

Tapez ici votre recherche



Toutes les rubriques



5 idées reçues sur les données de recherche

Imprimer

Partager

La gestion des données de recherche représente aujourd'hui un enjeu scientifique, économique et sociétal majeur. L'absence d'application des principes FAIR (Findable, Accessible, Interoperable, Reusable) entraîne non seulement des pertes financières considérables, mais également une perte de temps, une mauvaise valorisation des résultats scientifiques, et un frein à la reproductibilité.

Idée reçue n°1 : « Je suis propriétaire de mes données de recherche. »

Faux pour la recherche publique. La **Loi pour une République numérique** de 2016 assimile les données de la recherche à des données publiques lorsque les travaux sont financés pour plus de la moitié par des fonds publics.

Les institutions (universités, établissements publics à caractère scientifique et technologique, agences de financement) exercent généralement une responsabilité juridique et éthique sur ces données, qui dépassent l'individu qui les produit.

Les scientifiques sont responsables de leurs données mais rarement propriétaires. Ils doivent garantir leur traçabilité, leur conservation et, autant que possible, leur mise à disposition selon les principes FAIR.

Cette responsabilité s'inscrit désormais dans un contexte géostratégique en profonde mutation (tensions internationales, enjeux de souveraineté et de sécurité nationale). Ainsi, même si les principes FAIR encouragent le partage, certaines données peuvent faire l'objet de restrictions de diffusion, pour des raisons stratégiques.

Idée reçue n°2 : « Je peux gérer mes données moi-même facilement. »

Faux. La gestion des données est une compétence technique à part entière, au même titre que l'analyse statistique par exemple. Elle repose sur des connaissances variées : structuration, formats, métadonnées, documentation, plans de gestion des données, archivage, aspects juridiques et éthiques.

Une gestion individuelle risque d'entraîner des **pertes ou corruptions de données**, des confusions entre fichiers, un manque d'informations rendant les données inutilisables, des incohérences dans les formats.

De plus en plus d'institutions recrutent des data stewards, data managers ou ingénieurs en gestion des données, et proposent des formations. Cela offre un gain de temps précieux et une garantie de qualité.

Idée reçue n°3 : « La gestion des données a un gros impact environnemental. »

Vrai et faux. Le stockage des données consomme de l'énergie, et son impact dépend de la manière dont les données sont organisées et conservées.

De mauvaises pratiques augmentent l'empreinte carbone des données :

- Multiplication des copies redondantes qui nécessite plus d'espace de stockage.
- Absence de tri entre données utiles et données brutes obsolètes.
- Stockage dans des infrastructures non optimisées.

En centralisant, organisant et documentant correctement les jeux de données, il est possible de réduire significativement ces impacts. Par exemple **l'application des principes FAIR peut réduire directement de 20 % le nombre de copies redondantes de données**¹. Un stockage structuré permet de limiter les copies parasites qui apparaîtraient si chaque personne conservait sa propre version.

Idée reçue n°4 : « Avoir les données seules, c'est suffisant ! »

Faux. Sans métadonnées, les données sont inutilisables. Les métadonnées fournissent le contexte indispensable pour **comprendre, interpréter et réutiliser un jeu de données**. Avoir des données sans métadonnées, **c'est comme avoir une boîte de conserve sans étiquette** : on ne peut pas savoir ce qu'elle contient ou comment le contenu a été produit !

Quelques exemples de métadonnées indispensables : méthodes de collecte, protocoles expérimentaux, unité de mesure, contexte temporel et spatial...

Sans ces éléments, même vos propres données peuvent devenir incompréhensibles après quelques mois. Dans une célèbre enquête², réalisée par Monya Baker et publiée dans *Nature* en 2016, plus de 70 % des équipes de recherche déclaraient avoir essayé et échoué à reproduire les expériences d'autres scientifiques et 50 % déclaraient également avoir échoué à reproduire leurs propres expériences.

Idée reçue n°5 : « Mettre en œuvre les principes FAIR coûte trop cher et prend trop de temps... »

Faux. Si l'effort initial de documentation et de structuration semble parfois important, les

études montrent que c'est l'absence de principes FAIR qui génère les coûts les plus élevés.

Un rapport de la Commission européenne¹ a estimé le coût annuel de la non-conformité aux principes FAIR à **au moins 10,2 milliards d'euros**, dont 43,8 % liés au **temps perdu** par les scientifiques et 52,4 % au **stockage de données dupliquées inutilement**. À l'échelle des 302,9 milliards d'euros de dépenses de recherche en 2016, cela représente environ 3 % du budget total de la recherche européenne.

Mettre en place les principes FAIR permet de réduire :

- la perte de temps liée à la recherche d'information ;
- les risques de perte ou d'incompréhension des données ;
- les coûts d'infrastructure ;
- les risques d'erreurs et de retravail ;
- les duplications de données entraînant une facture environnementale et financière élevée.

Adopter une gestion rigoureuse des données n'est pas une contrainte supplémentaire, mais un investissement stratégique : gain de temps, réduction des coûts et meilleure qualité des recherches. Les idées reçues doivent laisser place à une culture de gestion de la donnée !

Pour plus d'informations sur la gestion des données, consultez le guide : <https://www.pepr-agroeconum.fr/ressources-utiles/guide-gestion-des-donnees>

Un grand merci à Frédéric de Lamotte (<https://orcid.org/0000-0003-4234-1172>), Data Steward (INRAE) pour son expertise et son aide à la rédaction de cet article.

¹Rapport de la Commission Européenne « Cost of not having FAIR research data » (2018). <https://op.europa.eu/en/publication-detail/-/publication/d375368c-1a0a-11e9-8d04-01aa75ed71a1/language-en>

²Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* 533, 452–454 (2016). <https://doi.org/10.1038/533452a>

Voir aussi

Pour aller plus loin :

Plus d'informations sur la diversité des métiers autour de la donnée : <https://hal-lara.archives-ouvertes.fr/hal-05265596v1>

Guide Passeport pour la science ouverte sur les données de la recherche : <https://www.ouvrirlascience.fr/wp-content/uploads/2024/03/24-02-22-Donnees-FR-WEB.pdf>

Date de modification : 02 décembre 2025 | Date de création : 02 décembre 2025 | Rédaction : AgroEcoNum