

Déposer des données dans un entrepôt, que doit-on anticiper ? L'exemple de GEO

 opencience.pasteur.fr/2025/01/08/deposer-des-donnees-dans-un-entrepot-que-doit-on-anticiper-lexemple-de-geo/

CeRIS - Institut Pasteur

8 janvier 2025

En décembre dernier, la chercheuse [Isabelle Stévant](#) a publié un article intitulé « [Chronique d'une soumission de données à GEO](#) » sur le blog participatif de bioinformatique francophone, [Bioinfo-fr.net](#). Cet article illustre étape par étape le **processus de dépôt de données de séquençage haut débit** dans l'[entrepôt](#) Gene Expression Omnibus ([GEO](#)).

Au-delà de son utilité directe pour les scientifiques qui souhaiteraient soumettre leurs données à GEO, cet article présente également un intérêt pour les chercheurs d'autres disciplines. En effet, il donne un aperçu de **tout ce qui doit être anticipé** avant de se lancer dans un dépôt de données, même si le processus varie d'un entrepôt à l'autre :

- **Prévoir le temps nécessaire** pour effectuer le dépôt, pour une éventuelle vérification des informations par un collègue, pour le transfert des données, et enfin pour la curation/validation du dépôt par les curateurs de l'entrepôt,
- **Identifier les fichiers** de données à déposer. Dans certains cas, l'entrepôt n'accepte des fichiers que dans un format particulier, il sera alors nécessaire de les convertir dans le bon format,
- Vérifier si l'entrepôt a des exigences ou des recommandations concernant l'**organisation** et le **nommage** des fichiers, afin de s'y conformer,
- **Documenter** au quotidien la façon dont les données ont été générées afin de pouvoir remplir un maximum d'informations au moment du dépôt,
- Si possible, identifier le plus tôt possible les **métadonnées** qui seront demandées au moment du dépôt (en téléchargeant les fichiers de métadonnées, en identifiant le standard de métadonnées utilisé...) afin de les compléter le plus tôt possible.

Dans le cas de GEO, l'autrice de l'article décrit le processus de dépôt comme « (relativement) long et pénible », tout en reconnaissant qu'il s'agit d'une étape obligatoire qu'il ne faut pas négliger. Le processus est souvent plus simple pour les autres entrepôts, mais il est indispensable de bien **connaître leurs spécificités** et d'**anticiper** les démarches pour se faciliter la tâche.