

Bonnes pratiques pour la gestion des données de la recherche



Bibliothèques
VetAgro Sup

En France, la **Loi pour une République numérique** de 2016 rend obligatoire la publication des données issues d'une activité de recherche financée au moins pour moitié par l'Etat. Ces données devront être librement réutilisables, à l'exception des restrictions liées à d'autres réglementations (données personnelles, etc.). Pour les projets sur financement de l'Union européenne, le programme Horizon Europe (2021 - 2027) rend obligatoire l'accès libre aux données de la recherche selon le principe « aussi ouvert que possible, aussi fermé que nécessaire ».

Cet objectif de la Science ouverte correspond à des principes fondamentaux en matière de recherche tels que la reproductibilité des études ou l'intégrité scientifique. En outre, des recherches actuelles montrent que le taux de citation des publications dont les données sont ouvertes est plus important (de l'ordre de 25 à 75 %).

Se former à la gestion de données

L'adoption de bonnes pratiques dans la gestion de leurs données s'impose de plus en plus comme un besoin pour les chercheurs soucieux d'inscrire leur activité dans une démarche de science ouverte. La structuration, le partage et l'ouverture des données de la recherche sont ainsi un des quatre axes du **deuxième Plan national pour la science ouverte**.

- **DoRANum** est une plateforme de formation en ligne sur la gestion et le partage des données de la recherche selon les principes FAIR (Facile à trouver, Accessible, Interopérable et Réutilisable), réalisée par l'**Inist-CNRS** et le **GIS Réseau URFIST**.
- **OSCAR** est une formation e-learning à suivre en auto-formation, à son rythme et à la carte. En cohérence avec sa politique d'ouverture, INRAE y donne accès à tous pour accompagner les changements de pratiques liés à l'ouverture de la science.

Trois modules sont disponibles : **Les fondamentaux de la science ouverte**, **Publications ouvertes**, **Gestion et partage des données**.

Les principales étapes en bref

- Choisir un entrepôt de données adapté
- Vérifier que l'entrepôt dans lequel les données sont déposées leur attribue un identifiant pérenne
- Documenter un maximum les données et utiliser des vocabulaires contrôlés de sa communauté disciplinaire
- Définir clairement les droits d'accès aux données au moment du partage
- Privilégier des formats de fichiers ouverts, non-propriétaires
- Attribuer une licence de réutilisation à ses jeux de données
- Mentionner les productions liées aux données (publications, thèses, autres jeux de données, codes et logiciels).

Le Minimum a FAIR : parcours interactif DoRanum, 1^{er} 2024

Elaborer un plan de gestion de données

DMP OPIDoR est un outil d'aide à la création en ligne de DMP (Data management Plan ou PGD Plan de Gestion de Données) qui vous permet d'anticiper la gestion des données dans le cadre de projets de recherche. Il permet également l'élaboration d'un DMP selon un des modèles mis à votre disposition.

Trouver un entrepôt

VetAgro Sup recommande à ses chercheurs de stocker à minima une version des données issues des projets de recherche sur un serveur informatique sécurisé, celui de l'établissement ou le cas échéant, des établissements partenaires (INRAE, CNRS...).

Vetagro Sup incite ses chercheurs à déposer leurs jeux de données dans des entrepôts certifiés et à but non lucratif, plutôt que dans les entrepôts des éditeurs scientifiques.

Les données doivent en priorité être déposées dans un entrepôt thématique.

Pour trouver un entrepôt thématique adéquat, vous pouvez :



- Contacter **les centres de références** liés à la thématique ou **L'Atelier G.O.T.O.DO UCA** de l'Université Clermont Auvergne*
- Consulter **l'annuaire re3data** qui recense un grand nombre d'entrepôts nationaux et internationaux.

* Les ateliers de la donnée sont des services proposés en proximité géographique des communautés scientifiques pour leur apporter une première assistance dans la gestion et l'ouverture de leurs données de recherche

Le Collège des Données de la recherche du Comité pour la science ouverte propose une **méthode d'identification des entrepôts thématiques de confiance**, ainsi qu'une première liste d'**entrepôts de confiance**. Ce **logigramme dynamique** a pour objectif d'aider le déposant à identifier où il pourra publier son jeu de données.

Si aucun entrepôt thématique n'est identifié, les producteurs de données peuvent déposer, par ordre de priorité :



- Si au moins un des co-auteurs est affilié à un établissement disposant d'un espace institutionnel, dans l'**espace institutionnel** de l'entrepôt pluridisciplinaire **Recherche Data Gouv** (dans le cas de l'existence de plusieurs espaces les co-auteurs s'accordent sur le choix).
- Dans le cas contraire dans l'**espace générique** de l'entrepôt Recherche Data Gouv.

Voir les avantages et inconvénients du **dépôt sur l'entrepôt ZENODO**.

Pour aller plus loin

Partager les données liées aux publications scientifiques : guide pour les chercheurs, 2022,

Ouvrir la science

Glossaire de la science ouverte

Contact : bibliothequesvas@vetagro-sup.fr