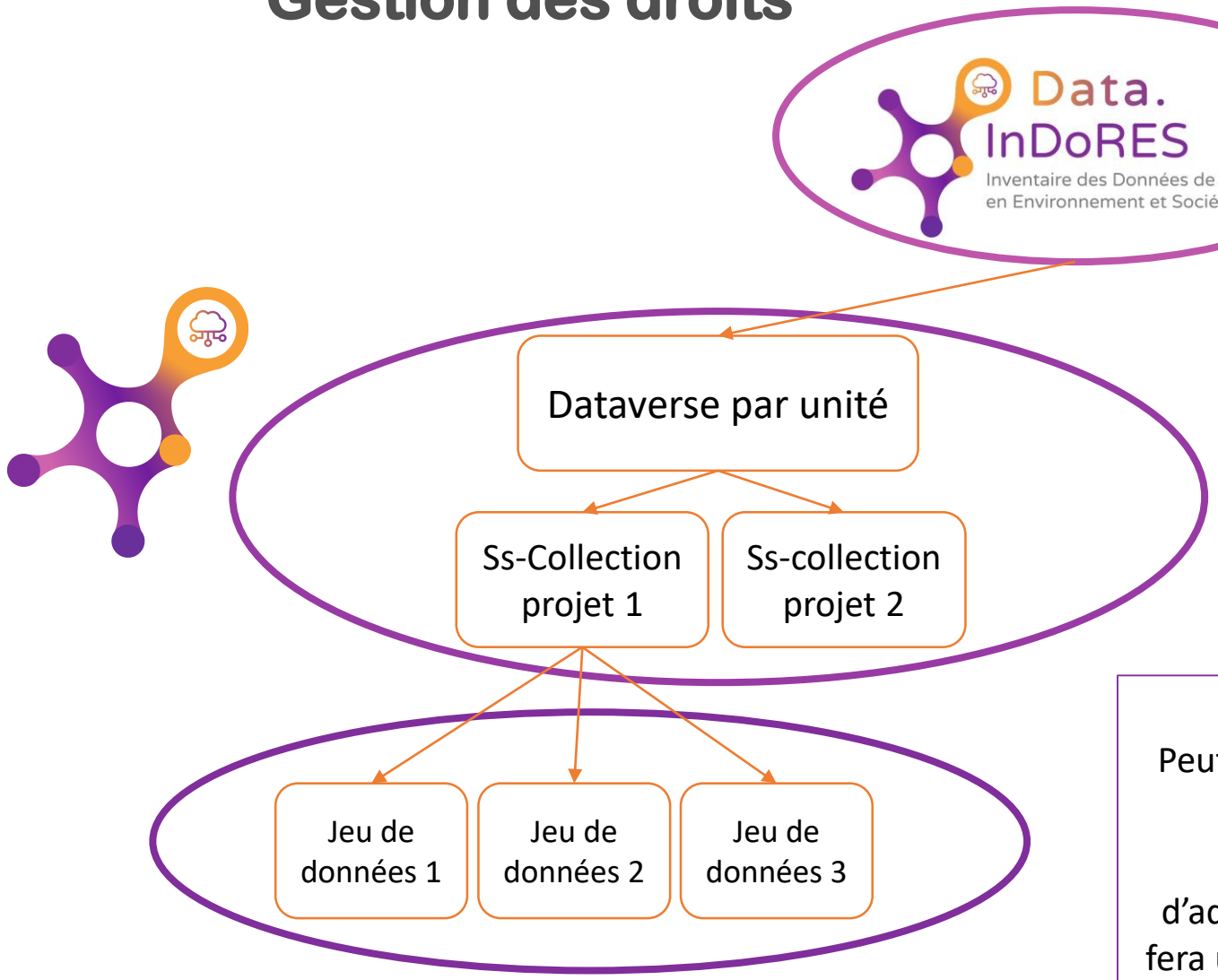
- A la création de votre compte (de préférence via Edugain, voir slide 16), vous ne pourrez pas ajouter et créer des jeux de données.
- Contactez un administrateur pour obtenir les droits de publication en cliquant sur le bouton « Contact » de la collection dans laquelle vous souhaitez publier.
  La liste des référents est également disponible **ici**.
- Une fois des droits de publication obtenus, vous ne pourrez ajouter des jeux de données **que** dans le dataverse de votre unité ou de votre projet.

Arborescence de Data.InDoRES

Inventaire des Données de la Recherche en Environnement et Sociétés



Gestion des droits



Géré par l'équipe data.InDoRES
Peut agir et administrer la totalité des Dataverse et jeux de données de l'entrepôt.
Pas de rôle de curation des données

Gérés par le référent OpenDoRES (ou à défaut par l'équipe data.InDoRES)
Peut agir et administrer la totalité des sous-collections et jeux de données du Dataverse.
Pas de rôle de curation des données

Gérés par l'auteur des jeux de données
Peut créer et modifier des jeux de données dans le Dataverse de son unité ou d'un projet uniquement.

Les publications sont soumises au référent (rôle d'administration), qui vérifiera les métadonnées, le format et fera une validation technique (*pas de curation des données)* ou demandera des modifications/compléments.

Publication d’un jeu de données

Fiche cat.InDoRES

Avant toute publication d'un jeu de données dans l'entrepôt data.InDoRES, il est nécessaire de compléter sa fiche de métadonnées dans cat.InDoRES

https://cat.indores.fr/

https://www.indores.fr/nouveau-jeu-de-donnees

Fiche cat.InDoRES

Les informations contenues dans les fiches du catalogue sont beaucoup plus détaillées que dans la description attendue du jeu de données dans l'entrepôt data.InDoRES. Elles se **rapportent à l'ensemble du projet et permettent de décrire finement le contexte d'acquisition et de production du jeu de données,**

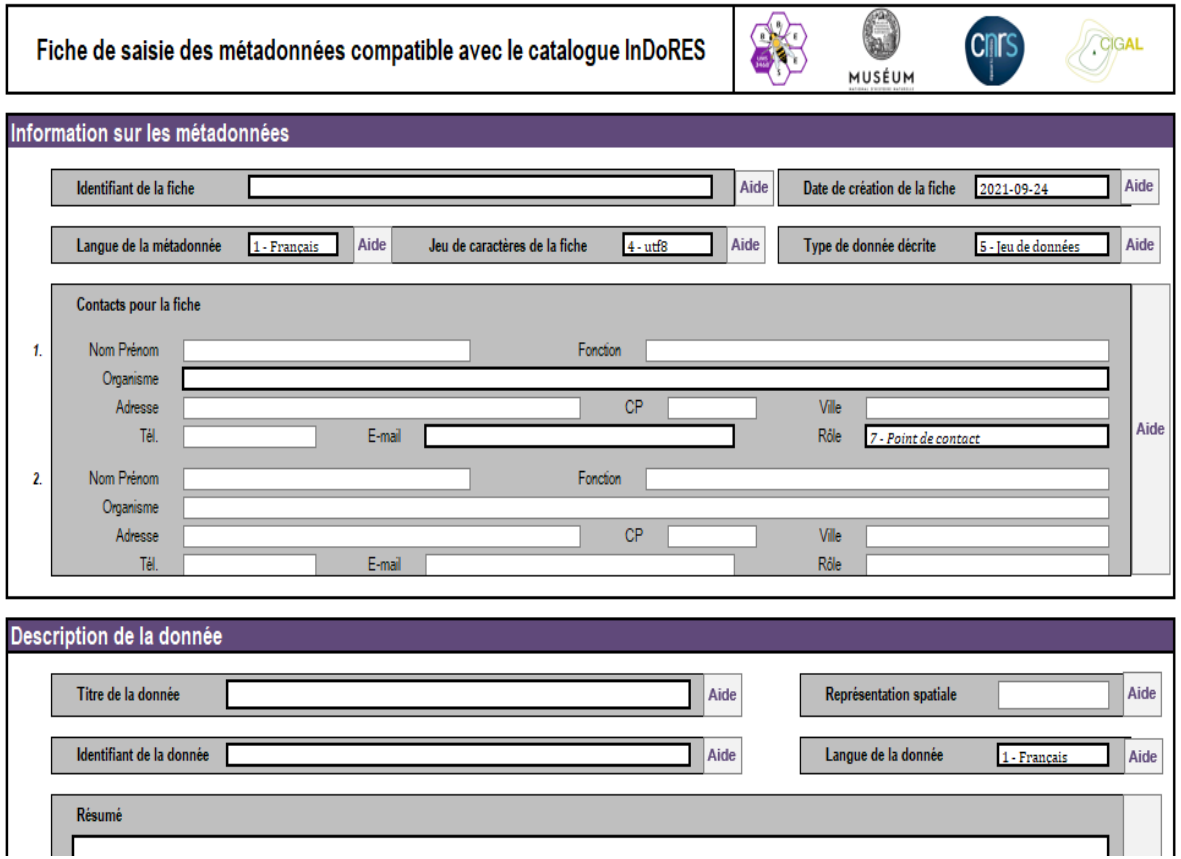
Un lien vers des publications ou tout autre ressources (site internet, protocole etc.) peut être associé dans la fiche cat.InDoRES

A noter : Nous travaillons actuellement à la mise à disposition d'un outil qui facilitera la saisie en une fois dans les deux outils (cat.InDoRES et data.InDoRES).

Fiche cat.InDoRES

- Feuille Excel à compléter.
  Elle reprend les standards (ISO 19115 et 19139), répondant ainsi aux exigences légales pour les données environnementales géolocalisées (Directive INSPIRE)

- Une aide est intégrée pour chaque champ (définition et exemple associé)



En cas de difficultés ou question : contact@indores.fr

La fiche remplie sera envoyée via le formulaire de contact sur le portail InDoRES ou directement à l'adresse: contact@indores.fr

Des bons usages de data.InDoRES

Avant toute publication de jeux de données, quelques recommandations indispensables :

- Préparer un fichier readme ou noter sur le premier onglet, si le fichier est de type tableur, le nom des variables ainsi que tout élément permettant une meilleure réutilisation du jeu de données.
- Privilégier les formats préconisés par InDoRES, qui faciliteront la pérennité des données. La liste des formats recommandés est disponible ici : https://www.indores.fr/sites/default/files/documents/04FormatsListe.pdf

Exemple d’un fichier ‘readme’

```
Ce jeux de données contient :

+ ce fichier a_lire.txt (et une traduction anglaise readme.txt)

+ le fichier ZAEU.zip contient le shapefile utilisé pour l'extraction des rasters

+ les rasters (.asc, WGS84) pour les variables BIOG (x1), GPW4 (x1), LC (x1) , LSTD (x240), LSTN (x240), NDVI (x237).
  le nom de ces rasters contient le mois et l'année.
  les informations sur la provenance et les traitements appliquées sont contenus dans les fichiers <xxxx>.json

+ les fichiers BIOG.json, GPW4.json, LC.json, LSTD.json, LSTN.json, NDVI.json contiennt la provenance des rasters originaux
  une description des traitements apliqués, des unités, des palettes de couleurs...

 + les fichiers ZAEU.json (verifié) et ZAEU.xml (non vérifié) contienent les indicateurs et la geometrie

 + les fichiers ZAEU-LSTD.csv, ZAEU-LSTN.csv, ZAEU-NDVI.csv contiennent les moyennes spatiales (aussi contenues dans ZAEU.json)

La publication de référence est
 + Isabelle Charpentier. From rasters to gradients of a social ecological research network. 2023. (hal-03942291v1)

La fiche parent est :
+ Collection of social ecological indicators for the assessment of gradients of the Metropolitan Zones Ateliers. Charpentier, Isabelle, 2023,
"Collection of social ecological indicators for the assessment of gradients of the Metropolitan Zones Ateliers", https://doi.org/10.48579/PRO/CJ3LHZ,
 data.InDoRES, V3

Des jeux de données similaires existent pour les autres zones ateliers
```

Exemple d’un fichier readme rédigé pour une fiche sur data.InDoRES par I. Charpentier

A noter : Il est possible de remplacer ce fichier, sur les tableurs, par une première feuille expliquant les noms de variables utilisées et toute information utile pour comprendre les données.

Des bons usages de data.InDoRES

Vérifications avant de publier vos données :

- Assurez-vous que votre ensemble de données respecte les critères d'anonymisation des données sensibles et personnelles dans le cadre du Règlement Général à la Protection des Données (RGPD).
- Si vous réutilisez des données dont vous n'êtes pas l'auteur, vérifiez la licence et vos droits à les rediffuser.

Parcours de publication : première connexion

Création et validation du compte
(diapositive suivante)

Envoi d'un mail à votre référent pour création de droits de publication *

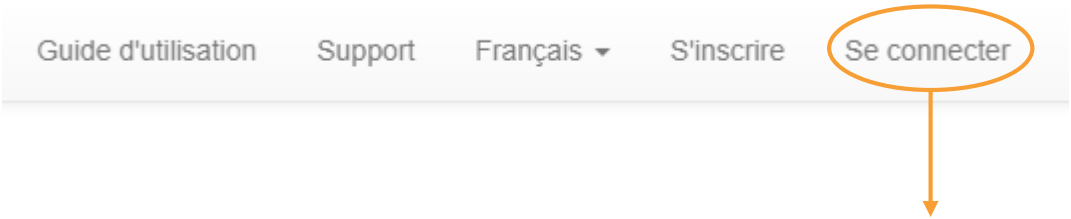
Mail de votre référent quand vos droits sont créés

Le bouton « Ajouter des données » est désormais visible.

☺

* **Lien vers la liste des référents**

Création d’un compte



Choisir de préférence l’option Edugain qui permet une plus grande sécurisation

A noter : pas de rôle par défaut donc même si le compte est créé, vous ne pouvez rien faire d’emblée.
Prévenir le référent de votre unité (ou les administrateurs de data.InDoRES) lors de la création de votre compte pour l’attribution d’un rôle. **Lien vers la liste de référents**.

Création d’un compte

A noter : Pensez à écrire au référent de votre unité pour l’ouverture des droits de publication en utilisant le bouton « Contact » sur la droite.

Une fois le compte créé, ce bouton apparait

Ajouter des données - saisie de la fiche de métadonnées

Description des champs obligatoires

Portail data.InDoRES (Muséum national d'Histoire naturelle et Institut Ecologie et Environnement (INEE) du CNRS)

Portail data.InDoRES >

Dataverse hôte

Changer le dataverse hôte effacera tous les champs dans lesquels vous pourriez avoir entré des données.

Portail data.InDoRES

Doit rester inchangé.
Il indique le Dataverse situé à la racine de l'entrepôt (ici data.InDoRES)

Modèle

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

Licence CC-BY

*Les astérisques indiquent les champs obligatoires.

Métadonnées de la référence bibliographique

Titre *

Enter title...

Ajouter "Données de réplication pour" au Titre

Auteur *

Nom *: Hénon, Amandine

Affiliation: Muséum national d'Histoire naturelle

Schéma de l'identifiant: Sélectionner...

Identifiant:

Personne-ressource *

Nom: Hénon, Amandine

Affiliation: Muséum national d'Histoire naturelle

Courriel *: amandine.henon@mnhn.fr

Ajouter des données - saisie de la fiche de métadonnées

Description des champs obligatoires

Portail data.InDoRES (Muséum national d'Histoire naturelle et Institut Ecologie et Environnement (INEE) du CNRS)

Portail data.InDoRES >

Dataverse hôte

Changer le dataverse hôte effacera tous les champs dans lesquels vous pourriez avoir entré des données.

Portail data.InDoRES

Modèle

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

Licence CC-BY

*Les astérisques indiquent les champs obligatoires.

Métadonnées de la référence bibliographique

Titre *

Enter title...

Ajouter "Données de réplication pour" au Titre

S'il existe un modèle particulier créé pour votre unité, vous pouvez le choisir ici.

Si non, le CC-BY reste celui par défaut.

A noter : Les 'templates' sont des modèles génériques. Celui proposé par data.InDoRES est intitulé 'Licence CC-BY 4.0', mais il se peut que les administrateurs de votre unité ou projet aient créé des templates spécifiques, notamment avec des champs pré-remplis.
Si c'est le cas, il est possible que le template par défaut soit différent.

Hénon, Amandine

Muséum national d'Histoire naturelle

+

Courriel *

amandine.henon@mnhn.fr

Ajouter des données - saisie de la fiche de métadonnées

Description des champs obligatoires

Portail data.InDoRES (Muséum national d'Histoire naturelle et Institut Ecologie et Environnement (INEE) du CNRS)

Portail data.InDoRES >

Dataverse hôte

Changer le dataverse hôte effacera tous les champs dans lesquels vous pourriez avoir entré des données.

Portail data.InDoRES

Modèle

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

Licence CC-BY

*Les astérisques indiquent les champs obligatoires.

Métadonnées de la référence bibliographique

Titre *

Enter title...

Ajouter "Données de réplication pour" au Titre

Auteur *

Nom *: Hénon, Amandine

Affiliation: Muséum national d'Histoire naturelle

Schéma de l'identifiant: Sélectionner...

Identifiant:

Personne-ressource *

Nom: Hénon, Amandine

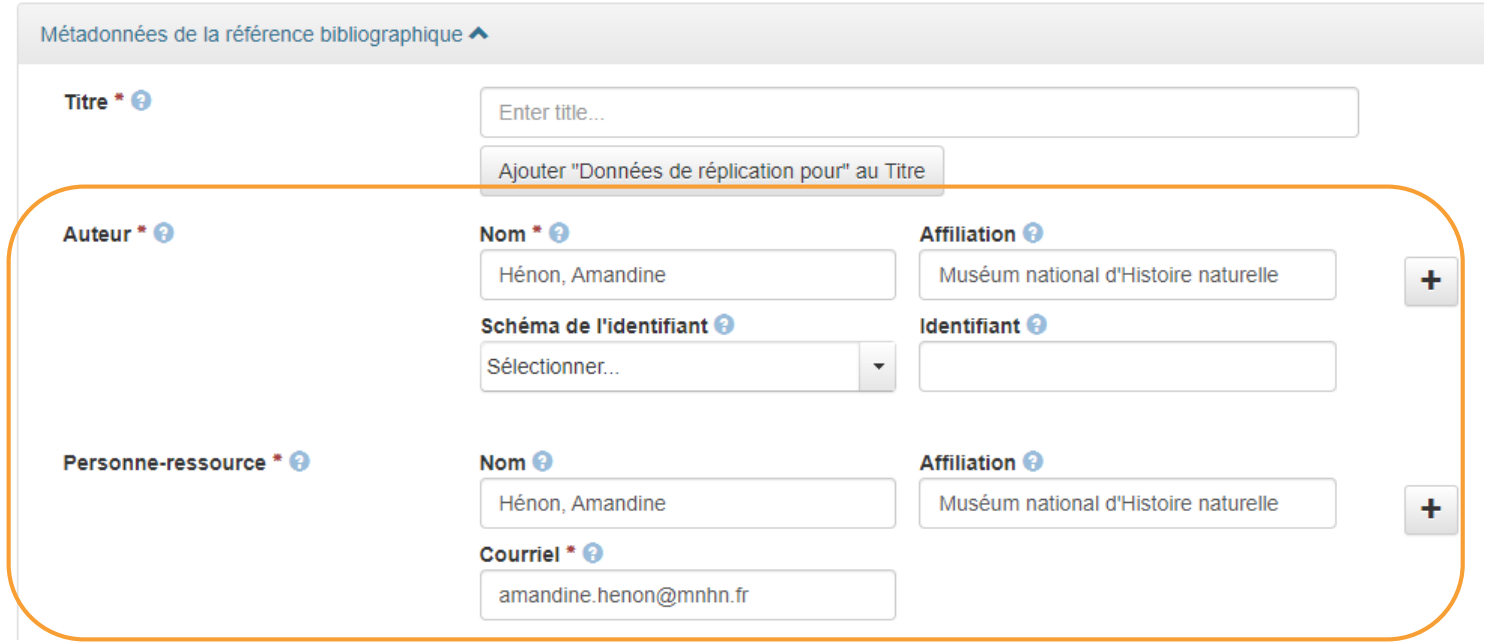
Affiliation: Muséum national d'Histoire naturelle

Courriel *: amandine.henon@mnhn.fr

Champs obligatoires. Autant que possible, faire figurer un identifiant unique type ORCID pour l'auteur.

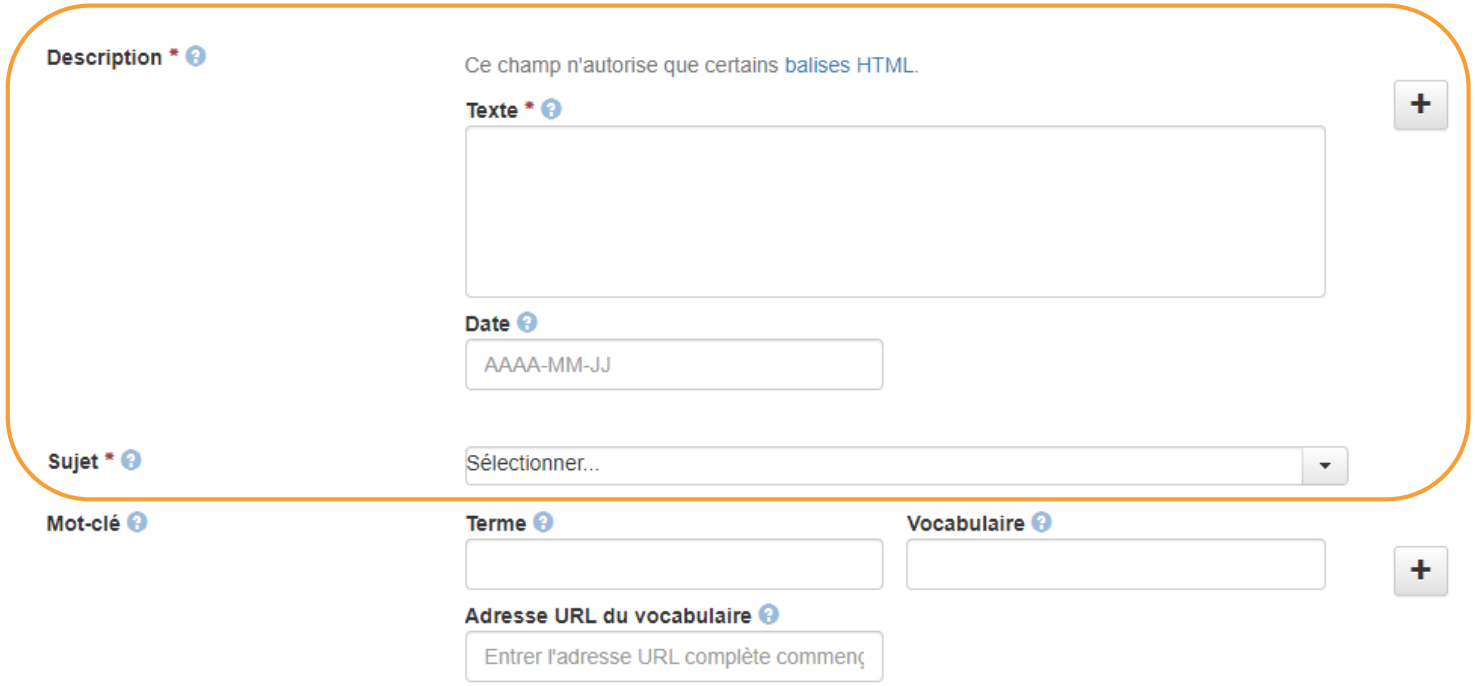
Ajouter des données - saisie de la fiche de métadonnées

Description des champs obligatoires



Les champs peuvent être dédoublés en appuyant sur le bouton + à droite. Pour « Auteur » et « Personne-ressource », indiquez autant de personnes que nécessaire.

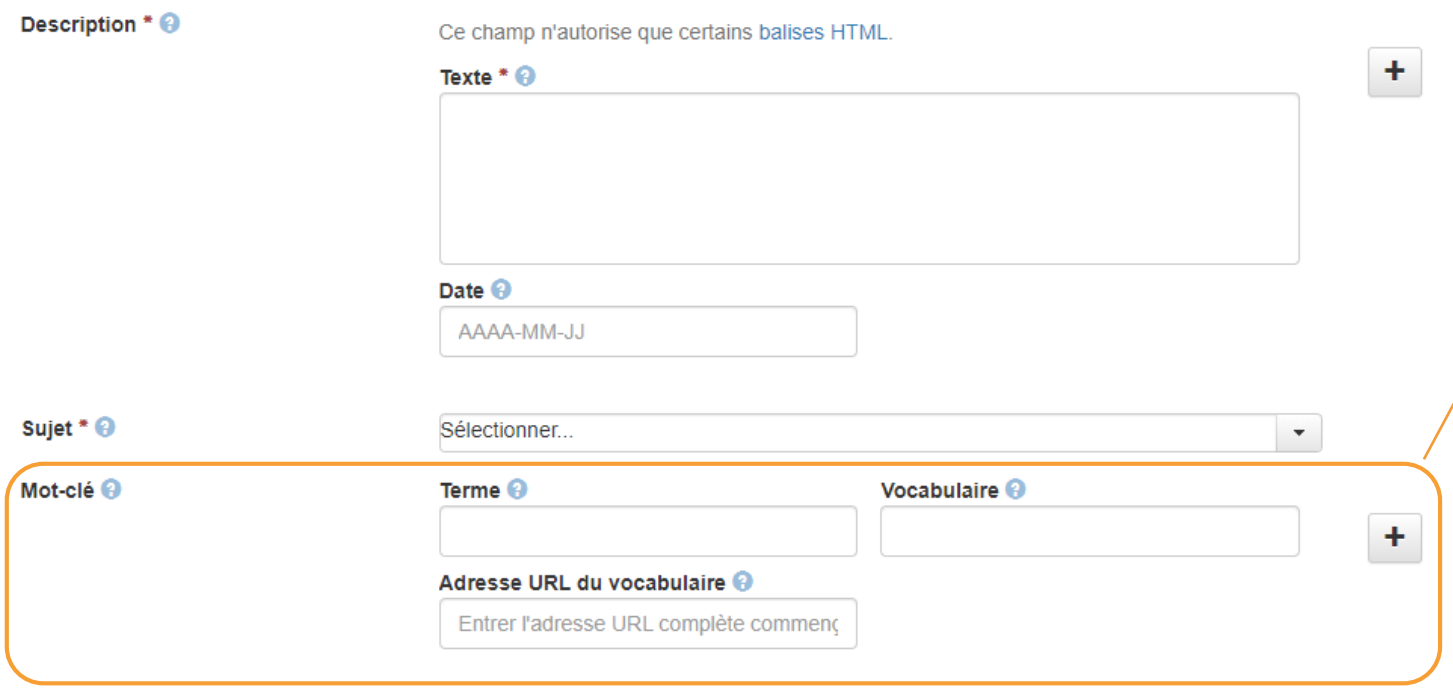
Saisie de la fiche par l'utilisateur - Description des champs de métadonnées obligatoires



La description et le sujet sont également des éléments obligatoires. Ne pas hésiter à mettre des détails dans la description

Ne pas hésiter à user du copier/coller de la fiche rédigée pour cat.InDoRES !

Saisie de la fiche par l'utilisateur - Description des champs de métadonnées obligatoires



A noter : Un seul mot-clé par ligne. Dédoubler la ligne à l'aide du « + » autant de fois que nécessaire.
Ne pas hésiter à user du copier/coller de la fiche rédigée pour cat.InDoRES !

Pas de thésaurus inclus par défaut.

Obligation d'inclure **au moins 2 mots clés** dans le Champ "Terme"

S'ils sont issus de thésaurus, indiquer le nom de celui-ci dans le champ "Vocabulaire".

Le champ "Adresse URL du vocabulaire" vous permet de noter l'adresse internet associée au mot clé dans le thésaurus utilisé - Ex : "faune" (http://data.loterre.fr/ark:/67375/BLH-D1K2VRK8-G)

Ajouter des données - saisie de la fiche de métadonnées

Autres champs

Publication connexe

Référence

+

Type d'identifiant

Sélectionner...

Numéro d'identification

Adresse URL

Entrer l'adresse URL complète commenç

Remarques

Déposant

Callou, Cécile

Date de dépôt

2023-05-03

Si le jeu de données est à mettre en lien avec une ou des publication(s) déjà existante(s), la ou les références peu(ven)t être ajoutée(s) (avec précision DOI, ISSN, ISBN etc.)

A noter : Dédoubler les publications à l'aide du « + » autant de fois que nécessaire.

Ajouter des données - saisie de la fiche de métadonnées

Dépôt du fichier pour publication

Déposer des fichiers ou dossier zip dans la **limite de 5Go**.

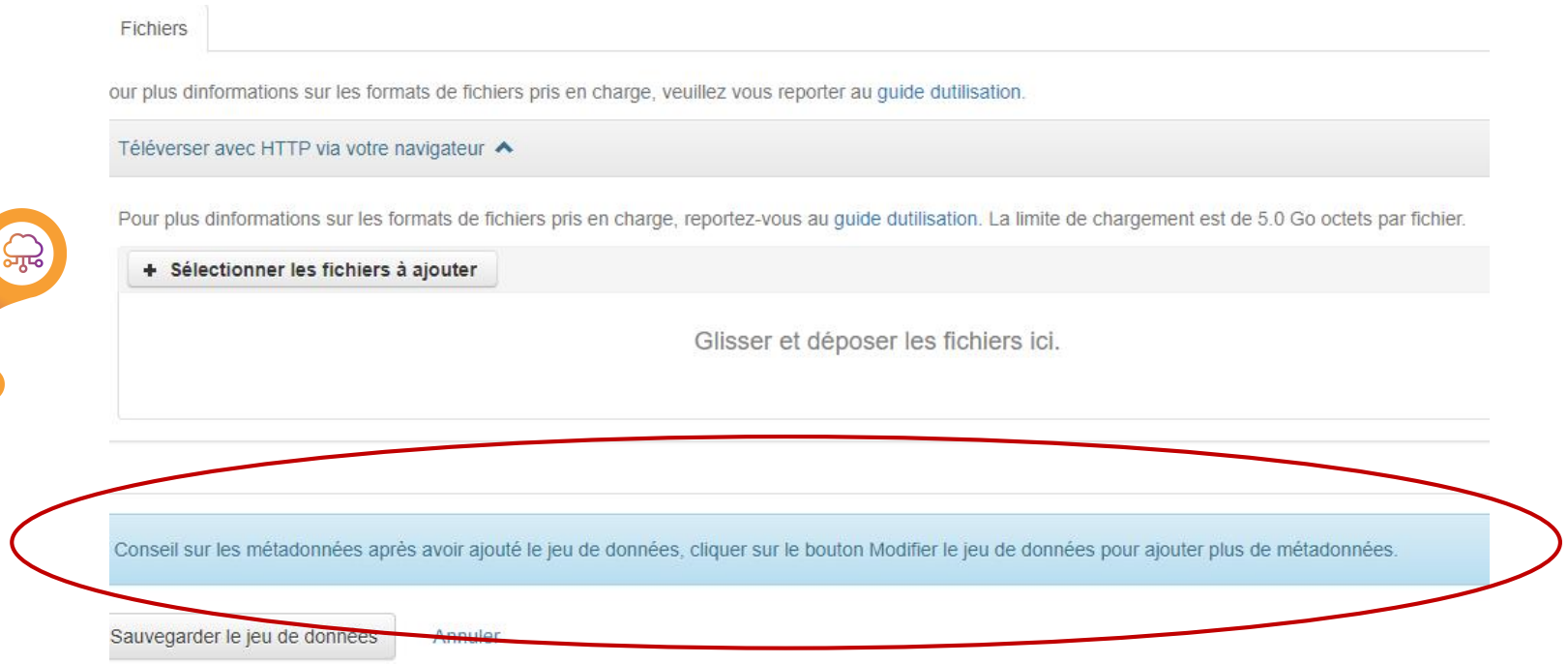
Les fichiers plus volumineux seront déposables ultérieurement de manière plus simple et rapide.

Puis penser à **sauvegarder**

A noter : Vous pouvez déposer autant de fichiers que nécessaire dans une même fiche. Attention toutefois si la taille totale des fichiers excède 80Go. Contacter l'équipe InDoRES au préalable.
En particulier, **intégrer un fichier readme** pour une meilleure réutilisation des données.
Ne pas oublier les formats recommandés.

Ajouter des données - saisie de la fiche de métadonnées

Description de métadonnées supplémentaires

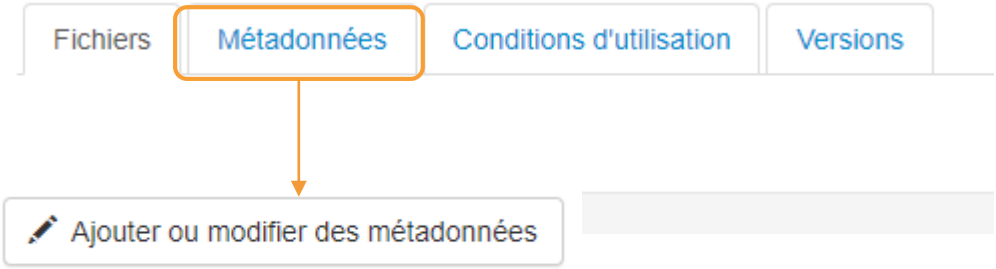


Une info-bulle rappelle la possibilité d'ajouter des métadonnées supplémentaires, qui sont plus détaillées et variées que les métadonnées obligatoires.
Cette action se fait une fois les fichiers déposés et sauvegardés.
En pratique : voir les pages suivantes

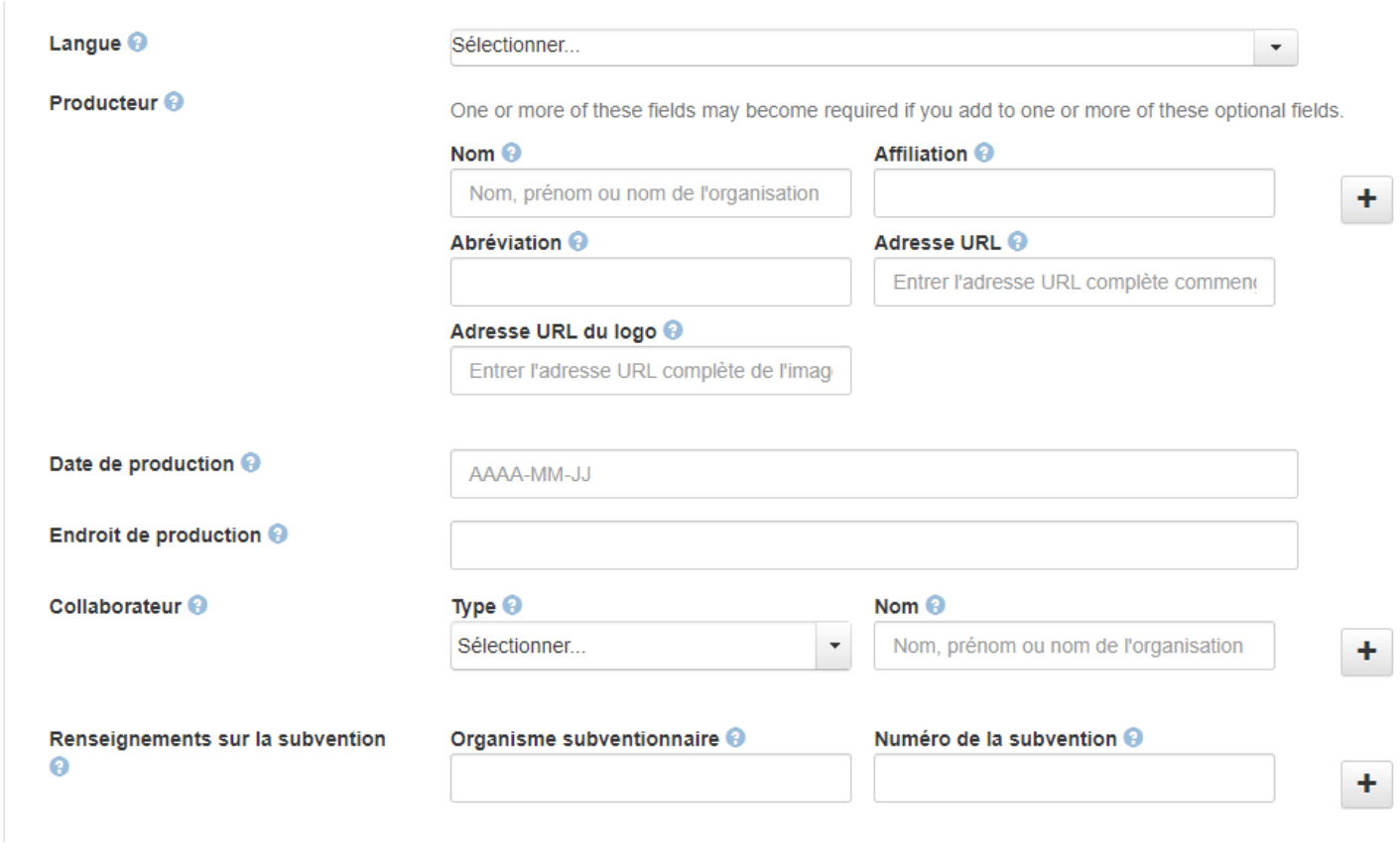
Description de métadonnées supplémentaires

Par défaut, le nombre de champs de métadonnées demandé au moment du dépôt est assez peu important.

Vous pouvez apporter plus d'informations (enrichir les métadonnées) à partir de la **page d'accueil du jeu de données**



Description de métadonnées supplémentaires (exemple de champs à renseigner)



A noter : Ne pas oublier d’enregistrer les modifications pour leur prise en compte.

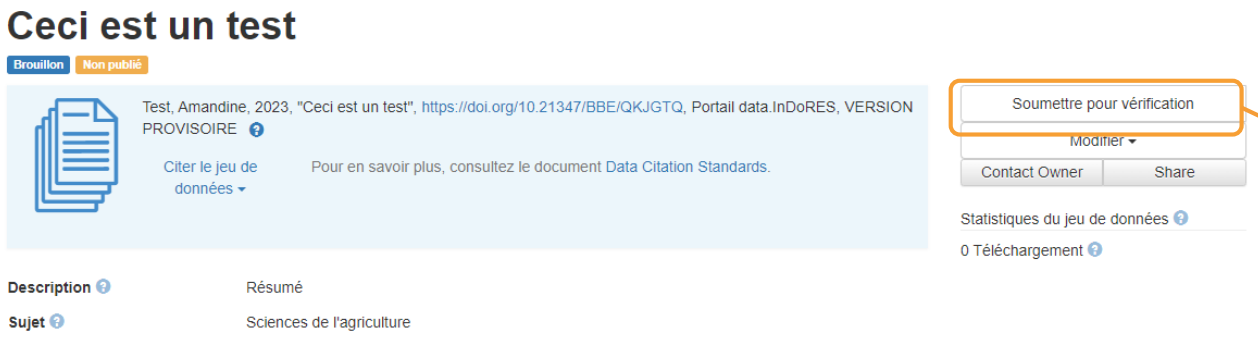
Page d’accueil du jeu de données - Conditions d’utilisation

Permet d’affiner ou de spécifier certaines conditions d’utilisation.

Les restrictions doivent être motivées par un argument justifié (données sous embargo dans le cadre d’une publication par exemple) car par défaut, elles sont directement accessibles.

Attention : La licence CC-BY 4.0 s’inscrit dans la démarche de Science Ouverte et répond aux obligations légales qui s’y réfèrent. Nous souhaitons donc qu’elle ne soit pas changée.

Soumission pour publication au référent

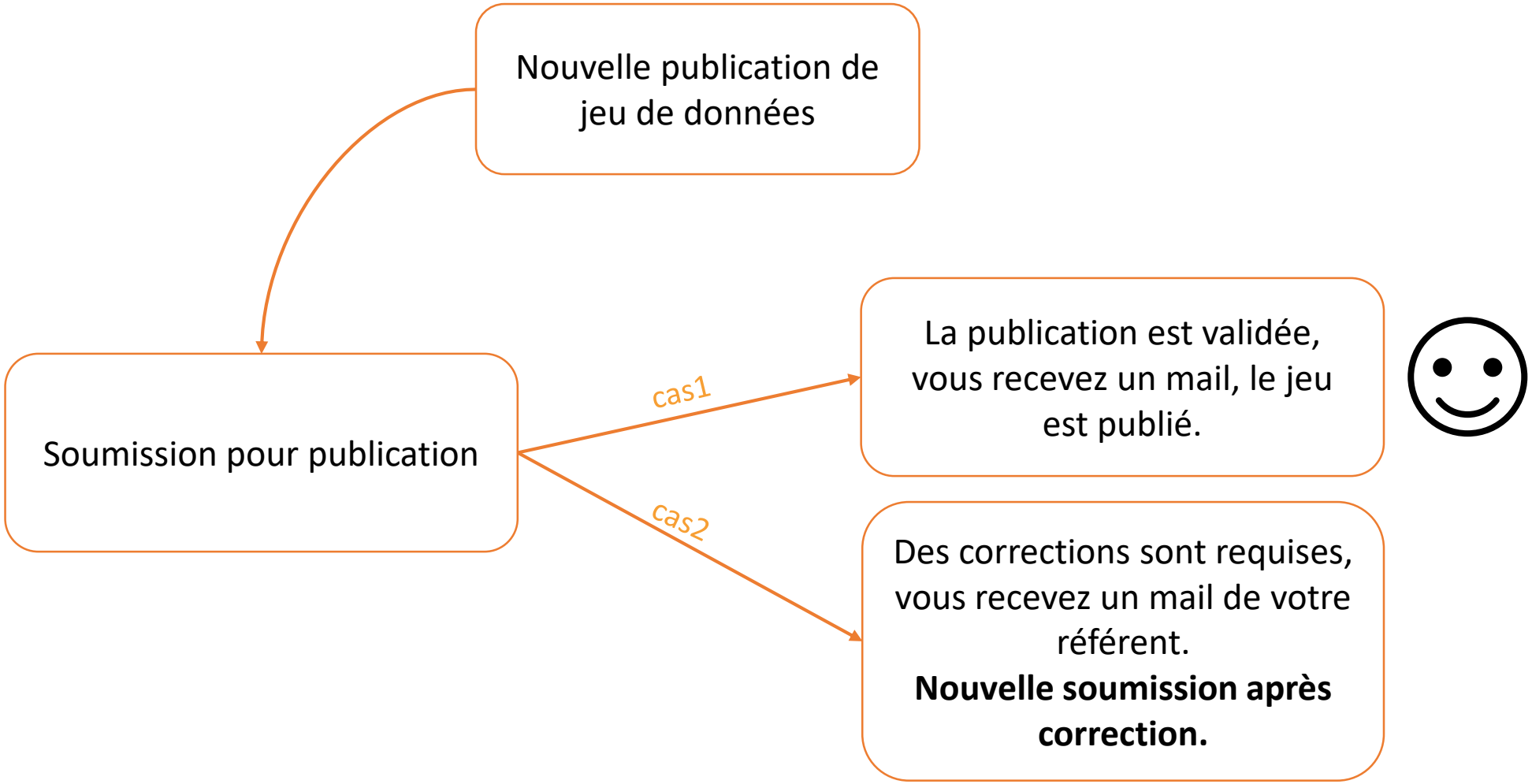


Lorsque vous avez terminé de remplir les métadonnées et de déposer les jeux de données, vous pouvez soumettre pour vérification aux administrateurs/référents.

A noter : Vous ne pouvez plus procéder à des modifications une fois la fiche soumise pour vérification.

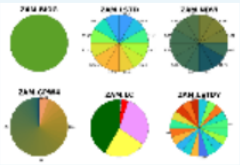
La vérification porte sur la complétude des métadonnées obligatoires et le format retenu (il ne s'agit donc que d'une validation technique) ; le contenu des fichiers relève de votre responsabilité.

Parcours de publication après ouverture des droits



Diffusion et réutilisation des jeux de données

Une fois le jeu de données publié



Charpentier, Isabelle, 2023, "Collection of rasters and social ecological indicators for Zone Atelier bassin de la Moselle", https://doi.org/10.48579/PRO/WRKT4O, data.InDoRES, V1, UNF:6:eLhH8/Efa3YAi8CPnL6GAw== [fileUNF]

Citer le jeu de données

Pour en savoir plus, consultez le document Data Citation Standards.

Un DOI est automatiquement ajouté au jeu de données, la référence peut être copiée/collée
Possibilité d'exporter la citation sous différents formats (XML, RIS, BibTeX)

Statistiques du jeu de données

1 Téléchargement

Nombre de téléchargement des jeux de données

A noter : D'autres statistiques seront disponibles dans une version ultérieure :

Modalités d'accès au dataset

Contact | Partager

Statistiques d'utilisation sur les datasets

170 consultations

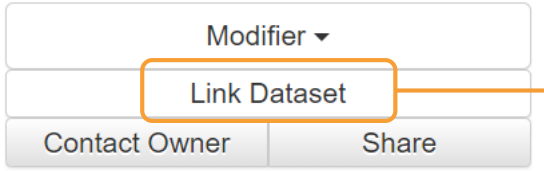
4 téléchargements

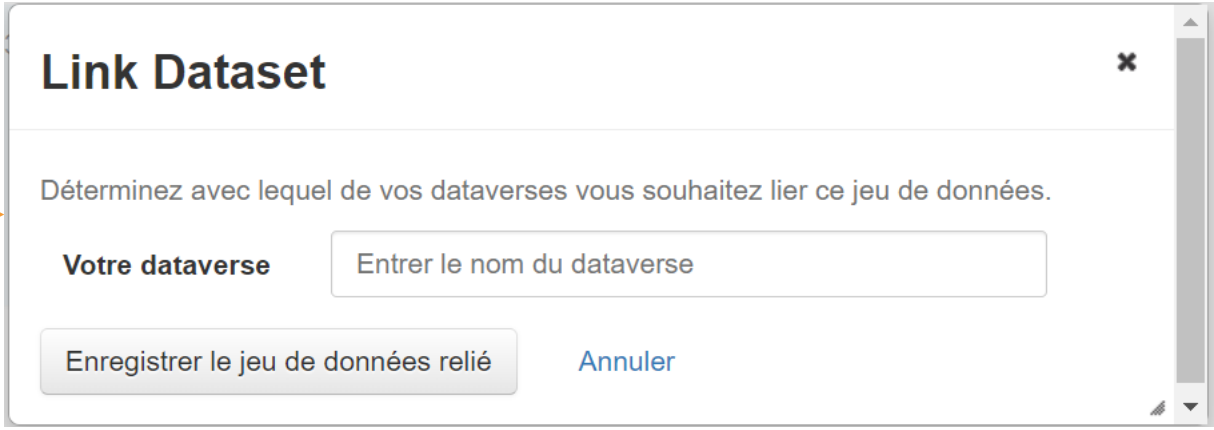
0 citation

Cas particuliers

Lier un jeu de données publié à deux sous-dataverse

Un jeu de données est publié dans un sous-dataverse et peut ensuite être visible dans un autre.
Exemple : un jeu produit dans le cadre d'un projet ANR (Dataverse dédié) et qui soit visible dans le Dataverse des UMR des participants.

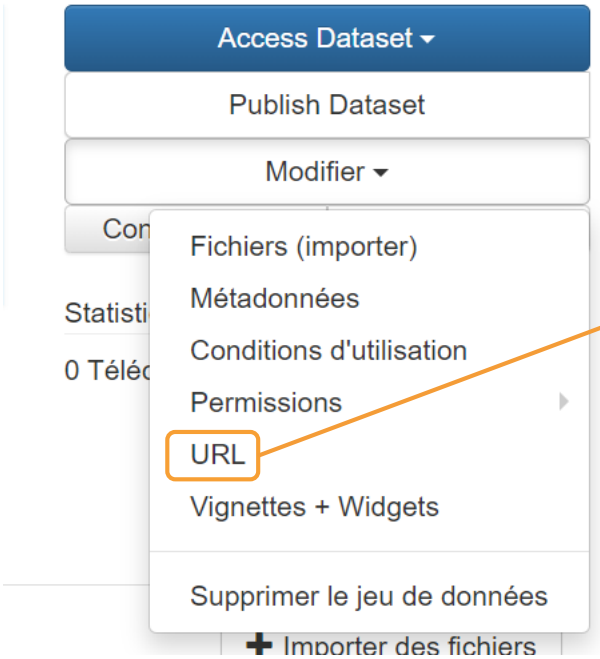




A noter : Cette opération s'effectue sur un **jeu de données publié**.

Cas des révisions d'article - Création d'URL privée

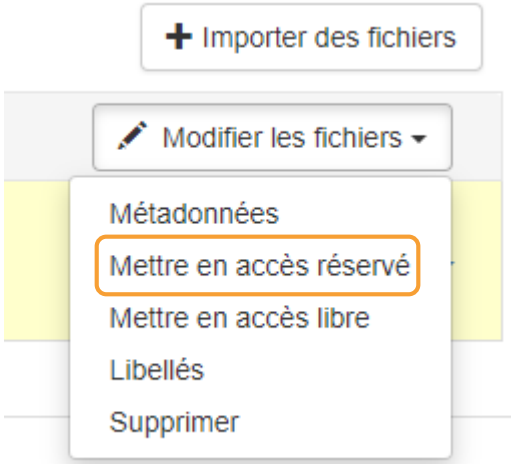
Pour un jeu de données associé à un article, dans une revue avec relecteurs, la procédure est la suivante :



A noter : Cette opération s'effectue sur **un jeu de données non publié**.
L'adresse créée sera à envoyer aux relecteurs (le DOI reste réservé le temps de la relecture).
Cette option est faite par un administrateur, n'hésitez pas à leur demander d'activer le lien.

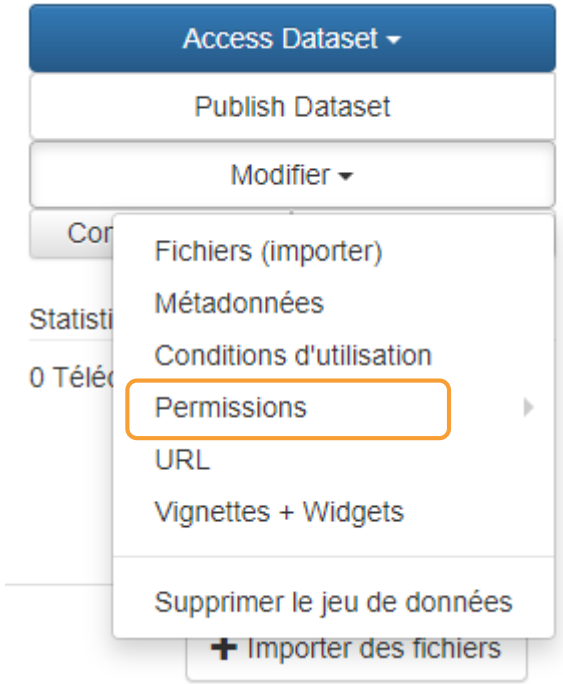
L’accès restreint – une utilisation exceptionnelle

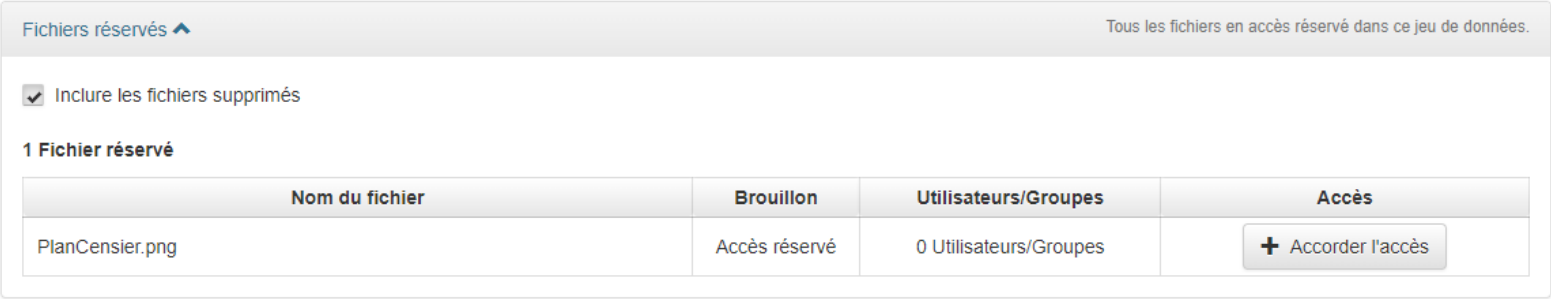
A noter : Cette option ne doit être utilisée que de **manière exceptionnelle.** Elle n’est **pas conforme aux principes de la Science ouverte**.
Nous demander conseil avant de l’appliquer.



Attention : L’accès restreint permet à l’utilisateur de limiter les téléchargements. Il ne doit être utilisé qu’en de rares cas. Il requiert une action de l’auteur de la fiche pour chaque demande !

L’accès restreint





Vous verrez apparaître la liste des personnes demandant l’accès aux fichiers réservés à cet endroit. Vous pourrez accepter ou refuser.

Si votre jeu de données est associé à une future publication et que vous souhaitez communiquer les informations aux relecteurs, ne pas utiliser cette technique mais la création d’URL privée expliquée précédemment.

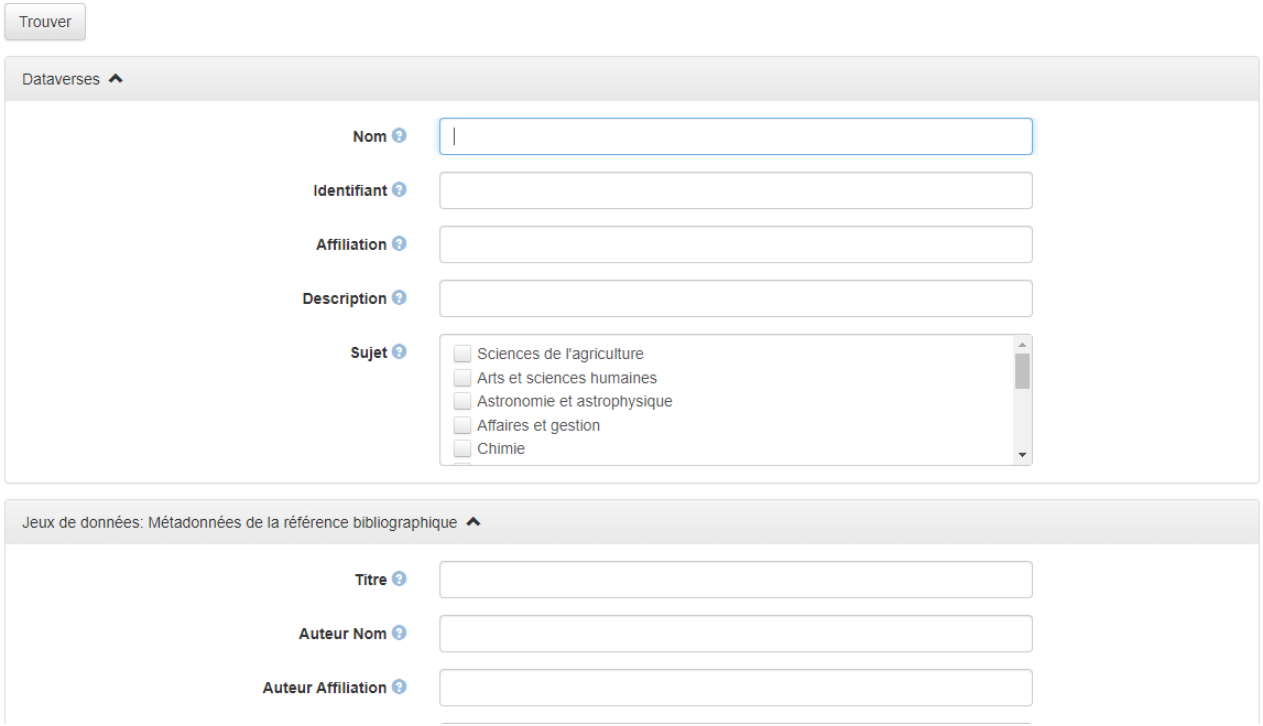
Recherche d’un jeu de données

Différents types de recherche

Recherche avancée

Recherche avancée sur des champs spécifiques.

Recherche simple en tapant un ou plusieurs mots-clés.



Différents types de recherche

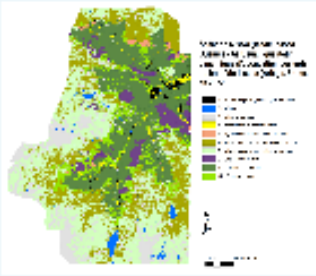
Recherche par facette

Réutilisation d’un jeu de données

Quelques points de vigilance pour une réutilisation

Comme dans tout entrepôt de données, vérifier les conditions d'utilisation dans l'onglet dédié. Par défaut, les jeux de données sur data.InDoRES sont en licence CC-BY 4.0 (réutilisation libre mais citation de la paternité obligatoire).

Quelques points de vigilance pour une réutilisation



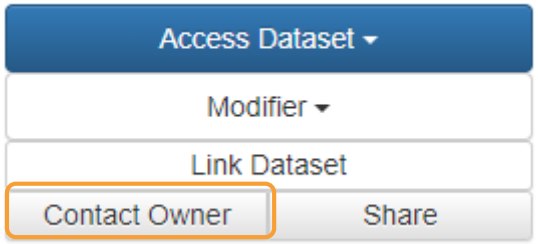
Houet, Thomas, 2023, "Simulated land cover maps for 2080 on the Haut Vicdessos (Pyrenees, France) considering two pastoral land use change scenarios and model parameterization", https://doi.org/10.48579/PRO/BWJSR9, data.InDoRES, V1

Citer le jeu de données

Pour en savoir plus, consultez le document Data Citation Standards.

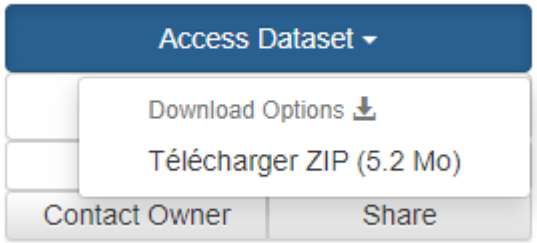
La citation se situe dans l’encart bleu en haut de chaque fiche.

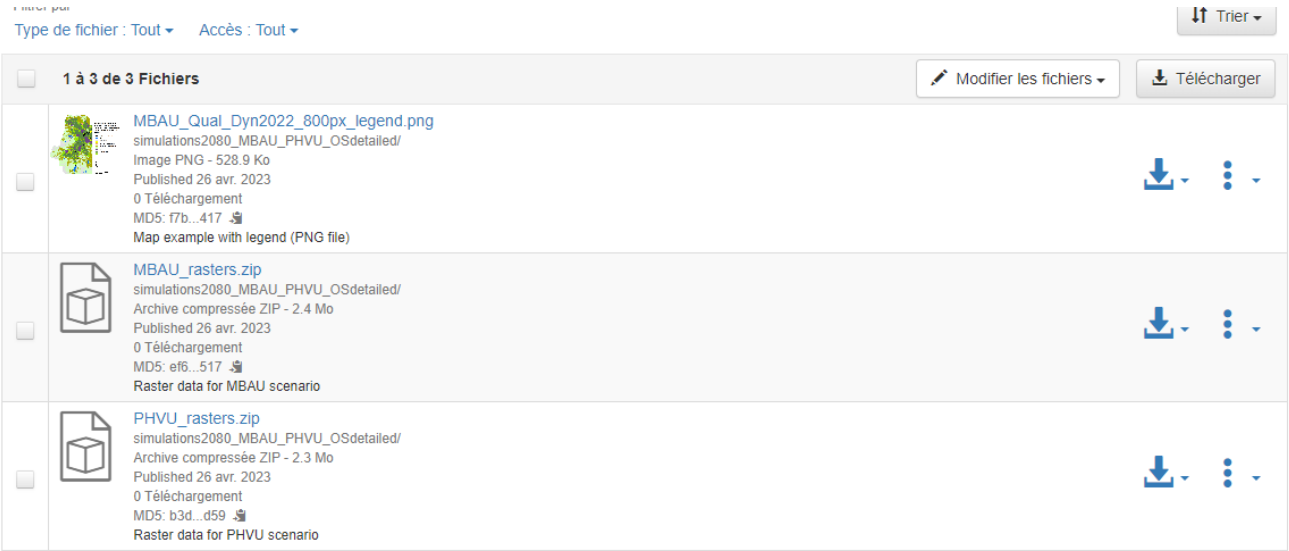
Contacter le producteur du jeu de données



Si vous souhaitez contacter le producteur d'un jeu de données, vous pouvez utiliser le bouton dédié.

Télécharger le(s) fichier(s)





- Vous pouvez télécharger l'ensemble des fichiers contenus dans un jeu de données via le bouton en haut à droite.
- Si vous ne souhaitez pas télécharger la totalité des fichiers, vous pouvez faire une sélection en cochant et cliquant sur le bouton télécharger.

Pour toute question : contact@indores.fr

Callou C. & Hénon A., *Guide utilisateurs data.InDoRES*, 2023

https://creativecommons.org/licenses/by/4.0/