

Comment rendre ses données FAIR ?

 opencience.pasteur.fr/2025/05/27/comment-faire-pour-rendre-ses-donnees-fair/

CeRIS - Institut Pasteur

27 mai 2025

[Mise à jour d'un article initialement publié en octobre 2020]

Les principes FAIR (que nous vous [expliquions ici](#)) peuvent être utilisés comme un **guide de bonnes pratiques** pour améliorer la découverte et la réutilisation des données. Pour répondre à une grande partie de ces principes, le plus simple est de déposer ses données dans un [entrepôt de données](#).

Il est donc indispensable de choisir un entrepôt qui permet d'atteindre un niveau suffisant de « FAIRisation » des données :

– Si vos données ne peuvent pas être rendues publiques, il vous faut **choisir un entrepôt adapté pour des données en accès restreint**, permettant d'en contrôler l'accès : vous pourrez ainsi signaler l'existence de vos données en rendant accessibles uniquement les métadonnées et en laissant la possibilité de faire une demande d'accès aux fichiers (exemple : l'entrepôt [European Genome-phenome Archive](#)).

– Il est conseillé de **choisir un entrepôt qui attribue des identifiants uniques et pérennes** (DOI, PURL, ARK, Handle...). Pour le savoir, vous pouvez chercher l'entrepôt dans le répertoire [re3data](#), regarder si la 4^e icône est bleue et consulter l'onglet *Standards* pour connaître le type d'identifiant attribué (exemple : l'entrepôt [ImmPort](#) attribue des DOI).

– Bien que ce soit une information plus difficile à trouver, il est également préconisé de **choisir un entrepôt qui suit un standard de métadonnées** et utilise des [vocabulaires standards](#) dans votre communauté. Pour cela, vous pouvez chercher l'entrepôt dans le répertoire [FAIRsharing](#), et consulter l'onglet *Related Standards* qui indique les standards de métadonnées, ontologies et formats standards mis en œuvre par l'entrepôt (exemple : l'entrepôt [Image Data Resource](#) suit le standard [REMBI](#) et utilise plusieurs ontologies).

Certaines bonnes pratiques liées aux principes FAIR ne dépendent que de vous :

– Pour des données en accès restreint, **bien préciser les conditions d'accès aux fichiers** : qui a le droit d'accéder aux données, comment et sous quel délai. Pour les données en accès ouvert, bien [choisir la licence de diffusion](#) à associer à vos jeux de données.

– **Décrire vos données de façon précise** [avec des métadonnées et de la documentation](#), en donnant le plus d'informations contextuelles possible sur les données, par exemple : la méthode de production des données, la version du logiciel utilisé, le paramétrage des instruments de mesure, le nom des variables...

– **Ajouter des liens vers des ressources complémentaires** (données, documentation, publications...) dans les métadonnées.

– **Privilégier les formats de fichiers ouverts ou largement utilisés**. Les formats ouverts, à l'inverse des formats propriétaires, sont interopérables : une donnée enregistrée dans ce type de format est indépendante du logiciel utilisé pour la créer ; elle pourra être lue et modifiée par tous les logiciels destinés à traiter ce type de fichier.

– Si les données ne peuvent être lues, traitées ou analysées que par un logiciel développé en interne, **mettre à disposition son code source** et le documenter.

Ressource utilisée pour la mise à jour de cet article :

Institut de l'Information Scientifique et Technique, Albaret, L., Bage, A.-S., Caporali, A., Lemeltier, D., MARINI, J., & Moreau, E. (2024). *DoRANum-Enjeux et bénéfiques : Le Minimum à FAIR*. DoRANum. <https://doi.org/10.13143/P4WF-4J63>