

Dépôt et entrepôts

Le dépôt des données de recherche répond à une logique de partage et d'ouverture. Il nécessite une gestion rigoureuse des données pour que celles-ci soient accessibles et réutilisables dans le temps.

Comment choisir un entrepôt de données ?

Suivez les recommandations

Dans le cas de disciplines structurées pour le partage des données (astronomie, génomique, etc.), les producteurs de données ont à disposition des **entrepôts spécifiques à leur discipline**.

La pratique de sa communauté est le meilleur guide, mais des **annuaires d'entrepôts** existent. L'annuaire [re3data](#) recense des entrepôts au niveau international et permet de **filtrer selon plusieurs critères** (discipline, attribution d'un identifiant pérenne, types de données ou formats acceptés, préservation à long terme des données, choix de la licence, certification, etc.).

En alternative, les producteurs de données pourront se tourner vers l'**entrepôt institutionnel** auquel ils sont affiliés.

Exemples

En génomique, la communauté scientifique utilise notamment [UniProt](#) ou [GenBank](#) comme base de référence.

En Sciences Humaines et Sociales, [Nakala](#) permet de partager, publier et valoriser tous types de données numériques documentées (fichiers textes, sons, images, vidéos, objets 3D, etc.) dans un entrepôt sécurisé afin de les publier en accord avec les principes FAIR.

Si aucun entrepôt disciplinaire ou institutionnel ne correspond à vos besoins, **il est recommandé de déposer dans l'entrepôt national pluridisciplinaire [Recherche Data Gov](#)**. Si vos données soutiennent une publication scientifique et pour des besoins de preuves, il est important de choisir cet entrepôt qui permet d'accéder de manière pérenne et fiable à la version du jeu de données qui a servi à l'obtention des résultats.

Le saviez-vous ?

Développé à partir de l'application web open source Dataverse, l'entrepôt **Recherche Data Gov** est organisé en espaces institutionnels de publication et de signalement des données des établissements qui souhaitent participer. Il propose également un espace générique dans lequel les déposants qui ne disposent pas d'un espace institutionnel peuvent déposer leurs jeux de données.

Comment déposer ses données ?

Préparez et vérifiez vos données

Pour que le partage des données soit efficace, il faut les préparer selon les principes FAIR, et ce, quelle que soit l'ouverture des données envisagée, totale ou partielle, ouverte ou sécurisée...



Remplissez le formulaire

Quand tout est prêt, connectez-vous sur le site de l'entrepôt, remplissez le formulaire de dépôt et importez vos fichiers.

Conseil : remplissez les champs obligatoires, mais pas uniquement ! Renseigner les champs supplémentaires permettra de faciliter la découverte de vos données, de les rendre plus interopérables et réutilisables.

Exemples

- Anonymiser les données à caractère personnel,
- Choisir une licence de diffusion appropriée pour vos données,
- Préparer les métadonnées et codes sources à associer à vos données,
- etc.

Où et comment déposer dans un entrepôt ?

 Pour choisir un entrepôt dans un annuaire, filtrez par critères

| | | |
|--|---|---|
| Quel type d'entrepôt privilégieriez-vous ? Disciplinaire ? Pluridisciplinaire ? | L'entrepôt est-il certifié, reconnu au niveau international ? | Quels types de données sont acceptés par l'entrepôt ? |
| Quels formats sont acceptés ? | L'entrepôt impose-t-il un standard de métadonnées ? | Impose-t-il une licence ? Vous convient-elle ? |
| Attribue-t-il un identifiant pérenne aux données ? | La plateforme est-elle simple d'utilisation ? Des aides sont-elles proposées ? | Pourrez-vous modifier les données déposées et accéder aux anciennes versions ? |
| L'entrepôt est-il localisé en France ? | L'entrepôt permet-il de restreindre l'accès aux données ? De mettre un embargo ? | L'entrepôt s'engage-t-il à préserver l'accès aux données sur le long terme ? |
| Le dépôt est-il payant ou gratuit ? | L'entrepôt propose-t-il un lien entre les données et les publications associées ? | Des statistiques d'utilisation, de consultation, de téléchargement sont-elles disponibles ? |

Les réponses à ces questions peuvent être trouvées sur les pages de politique des entrepôts ou être filtrées dans des annuaires d'entrepôt

✓ Vérifiez vos données avant de les déposer

- ✓ Les jeux de données à partager ont été sélectionnés
- ✓ Les principes éthiques sont respectés
- ✓ Les droits de diffusion sont vérifiés
- ✓ Les modalités d'accès sont définies
- ✓ Les fichiers sont organisés et nommés de manière explicite,
- ✓ Les fichiers sont dans des formats pérennes et ouverts
- ✓ Le volume des fichiers ne dépasse pas la taille maximale
- ✓ Les données sont décrites et documentées (en suivant, si possible, des standards)
- ✓ Un identifiant unique et pérenne est attribué aux données (et aux codes sources)
- ✓ Une licence est attribuée aux données