



ÉCOLOGIE &
ENVIRONNEMENT

Publier ses données dans un entrepôt

Amandine Hénon – UAR DoHNEE



Entrepôt, catalogue : quelles différences ?



L'entrepôt permet de stocker, d'accéder et d'utiliser des données. Il peut être thématique (data.InDoRES), institutionnel (data.INRAE) ou générique (Zenodo). Il assure la sécurisation des données et leur stockage sur le moyen terme.

Le catalogue de métadonnées est un annuaire regroupant des fiches « carte d'identité » autour de projets, de bases ou de jeux de données. Il permet de rassembler les informations et les ressources externes au même endroit.

Publication dans un entrepôt : non modifiable ou supprimable contrairement au fiche du catalogue qui évolue avec le projet.



Pourquoi publier ses données dans un entrepôt ?

Stockage pérenne vs. diffusion

- **Finalité différente** : stockage pour sauvegarder ses données mais publication pour partager de manière sécurisée et permettre la réutilisation.

Pourquoi publier ses données dans un entrepôt ?

Stockage pérenne vs. diffusion

- **Finalité différente** : stockage pour sauvegarder ses données mais publication pour partager de manière sécurisée et permettre la réutilisation.
- « **Choix** » **des données** : on ne publie pas toutes les données, on pourrait vouloir tout sauvegarder.

Pourquoi publier ses données dans un entrepôt ?

Stockage pérenne vs. diffusion

- **Finalité différente** : stockage pour sauvegarder ses données mais publication pour partager de manière sécurisée et permettre la réutilisation.
- « **Choix** » **des données** : on ne publie pas toutes les données, on pourrait vouloir tout sauvegarder.
- **Nettoyage et documentation** : la publication dans un entrepôt demande un travail de documentation (métadonnées) et une structuration des fichiers.

Pourquoi publier ses données dans un entrepôt ?

Stockage vs. diffusion

- **Finalité différente** : stockage pour sauvegarder ses données mais publication pour partager de manière sécurisée et permettre la réutilisation.
- « **Choix** » **des données** : on ne publie pas toutes les données, on pourrait vouloir tout sauvegarder.
- **Nettoyage et documentation** : la publication dans un entrepôt demande un travail de documentation (métadonnées) et une structuration des fichiers.
- **Un timing différent** : la publication dans un entrepôt intervient souvent à une étape ultérieure du cycle de vie alors que le stockage se fait dès la collecte.

**Quels critères
pour choisir
un entrepôt ?**

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Politique de l'entrepôt

Possible incompatibilité

Métadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à
tous ?

Identifiants uniques

Présence de DOI, ORCID

Interopérabilité

Liens avec les portails
européens et les projets
fédérateurs

Pérennité

L'entrepôt est-il soutenu ?

**Quels critères
pour choisir
un entrepôt ?**

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Si un entrepôt disciplinaire existe :
métadonnées, thésaurus ou formats de
données acceptés pourront être plus proches
de vos données.

Coût et accès

Gratuité ? Accès ouvert à
tous ?

Identifiants uniques

Présence de DOI, ORCID

Interopérabilité

Liens avec les portails
européens et les projets
fédérateurs

Pérennité

L'entrepôt est-il soutenu ?

**Quels critères
pour choisir
un entrepôt ?**

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Politique de l'entrepôt

Possible incompatibilité

Étatadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à
tous ?

Identifia unique

Présence de DOI,

- Formats de données acceptés
- Taille maximale par fichier/Volumétrie
- Respect RGPD
- Quelle est la licence apposée au jeu de données ?
- Y a-t-il une étape de modération et un accompagnement ?
- Quid des partenaires non académiques ?
- Où sont stockées les données ?

ennité

est-il soutenu ?

Quels critères pour choisir un entrepôt ?

- Standard interopérable de métadonnées ?
- Informations minimales requises
- Préconisation de thésaurus pour les mots-clé
- Présence d'un ReadMe pour accompagner les données

Métadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à tous ?

Identifiants uniques

Présence de DOI, ORCID

Interopérabilité

Liens avec les portails européens et les projets fédérateurs

Pérennité

L'entrepôt est-il soutenu ?

**Quels critères
pour choisir
un entrepôt ?**

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Politique de l'entrepôt

Métadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à
tous ?

- La gratuité du service est-elle garantie ?
- L'entrepôt est-il ouvert à tous sans conditions ?
- Un compte est-il nécessaire pour accéder aux données ?

Pérennité

L'entrepôt est-il soutenu ?

**Quels critères
pour choisir
un entrepôt ?**

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Politique de l'entrepôt

Possibilité d'interopérabilité

Métadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à
tous ?

Identifiants uniques

Présence de DOI, ORCID

- Un DOI est-il attribué à chaque jeu de données ?
- Ce DOI est-il « versionnable » ?
- L'ORCID est-il implémenté ou recommandé ?
- Identifiant d'établissement et de structure : ROR ?

Quels critères pour choisir un entrepôt ?

Disciplinaire

Spécificités disciplinaires ou généralistes ?

Politique de l'entrepôt

Possible incompatibilité

Métadonnées

Qualité de l'information

- L'entrepôt est-il moissonné par les portails nationaux et européens ?
- L'interopérabilité des systèmes est-elle prise en compte (API ?) ?
- Présence de thésaurus ?
- Possibilité de lier des ressources

Interopérabilité

Liens avec les portails européens et les projets fédérateurs

Pérennité

L'entrepôt est-il soutenu ?

**Quels critères
pour choisir
un entrepôt ?**

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Politique de l'entrepôt

Possible incompatibilité

Métadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à
tous ?

- L'entrepôt est-il soutenu par un organisme de recherche ?
- Y a-t-il des personnels statutaires dédiés ?
- Une réflexion de pérennisation est-elle engagée par l'entrepôt ?

Pérennité

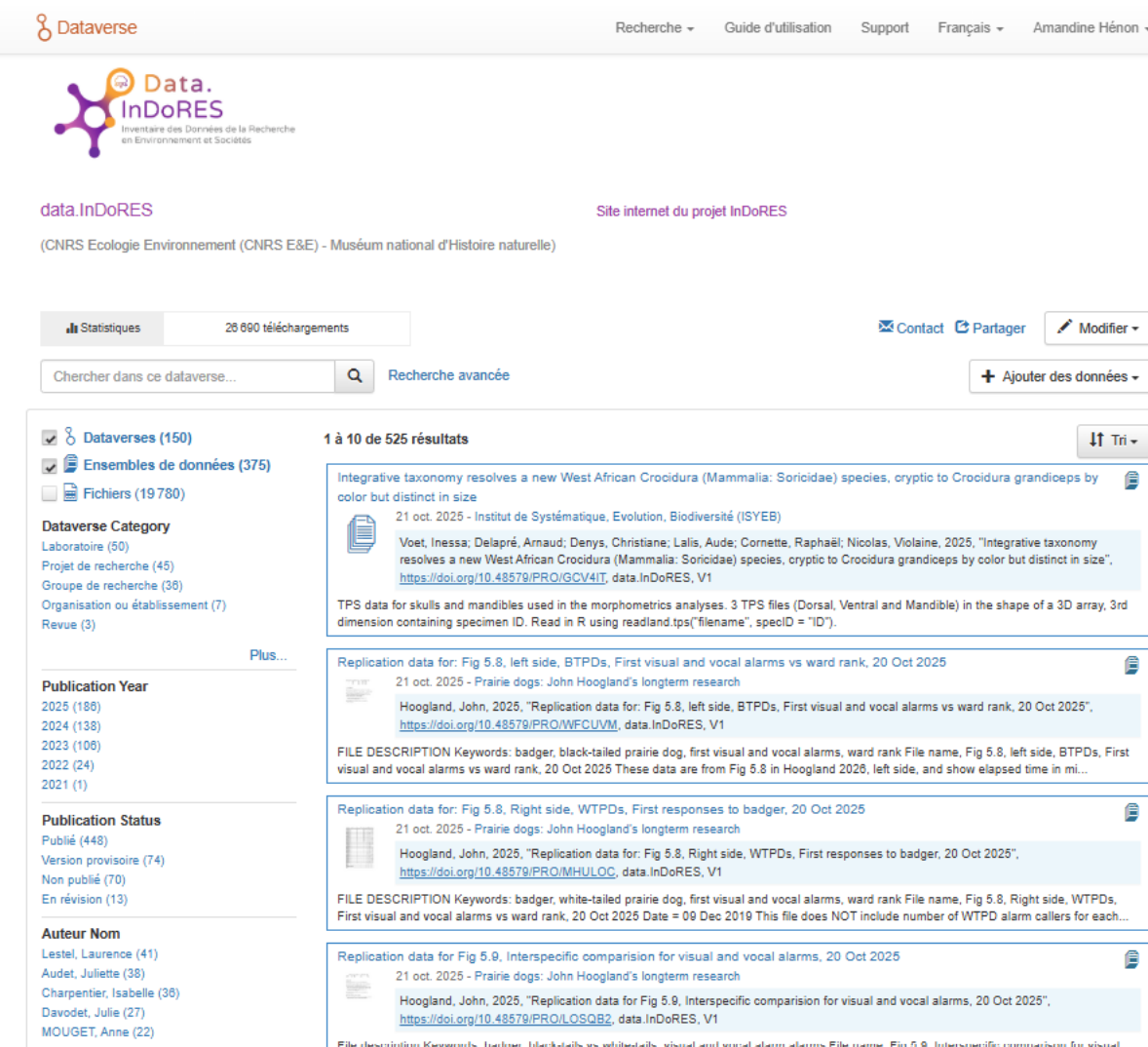
L'entrepôt est-il soutenu ?



Data.InDoRES : entrepôt de données sécurisé

Fonctionnalités

- Solution adoptée : Dataverse déploiement en adéquation avec les entrepôts déjà existants
- Stockage des données de manière sécurisée et pérenne au Centre de Calcul de l'IN2P3 (infrastructure nationale de recherche).
- Structuration par unité ou par projet pour une visibilité accrue.
- Interface français/anglais



The screenshot displays the DataInDoRES website, which is part of the Dataverse ecosystem. The header includes the Dataverse logo and navigation links for Recherche, Guide d'utilisation, Support, Français, and Amandine Hénon. The main content area shows the DataInDoRES logo and its affiliation with CNRS Ecologie Environnement and the Muséum national d'Histoire naturelle. Below this, there are statistics (28 690 téléchargements) and buttons for Contact, Partager, and Modifier. A search bar is present with the text "Chercher dans ce dataverse..." and a "Recherche avancée" link. The search results section shows a list of datasets with filters for Dataverses (150), Ensembles de données (375), and Fichiers (19 780). The results are sorted by Publication Year, with 2025 having 186 entries. The first result is titled "Integrative taxonomy resolves a new West African Crocidura (Mammalia: Soricidae) species, cryptic to Crocidura grandiceps by color but distinct in size" and includes a DOI link. The second result is titled "Replication data for: Fig 5.8, left side, BTPDs, First visual and vocal alarms vs ward rank, 20 Oct 2025" and includes a DOI link. The third result is titled "Replication data for: Fig 5.8, Right side, WTPDs, First responses to badger, 20 Oct 2025" and includes a DOI link. The fourth result is titled "Replication data for Fig 5.9, Interspecific comparison for visual and vocal alarms, 20 Oct 2025" and includes a DOI link.

Fonctionnalités

- DOI (Digital Object Identifier) délivré par DataCite
- Licence CC-BY 4.0
- Possibilité embargo jusqu'à publication article
- Diversité de formats acceptés (liste de recommandations de formats disponibles sur InDoRES)
- Possibilité de lier un jeu de données à plusieurs sous-entrepôts (un projet ANR et des UMR par exemple)



Data for "Reviewing Seas of Data: Integrating Image-Based Bio-logging and Artificial Intelligence to Enhance Marine Conservation"

Version 1.1

KATO, Akiko; ROPERT-COUDERT, Yan, 2025, "Data for "Reviewing Seas of Data: Integrating Image-Based Bio-logging and Artificial Intelligence to Enhance Marine Conservation", <https://doi.org/10.48579/PRO/GU2J> [CY](#), data.InDoRES, V1

Citer l'ensemble de données ▾ Pour en apprendre davantage sur le sujet, consulter le document [Data Citation Standards \[en\]](#).

Modalités d'accès à l'ensemble de données ▾

Modifier l'ensemble de données ▾

Lier l'ensemble de données

Contacteur le propriétaire Partager

Statistiques d'utilisation sur l'ensemble de données ⓘ

57 téléchargements ⓘ

Description ⓘ

This dataset contains data and Python codes associated to the manuscript: Reviewing Seas of Data: Integrating Image-Based Bio-logging and Artificial Intelligence to Enhance Marine Conservation.

Data: Underwater images and sample video collected from Adélie penguins (*Pygoscelis adeliae*) at the colony on Ile des Pétrels, at the Dumont d'Urville station, in Terre Adélie, East Antarctica (66°40' S; 140°01' E), over the 2017-2018 breeding season. Birds were equipped with video loggers (Little Leonardo, DVL400M, 61 mm × 21 mm × 15 mm, 29g) during the chick guarding stage, in which one parent guards the chicks on the nest while the other is foraging at sea to bring food back to its offspring. Only one member of a pair was captured on or when leaving the nest, when both adults were attending the nest half-attended is a compromise. This data collection is part of the research team.

[Lire la suite de Description \(+\)](#)

Sujet ⓘ

Sciences de la terre et de l'environnement

Mot-clé ⓘ

ecology, conservation, marine ecosystems, big data, artificial intelligence, computer vision, animal behaviour, data recording, underwater image, foraging

Licence/Accord d'utilisation des données

CC BY 4.0

Fichiers Métadonnées Conditions Versions

Changer l'affichage Tableau Arborescence

Chercher dans cet ensemble de données ... 🔍

+ Téléverser des fichiers

Filtrer par

Type de fichier : Tout ▾ Accès : Tout ▾

1 à 10 de 17 Fichiers

Modifier les fichiers ▾ Télécharger

	dataset.yaml Inconnu - 248 o Publié 31 mars 2025 4 téléchargements MD5: 62b...855	
	How_To_Image_Based_Bio-logging.ipynb Notebook Jupyter - 3.6 Mo Publié 31 mars 2025 5 téléchargements MD5: 0c9...c79	

Le réseau de référents OpenDoRES



- Formation à la Science ouverte et aux principes FAIR pour accompagner les structures (faire remonter demandes et questions, les réponses venant enrichir le portail InDoRES).
- Implication dans **l'administration de l'entrepôt de la structure** (Structure opérationnelle de recherche ou projet) pour permettre une autonomie de publication des données. La gestion technique est prise en charge par l'UAR BBEES en collaboration avec les équipes du CCIN2P3
- Liste de diffusion pour faciliter le dialogue, afin de poser les questions et partager les ressources & initiatives.

Contact : contact@indores.fr

Checklist des critères pour data.InDoRES

Disciplinaire

Spécificités disciplinaires
ou généralistes ?

Politique de l'entrepôt

Possible incompatibilité

Métadonnées

Qualité de l'information

Coût et accès

Gratuité ? Accès ouvert à
tous ?

Identifiants uniques

Présence de DOI, ORCID

Interopérabilité

Liens avec les portails
européens et les projets
fédérateurs

Pérennité

L'entrepôt est-il soutenu ?

Démo et prise en main

Entrepôt test « bac à sable » : <https://dataverse-test.in2p3.fr/>

Licence

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



Attribution 4.0 International (CC BY 4.0)

Vous êtes autorisé à :

- **Partager** — copier, distribuer et communiquer le matériel par tous moyens et sous tous formats
- **Adapter** — remixer, transformer et créer à partir du matériel pour toute utilisation, y compris commerciale.

Selon les conditions suivantes :

Attribution — Vous devez créditer l'Œuvre, intégrer un lien vers la licence et indiquer si des modifications ont été effectuées à l' Œuvre. Vous devez indiquer ces informations par tous les moyens raisonnables, sans toutefois suggérer que l'Offrant vous soutient ou soutient la façon dont vous avez utilisé son Œuvre.

Pas de restrictions complémentaires — Vous n'êtes pas autorisé à appliquer des conditions légales ou des [mesures techniques](#) qui restreindraient légalement autrui à utiliser l' Œuvre dans les conditions décrites par la licence.

Attribution : A. Henon – Ecole thématique Data SEE-Life - 2025