

LES PRINCIPES FAIR

RECOMMANDATIONS

- Utiliser des identifiants pérennes uniques comme le DOI, ORCID
- Décrire les données à l'aide de métadonnées riches, en utilisant du vocabulaire contrôlé et en respectant des normes reconnues
- Choisir des formats de fichiers ouverts et non propriétaires
- Définir les conditions d'accès aux données
- Déposer ses données dans un dépôt de données ouvert
- Attribuer une licence de réutilisation

Les principes FAIR définissent des caractéristiques permettant aux données et aux métadonnées d'être **faciles à trouver, accessibles, interopérables et réutilisables**. Des données FAIR ne sont pas forcément synonymes de données ouvertes. Certaines données bien que déposées en accès ouvert peuvent ne pas être conformes aux principes FAIR. A contrario certaines données en accès restreint ou fermé peuvent être en accord avec les principes FAIR.

FINDABLE

Pour rendre les données et les métadonnées faciles à trouver par les humains et les systèmes informatiques (GO FAIR, s. d.; Swiss National Science Foundation, s. d.) il faut :

- Avoir un [identifiant pérenne unique](#) (PID) (ex. DOI, ORCID...)
- Décrire les données à l'aide de métadonnées riches qui respectent des normes reconnues. Elles documentent le contexte de création, les conditions de partage (licence) et de réutilisation afin de permettre une bonne interprétation. Ces informations doivent être lisibles par le système informatique.
- Mentionner clairement l'identifiant pérenne unique dans les métadonnées afin que l'association entre le fichier de métadonnées et l'ensemble des données soit explicite
- Enregistrer les données et métadonnées dans un dépôt qui indexe les métadonnées et qui permette de les rechercher.

ACCESSIBLE

Pour faciliter l'accès, la consultation et le téléchargement des données et des métadonnées par des humains et des systèmes informatiques, il est nécessaire de les stocker durablement.

(Centre pour la communication scientifique directe, s. d.; GO FAIR, s. d.; Swiss National Science Foundation, s. d.). Accessible ne veut pas forcément dire ouvert ou libre mais seulement que les



conditions exactes de l'accès sont clairement mentionnées (GO FAIR, s. d.). Pour se faire il est nécessaire de :

- Choisir un dépôt de données qui utilise un protocole de communication standardisé, gratuit et ouvert afin de faciliter la récupération des données (ex. http, ftp, smtp...)
- Définir les conditions d'accès : licence, droit de réutilisation.
- Rendre les métadonnées accessibles même si les données ne peuvent être ou ne sont plus mises à disposition. (ex. les métadonnées auteurs, institutions, publications associées peuvent être utiles même si les données sont manquantes). Cette solution permet d'assurer la découverte des données et contribue à limiter les coûts de stockage.
- S'assurer que les métadonnées et les données sont des fichiers physiquement séparés afin de faciliter la continuité de la découverte des métadonnées en cas de suppression des données.
- Créer un compte sur le répertoire afin d'être authentifié en tant qu'auteur et laisser la possibilité d'être contacté pour obtenir davantage d'informations ou une autorisation d'utilisation si l'accès est mis sous conditions.

INTEROPERABLE

Pour rendre les données et les métadonnées combinables entre elles, utilisables et interprétables par les différents systèmes informatiques - on parle d'interopérabilité technique et sémantique (Ceris-Institut Pasteur, 2022) – il est faut :

- Choisir des [formats de fichiers ouverts et non-propriétaires](#)
- Choisir des standards de métadonnées ainsi que des formats adaptés
- Utiliser un vocabulaire contrôlé, des ontologies, des thésaurus pour décrire les données et les métadonnées ([ayant des PID](#)) (Deng, s. d.; Paquette-Bigras, s. d.). Consulter la guideline métadonnées pour obtenir davantage d'informations.
- Faites des références à d'autres métadonnées existantes. Décrire le lien ou la relation scientifique entre des ensembles de données en utilisant les PID.

REUSABLE

Pour rendre les données et les métadonnées réutilisables pour de futures recherches, elles doivent être suffisamment décrites, notamment sur la provenance des données et les conditions d'utilisation (GO FAIR, s. d.; SNF, s. d.). Il faut donc veiller à :

- Décrire aussi bien l'ensemble de données (contenu de l'ensemble de données, comment elles sont générées, traitées et réutilisées...) qu'au niveau des données (informations pour utiliser les données, ex. définition des noms de variables...) (SNF, s. d.). Quelques points d'attention proposés par le SNF et GO FAIR :
 - Décrire «la portée de vos données : dans quel but ont-elles été générées/collectées ?
 - Mentionner les particularités ou limitations des données dont les autres utilisateurs devraient avoir connaissance.
 - Préciser la date de génération/collecte de l'ensemble de données, conditions de laboratoire, qui a préparé les données, réglages des paramètres, nom et version du logiciel utilisé





- S'agit-il de données brutes ou traitées ?
 - S'assurer que tous les noms de variables sont expliqués ou auto-explicatifs (c'est-à-dire définis dans le vocabulaire contrôlé du domaine de recherche).
 - La version des données archivées et/ou réutilisées est clairement spécifiée et documentée.»
- Décrire dans les métadonnées la licence de réutilisation des données. Les licences suggérées sont CC BY, CC 0... (SNF, s. d.)

RÉFÉRENCES

Centre pour la communication scientifique directe. (s. d.). *Principes FAIR*. CCSD. <https://www.ccsd.cnrs.fr/principes-fair/>

Ceris-Institut Pasteur. (2022, juin 3). *Comment rendre ses données interopérables ?* Open science : évolutions, enjeux et pratiques. <https://openscience.pasteur.fr/2022/06/03/comment-rendre-ses-donnees-interoperables/>

Deng, S. (s. d.). *UCF research guides: metadata: dataset metadata checklist*. https://guides.ucf.edu/metadata/datasetmetadata_checklist

GO FAIR. (s. d.). *FAIR Principles*. GO FAIR. <https://www.go-fair.org/fair-principles/>

Observatoire global du Saint-Laurent. (2019, octobre 17). *Principes FAIR*. OGSL. <https://ogsl.ca/fr/principes-fair/>

Paquette-Bigras, E. (s. d.). *Documentation des données*. Bibliothèques - Université de Montréal. <https://bib.umontreal.ca/gerer-diffuser/gestion-donnees-recherche?tab=2305>

Swiss National Science Foundation. (s. d.). *Explanation of the FAIR data principles*. https://www.snf.ch/SiteCollectionDocuments/FAIR_principles_translation_SNSF_logo.pdf

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

| Date de mise à jour | Qui | Changements et commentaire | Version |
|---------------------|---|----------------------------|---------|
| 06.06.2022 | Pauline Melly, HES-SO VS | Création de la guideline | V1 |
| 06.06.2022 | Raphaël Grolimund, HEG-GE Jean-Paul Calbimonte, HES-SO VS Laurent Amiotte-Suchet, HESAV | Relecture | V1 |



Les principes FAIR © 2022 by Groupe de travail Guidelines de la Communauté Open Science HES-SO is licensed under [Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

