

Comment publier ses données de recherche

Pour accéder à la ressource : https://doranum.fr/depot-entrepots/comment-publier-ses-donnees-de-recherche_10_13143_ajk3-ea67/

Date de publication : 25/01/2024

1. Introduction

Les données sont habituellement intégrées partiellement dans les articles scientifiques (sous forme de tableaux ou graphiques...).

Il existe trois actions complémentaires qui permettent un accès à l'intégralité de ses données scientifiques :

- Déposer ses données dans un **entrepôt**
- Publier ses données dans un **data paper**
- Lier ses données aux publications

2. Déposer ses données dans un entrepôt

C'est le **premier réflexe à avoir pour publier ses données**. Les entrepôts offrent un environnement dédié à la diffusion et la documentation des données, pour une ouverture et une réutilisation optimales de celles-ci.

Point de vigilance : attention à bien choisir l'entrepôt ! En fonction des entrepôts et des standards de métadonnées utilisés, la description des métadonnées peut être plus ou moins riche.

Avantages

- **Qualité** : nombreux entrepôts connus et reconnus de la communauté scientifique (dont certains certifiés CoreTrustSeal)

- **Orientation disciplinaire** : diffusion orientée vers une communauté d'intérêt spécifique si on choisit un entrepôt disciplinaire
- **Principes FAIR** : plateformes permettant de partager ses données selon les principes FAIR
- **Identifiant pérenne** : la plupart des entrepôts attribue automatiquement un identifiant pérenne
- **Standard de métadonnées** : la plupart des entrepôts impose un standard de métadonnées pour favoriser l'interopérabilité.
- **Format ouvert** : certains entrepôts proposent une conversion automatique de vos fichiers vers un format ouvert
- **Durée de conservation** : garantie sur plusieurs années, variable selon les entrepôts
- **Volumétrie** : pas/peu de restriction en volume de données
- **Conditions d'utilisation** : les entrepôts permettent d'attribuer des licences et de définir des conditions d'accès sur certains fichiers (restrictions, embargos...)
- **Citation** : génération d'une citation standardisée du jeu de données
- **Visibilité** : découverte des données facilitée
- **Réutilisation des données facilitée**

3. Publier ses données dans un data paper

Définition : « Un data paper est une publication dont le but est la description d'un jeu de données scientifiques. Contrairement à un article de recherche classique, le data paper consiste en une description détaillée des données scientifiques, de leurs métadonnées, ainsi que des circonstances et méthodes de leur collecte, mais sans analyse ou interprétation de ces données. »

MESR, Ministère de l'Enseignement Supérieur et de la Recherche. Partager les données liées aux publications scientifiques. Guide pour les chercheurs. Mars 2022. <https://www.ouvrirlascience.fr/partager-les-donnees-liees-aux-publications-scientifiques-guide-pour-les-chercheurs/>

Un **data paper** peut paraître dans des revues scientifiques traditionnelles ou dans des revues spécifiques (data journals).

Le data paper s'intègre dans le processus éditorial des articles classiques et reste **le moyen valorisant le mieux les données**. En fonction des disciplines, il peut venir en complément d'une publication classique ou la remplacer.

Avantages

- **Binôme data paper / entrepôt** : le fait de déposer au préalable dans un entrepôt les données décrites dans le data paper permet de s'assurer d'avoir des données standardisées et accessibles
- **Identifiant pérenne** : attribution d'un identifiant pérenne pour le data paper (par l'éditeur) qui fait le lien avec celui des données (via l'entrepôt dans lequel elles ont été déposées)
- **Contexte de création des données ou de réutilisation de données préexistantes** : description et documentation très précise des données et des métadonnées
- **Attribution** : paternité des données / créditer les auteurs
- **Peer-reviewing** : même processus que pour une publication classique
- **Volumétrie** : pas de restriction en volume de données (hormis une éventuelle restriction de l'entrepôt)
- **Visibilité accrue** : indexation dans les bases de données, double citation (du data paper et du jeu de données)
- **Réutilisation des données** : elle est facilitée par la richesse des métadonnées.

4. Lier ses données à ses publications

Il y a **deux actions réciproques à effectuer pour lier un article et des données** :

- Avant la publication, déposer ses données dans un entrepôt et citer leur identifiant pérenne dans l'article.
- Après la publication, renseigner dans l'entrepôt l'identifiant pérenne de l'article associé aux données.

Point de vigilance :

Indiquer le lien vers les données après la publication d'un article est le plus souvent impossible à l'heure actuelle.

La seule alternative consiste à déposer son article dans une archive ouverte (HAL par exemple) et de renseigner l'identifiant pérenne de l'article et des données dans les champs spécifiques dédiés aux données associées de l'enregistrement.

Avantages

- **Visibilité accrue des données et de l'article** : citation, augmentation du potentiel de réutilisation
- **Attribution** : paternité des données / créditer les auteurs
- **Transparence et reproductibilité du processus scientifique** : données plus accessibles
- **Peer-reviewing facilité**
- **Volumétrie** : pas de restriction en volume de données

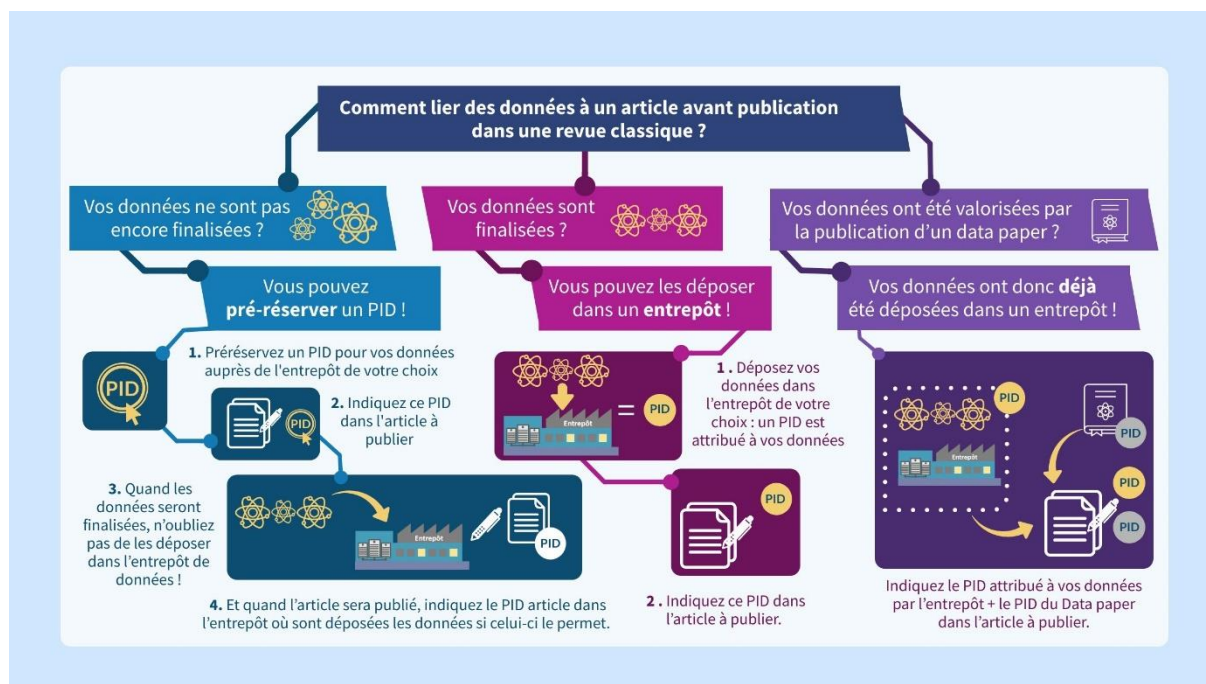
Il est recommandé de ne pas publier ses données dans les « supplementary data » ou « supplementary materials ». Les données ne sont pas correctement documentées et cela ne facilite pas leur réutilisation. De plus, cela évite une demande de transfert exclusif de droits à l'éditeur.

5. Processus idéal

L'idéal est de procéder ainsi :

- Déposer ses données dans un entrepôt
- Rédiger et publier un article scientifique en liant bien son article aux données (déposées dans un entrepôt)
- Rédiger et publier un data paper.

6. Lier ses données à un article avant publication dans une revue classique



Webographie :

- Bracco Laetitia, Bouchet-Moneret Florence, Jouneau Thomas. *Le chemin de la publication article / données de la recherche*. 19 juin 2023. <https://doi.org/10.5281/zenodo.8054302>
- Couperin. *GTSO Données. Data papers : quand ? comment ? pourquoi ?* https://doranum.fr/data-paper-data-journal/data-papers-quand-comment-pourquoi_10_13143_sx95-cf48/
- INRAE. *Publier un Data Paper pour valoriser ses données*. <https://science-ouverte.inrae.fr/offre-service/formations/open-class/open-class-publier-un-data-paper>
- MESR, Ministère de l'Enseignement Supérieur et de la Recherche. *Partager les données liées aux publications scientifiques. Guide pour les chercheurs*. Mars 2022. <https://www.ouvrirlascience.fr/partager-les-donnees-liees-aux-publications-scientifiques-guide-pour-les-chercheurs/>