

Où diffuser ses données de recherche ?

0. Qu'est-ce qu'un *repository* ?

Un *repository* est un dépôt de données de la recherche, c'est-à-dire une infrastructure numérique mise en place pour stocker, partager, accéder et archiver les jeux de données des chercheurs.

1. Pourquoi partager des données dans un *repository* ?ⁱ



A. Exigence de partage de données sources parfois imposée par

- Bailleurs de fonds
- Revues scientifiques
- Institutions et universités

B. Bénéfices pour vous

- Citations
- Reconnaissance académique et professionnelle
- Nouvelles collaborations

C. Bénéfices pour la communauté de recherche

- Réutilisation pour l'enseignement
- Plus grande facilité à trouver des données utiles
- Archivage et préservation des données pour le futur
- Possibilité de nouvelles recherches
- Recherche plus efficiente

D. Bénéfices pour le public

- Opportunités pour la science citoyenne
- Bénéfices économiques dans le secteur privé
- Renforcement de la confiance publique dans la science
- Validation des données dans les publications de recherche

2. Où partager ?



Il existe plusieurs solutions pour ouvrir ses données :

A. *Repository* générique

- + Convient pour tous les types de données de recherche, est destiné à garantir la préservation à long terme, fournit un DOI pour un accès et un référencement permanent
- Vérifier que les termes et conditions sont bien conformes avec la politique du bailleur de fond, du journal ou de l'institution (droits d'auteur, préservation long terme)

B. *Repository* disciplinaire

- + Spécialisé dans un domaine, offre une expertise sur le stockage et l'utilisation spécifiques au type de données
- Plus sélectif (standards de métadonnées et de documentation plus élevés)

C. *Repository* institutionnel

- + Accès à long terme, assurance de conformité à la politique de l'institution
- N'existe pas toujours, peut avoir moins d'impact en matière de visibilité

Les options suivantes n'offrent pas de garanties suffisantes sur le long terme. Elles peuvent être utilisées mais devraient l'être uniquement en complément des *repositories* ci-dessus.

🌟 Service proposé par une revue (*supplementary materials*)

- Pas de garantie de pérennité à long terme (selon l'éditeur), peut être onéreux (tant pour le dépôt que pour la consultation), peu de flexibilité sur les formats acceptés, uniquement pour les données directement liées à une publication.

🌟 Page internet personnelle

- N'est pas visible pour les nouveaux utilisateurs, ne permet pas un accès à long terme.

3. Comment partager les données : principes FAIR



Data FAIRport est une initiative lancée en 2014. Son objectif est d'assurer de bonnes conditions pour la découverte et la réutilisation des données tant par les humains que par les ordinateurs.

A. Findable (trouvable)

Les données sont décrites avec des métadonnées, un identifiant unique et pérenne et sont enregistrées et indexées dans un système permettant des recherches.

B. Accessible (accessible)

Les (méta)données peuvent être récupérées avec leur identifiant en utilisant un protocole de communication standardisé ouvert, gratuit et universel. Si nécessaire, le protocole offre une fonction d'authentification. Les métadonnées sont accessibles en tout temps.

C. Interoperable (interopérable)

Les (méta)données utilisent un langage formel, accessible, partagé et largement applicable, un vocabulaire FAIR, et contiennent des références contrôlées vers d'autres (méta)données.

D. Reusable (réutilisable)

Les données et les métadonnées sont suffisamment décrites pour permettre la réutilisation des données dans les recherches futures : contenu, contexte et condition de production, variables, unités, licence d'usage claire et accessible, etc. Elles respectent les standards définis par la communauté scientifique.

4. Trouver le bon *repository* : options à disposition



A. Yareta

Yareta est le *repository* institutionnel de l'Université de Genève, développé grâce au projet DLCM et au Projet de loi 12146. Il permet la préservation à long-terme et la mise en ligne des données issues de la recherche selon la norme ISO 14721 et les principes FAIR. Chaque dépôt est doté d'un DOI. Le service est ouvert aux chercheurs de l'UNIGE, de l'IHEID, des HUG et de la HES-SO Genève. Gratuit jusqu'à 20 Go, un devis peut être demandé pour l'archivage de volumes plus importants. Les frais de dépôt sont éligibles par le FNS.



B. Zenodo

Zenodo est un *repository* créé en 2013 par OpenAIRE et le CERN. C'est un dépôt de données dit « general-purpose » qui convient à toutes les disciplines scientifiques. Il assure la préservation long terme des données déposées, leur octroie un DOI, offre un choix de licences, des métriques sur l'usage des données, etc. Le dépôt est gratuit. Les fichiers peuvent faire jusqu'à 50GB. Il est considéré comme FAIR par le FNS.



C. Re3data.org (Registry of Research Data Repositories)

Re3data.org a été créé en 2012 et fournit un aperçu des *repositories* existants. Ce registre contient des *repositories* de toutes disciplines académiques.



Les chercheurs peuvent facilement accéder aux conditions d'accès et d'utilisation et autres caractéristiques de chaque entrepôt de données, notamment grâce aux icônes.

A noter que chaque *repository* est examiné avant d'être inclus au registre.

ⁱ Selon: Hole, Brian. *Open Access: Advantages, Funding, Opportunities*. 24 Oct 2012. Available: <https://www.slideshare.net/brianhole/open-access-advantages-funding-opportunities>
Images Icones: <https://icons8.com/icon/pack/free-icons/carbon-copy>

