


Critères pour choisir un dépôt de données

Recommandations

- 📄 Partagez vos données dans un dépôt de données FAIR, non-commercial et persistant qui assure la préservation à long terme (Exemple SWISSUbase, OLOS, Zenodo, ...)
- 📄 Choisissez des dépôts qui proposent d'attribuer automatiquement des DOI (Digital Object Identifier) ou qui permettent leur ajout
- 📄 Choisissez des dépôts qui proposent des licences Creative Commons (CC0 ou CC BY)

Afin de vous aider à choisir le dépôt de données le plus adapté à vos besoins mais qui soit également conforme aux [principes FAIR](#), voici une liste de **11** critères à considérer :






1. Le dépôt propose-t-il d'attribuer un identifiant numérique unique et pérenne au jeu de données ?

- [L'identifiant unique et durable](#) ou PID (Persistant Identifier), permet de référencer, citer et fournir un lien stable vers un fichier en ligne (Deboin, 2017). Un lien hypertexte unique et pérenne est créé et permet de retrouver la ressource en tout temps, même si l'adresse URL de la page change.
 - Il existe plusieurs types de PID. Le plus connu pour identifier des jeux de données ou des articles de revues est le **DOI** (Digital Object Identifier).
 - Exemple d'un DOI rattaché à une ressource sur disponible sur le dépôt de données Zenodo 
 - L'identifiant pérenne est intégré dans la référence bibliographique. Ainsi si le jeu de données est réutilisé, il sera facilement retrouvé depuis la citation.

Exemple : Valérie Gadrat, Yvette Lafosse, Claire Sowinski, & Coralie Wysoczynski. (2019, April). Jeu GopenDoRe : le jeu coopératif pour acquérir des bonnes pratiques de gestion et de partage des données de la recherche (Version 2). Zenodo. <http://doi.org/10.5281/zenodo.2657316>

2. La licence d'utilisation sous laquelle les données seront accessibles est-elle clairement mentionnée ou l'utilisatrice-teur peut-il télécharger/choisir une licence ?

- Certains dépôts imposent une licence alors que d'autres proposent un choix de licences.
- Lorsque les données sont liées à un article scientifique, il est nécessaire de vérifier si la revue applique des conditions spécifiques pour la licence de diffusion des données. Pour cela, consultez les instructions aux auteurs.

- [Les licences Creative Commons](#) ont été créées en 2002 pour la diffusion de contenus numériques comme le texte, les images et les films. Mais elles peuvent également être utilisées pour une diffusion sur papier. Elles permettent aux auteurs d'indiquer facilement les droits qu'ils veulent conserver et les droits auxquels ils renoncent afin de permettre à d'autres de réutiliser leur œuvre (Fly, 2015).
 - Quatre clauses peuvent être combinées en fonction de ses besoins :
 -  *Attribution* (sigle **BY**) : paternité, c'est-à-dire citation de l'auteur initial (obligatoire).
 -  *Non Commercial* (sigle **NC**) : interdiction de tirer un profit commercial de l'œuvre sans autorisation de l'auteur.
 -  *No Derivative Works* (sigle **ND**) : impossibilité d'intégrer tout ou partie dans une œuvre composite.
 -  *Share Alike* (sigle **SA**) : partage de l'œuvre, avec obligation de rediffuser selon la même licence ou une licence similaire.
 - Ou il est possible d'utiliser la licence du domaine public
 -  Universel (sigle **0**) : transfert dans le domaine public, renoncement à tous les droits, aucune restriction de réutilisation. Licence créée pour faciliter la réutilisation des jeux de données.

Pour aller plus loin

- [Présentation des creatives commons en vidéo](#)
- [Connaitre et utiliser les licences creatives commons](#) : guide CoopIST

3. Peut-on charger et/ou rajouter des métadonnées (p. ex. nom de l'auteur-e, contenu du set de données, publications associées, etc.) ?

- **Les métadonnées** sont des éléments qui servent à décrire de manière normalisée et structurée, le but, l'origine, les caractéristiques temporelles, l'emplacement géographique, la paternité, les conditions d'accès et les conditions d'utilisation d'une ressource, comme un jeu de données (Corti, 2020). Elles facilitent la recherche du jeu de données ainsi que sa compréhension. Plus on communique des informations sur le jeu de données et plus il sera compréhensible et facile à trouver.
 - Il existe différents schémas de description, dont le [Dublin Core](#), [Data Documentation Initiative](#) (DDI), [Metadata Encoding and Transmission Standard](#) (METS), le [schéma de métadonnées DataCite](#)...

4. Les citations et les métadonnées sont-elles toujours accessibles publiquement (même dans le cas de sets de données à accès restreint) ?

- Rendre les métadonnées accessibles même si les données ne peuvent être ou ne sont plus mises à disposition permet de répondre aux critères FAIR *Accessible* et *Findable*. (ex. les métadonnées auteurs, institutions, publications associées peuvent être utiles même si les données sont manquantes). De plus cette solution permet de limiter les coûts de stockage

5. Le dépôt de données fournit-il un formulaire de soumission demandant que les métadonnées intrinsèques respectent un format spécifique (afin d'assurer leur utilisation automatique/interopérabilité) ?

- Afin qu'elles puissent être interprétables par des machines, les métadonnées (descriptives, administratives et structurelles) doivent être conformes à des schémas standards comme DataCite.

6. Le dépôt de données dispose-t-il d'un plan de préservation à long terme des données archivées ?

- Le plan de préservation va fournir des informations sur les actions mises en place par le dépôt pour assurer sa gestion et son financement à long terme ainsi que la préservation des données. " Plus le dépôt s'oriente vers la préservation, plus son financement, la gestion de ses ressources et sa gouvernance doivent être solides" (Guirlet, 2020, p. 27)
- Le cadre COAR de bonnes pratiques en matière d'entrepôts propose 4 caractéristiques (2 essentielles, 2 recommandées) pour l'objectif de préservation (Confederation of Open Access Repositories, 2020):
 - "L'entrepôt (ou l'organisme qui gère l'entrepôt) a un plan à long terme pour la gestion et le financement de l'entrepôt.
 - L'entrepôt fournit de la documentation ou a une politique qui définit la durée pendant laquelle les ressources seront gérées sur le long terme et de la documentation sur les pratiques de préservation.
 - L'entrepôt a une approche documentée de la préservation, qui adopte des pratiques reconnues de préservation
 - L'accord entre le déposant et l'entrepôt prévoit toutes les actions nécessaires pour assurer les responsabilités de préservation - par exemple les droits de copie, de transformation et de stockage des ressources".

7. Le dépôt de données est-il adapté au type de données que vous allez déposer ?

- Le dépôt accepte-t-il tous les formats de fichier ?
- Existe-t-il une limitation en termes de taille ?
- Tous les formats de données sont-ils acceptés ?

8. Répond-il aux recommandations du bailleur de votre projet, de votre institution, ou de la revue dans laquelle vous publiez ?

- FNS : les données de la recherche devraient être libres d'accès
 - Obligation d'inclure un data management plan (DMP) lors de la soumission d'une requête pour la plupart des instruments d'encouragement
 - Archiver en libre accès les données produites durant la recherche sur un dépôt de données ouvert, pour autant qu'aucune clause juridique, éthique, concernant le droit d'auteur ou autre ne s'y oppose
 - Le [FNS octroie des fonds jusqu'à CHF 10'000](#) pour la préparation des données, la mise en ligne et la validation. Le téléchargement n'est pas pris en compte.
- [Les autres pays européens](#)

9. Est-il reconnu dans votre discipline et par la communauté scientifique ([certains entrepôts, encore peu nombreux, sont certifiés](#)) ?

- Le FNS propose une [liste d'entrepôts](#) qui répondent à leurs exigences en matière d'ORD et qui sont fréquemment utilisés par la communauté.
 - Généralistes : Zenodo, SWISSUbase, OLOS, GitHub...
 - Disciplinaires : DaSCH, ENA, PlaTec...
- La liste de dépôts recommandés par la revue [Nature](#)
- Consulter des répertoires d'entrepôts pour choisir le dépôt le plus adapté : [Re3Data](#), [Fairsharing](#), [OpenAIRE](#)

10. Est-il gratuit (la plupart des entrepôts) ou pratique-t-il des coûts de dépôt de données acceptables ?

11. Propose-t-il des modalités d'accès aux données, adaptés à vos besoins

- Accès libre, après enregistrement, accès restreint, sur demande, différé par un embargo ?

Tutoriels

- Attribuer un DOI à un jeu de données : Il est possible d'obtenir un DOI en envoyant un mail à l'adresse open(at)hes-so.ch
- [Sélectionner une licence à l'aide du sélectionneur en ligne](#)

Références

Confederation of Open Access Repositories. (2020, novembre 6). *Cadre commun COAR de bonnes pratiques en matière d'entrepôts*. Zend.

<https://doi.org/10.5281/zenodo.4118380>

Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2020). *Managing and sharing research data : A guide to good practice* (Second edition). SAGE

Dedieu, L. ; Barale, M. 2020. Déposer des données dans un entrepôt, en 6 points. CIRAD.

<https://doi.org/10.18167/coopist/0070>

Deboin, M. C. (2017). *Identifier et rechercher une publication ou un jeu de données par son DOI*. CIRAD. <https://doi.org/10.18167/COOPIST/0005>

Fily, M.-F. (2015). *Connaitre et utiliser les licences Creative Commons*. CIRAD.

<https://doi.org/10.18167/XTNV-D457>

Fonds national suisse. (2022). Open Research Data. FNS.

<https://www.snf.ch/fr/dMILj9t4LNk8NwyR/dossier/open-research-data>

Guirlet, M. (2020). *Guide décisionnel et vade-mecum pour la mise à disposition d'un dépôt de données de recherche ouvertes en Suisse* [Haute école de gestion. Travail de Master]. https://doc.rero.ch/record/329696/files/Guirlet-M_moire-Vdef.pdf

Scientific Data. (2022a). *Data Policies*. Nature. <https://www.nature.com/sdata/policies/data-policies>

Scientific Data. (2022b). *Data Repository Guidance*. Nature.

<https://www.nature.com/sdata/policies/repositories>



Critères pour choisir un dépôt de données © 2021 by Groupe de travail
Guidelines de la Communauté Open Science HES-SO is licensed under [Attribution
4.0 International](#)

