

Supprimer des données : pourquoi ? lesquelles ?

Pour accéder à la ressource : <https://doi.org/10.13143/TSFJ-6Q60>

Date de publication : 13/01/2026

1. Enjeux

Ces dernières années, la quantité de données de recherche produites a connu une augmentation exponentielle, ce qui a rendu d'autant plus essentiel le besoin d'en assurer une bonne gestion, notamment au niveau de la question du stockage.

Pour comprendre les enjeux, il est important de bien différencier les notions de stockage, de partage et d'archivage pérenne (pour plus d'informations, vous pouvez consulter [la ressource DoRANum](#) sur ce sujet).

Le stockage sécurisé des données est une étape essentielle durant le projet, il permet de garantir la sécurité des données et d'en assurer l'accès pour l'ensemble des collaborateurs du projet.

Il ne faut pas le confondre avec le partage, qui consiste à déposer ses données dans un entrepôt afin de les rendre accessibles et réutilisables par la communauté scientifique, en respectant les usages et principes de son domaine scientifique.

L'archivage pérenne, quant à lui, intervient après le projet et a pour objectif de conserver les données, d'en garantir l'accès et d'en préserver l'intelligibilité sur le long terme. Attention : toutes les données ne sont pas éligibles à l'archivage pérenne. Celles-ci doivent, par exemple, présenter une valeur scientifique reconnue par la communauté dont elles proviennent.

Ces différentes étapes génèrent non seulement des coûts financiers importants, mais également des coûts écologiques. Bien gérer ses données, tout au long d'un projet de recherche, apparaît essentiel afin de limiter ces coûts, mais également de gagner en efficacité et en temps. En effet, plus il y a de données, plus il devient difficile de trouver les données qui nous intéressent réellement.

Pour cela, il est important de savoir sélectionner les données à conserver et, par conséquent, les données à supprimer.

2. Quelles données supprimer ?

Tout d'abord, il est important de :

- Organiser régulièrement une revue des données afin de décider s'il est toujours nécessaire de les maintenir disponibles. La revue de données peut se faire, par exemple, à travers la mise en place d'un code logiciel qui va vérifier quels jeux de données ne sont pas utilisés depuis un temps défini.
- Distinguer clairement la copie principale de ses dérivés. La copie principale est la version sur laquelle on travaille. Les autres copies sont là en secours, au cas où l'on rencontre des problèmes avec la version de travail ([voir cette ressource sur la stratégie de sauvegarde 3-2-1](#)). Lorsqu'on voudra supprimer des données pour libérer de l'espace, il ne faudra pas oublier d'aller supprimer les différentes copies.

Il n'existe pas de consensus quant au choix des données qu'il est nécessaire de conserver à long terme. Les pratiques diffèrent d'une discipline à l'autre mais également d'un établissement à l'autre, selon la politique de conservation de données mise en œuvre. Toutefois, il faut noter que certains établissements n'ont pas de politique établie de conservation des données. De plus, dans certains domaines comme la Génomique, les entrepôts internationaux comme l'[ENA](#) (European Nucleotide Archive) ou le [SRA](#) (Sequence Read Archive) assurent à la fois le partage et l'archivage des données.

Plusieurs questions permettent tout de même d'être guidé dans le choix de la conservation ou de la suppression des données :

- Quelles sont vos obligations quant à la durée de rétention des données notamment vis-à-vis de votre institution ou des bailleurs de fonds ? À l'inverse, quelles sont vos obligations à NE PAS préserver des données ? (par exemple : en principe, les données personnelles font l'objet d'une obligation de durée de conservation limitée, comme l'explique [cet article](#) de la CNIL).
- Quelles sont les raisons pour conserver ces données ? (Permettre la réutilisation des données, permettre la vérification des données et des découvertes s'appuyant sur elles, conservation pour des questions d'héritage, coût de reproduction, caractère unique et non réplicable, valeur scientifique, historique, culturelle...)
- Quels coûts sont liés au stockage de ces données ? Comment sont-ils financés ?
- S'il est nécessaire d'archiver les données, dans quelles conditions allez-vous le faire et quelle documentation allez-vous produire pour permettre la conservation de ces données dans le long terme ? Il est important de noter que stocker des données mal documentées voire non documentées s'avère inutile : celles-ci ne seront pas réutilisables.
- Quand vous partirez (par exemple : vous partez à la retraite ou vous quittez l'équipe ou encore le projet), que se passera-t-il concernant les données que vous avez laissées ?

Actions concrètes à mettre en place

- Rédigez un plan de gestion de données (PGD) pour anticiper ces questions et mieux vous organiser.
- Organisez vos fichiers en distinguant les données qui auront vocation à être supprimées, partagées ou pérennisées et indiquez bien les numéros de version.
- Documentez vos données en mentionnant sous forme de métadonnées toutes les informations en lien avec le partage des données durant le projet (propriétaire, licence, version...).

- Tournez-vous vers des infrastructures de stockage dédiées, car elles ont une politique de sauvegarde professionnelle et cohérente (stratégie de sauvegarde 3-2-1, gestion des droits d'accès, sécurité...).

En conclusion...

La suppression est une étape à ne pas négliger pour une bonne gestion des données de recherche. Même si les pratiques diffèrent en fonction des disciplines et des politiques mises en œuvre, vous disposez désormais de quelques clés pour mieux anticiper et gérer cette étape.

Webographie

DATAcc (2022). *Que garder, que supprimer ?* <https://www.datacc.org/vos-besoins/conserver-ses-donnees/que-garder-que-supprimer/>

DCC (2014). *Five steps to decide what data to keep: a checklist for appraising research data*. Edinburgh: Digital Curation Centre.

<https://www.dcc.ac.uk/sites/default/files/documents/publications/Five%20Steps%20to%20decide%20what%20data%20to%20keep.pdf>