

# Comment organiser et nommer ses données ?

---

 [opencscience.pasteur.fr/2024/01/18/comment-organiser-et-nommer-ses-donnees/](https://opencscience.pasteur.fr/2024/01/18/comment-organiser-et-nommer-ses-donnees/)

CeRIS - Institut Pasteur

18 janvier 2024

*[Mise à jour d'un article initialement publié en avril 2021]*

Vous est-il déjà arrivé d'ouvrir un dossier partagé avec vos collègues sans parvenir à mettre la main sur les données que vous cherchiez ? Avez-vous déjà été confronté à cette même problématique en ouvrant un vieux disque dur contenant vos propres données ? « *Les données sont là, j'en suis sûr(e) ! Mais alors dans quel dossier et sous quel nom... ?* »

En adoptant des **bonnes pratiques d'organisation et de nommage** de vos fichiers de données, il devrait être plus facile de retrouver les données et la documentation pertinentes pour comprendre exactement ce qui a été fait et pouvoir réutiliser les données.

## Mettre en place un plan de classement

Un plan de classement est un outil conçu pour faciliter le classement et le repérage des données et documents. Il permet d'organiser les dossiers et fichiers en affectant chaque donnée, dès sa création, à un dossier précisément identifié.

Dans vos espaces de stockage collaboratifs, il est donc conseillé de **créer une arborescence hiérarchisée** qui reflète votre activité, en allant du général au particulier. Évitez si possible de créer des dossiers portant le nom des collaborateurs. Une telle organisation peut compliquer le repérage des données car l'information sur « qui a produit quelles données » peut se perdre au fil du temps. Préférez une organisation par projet comme dans l'exemple ci-dessous.

*Exemple de plan de classement issu de la [formation IFB Science Ouverte](#), sous licence [CC-BY-SA](#)*

## Adopter des règles de nommage précises et communes à l'entité ou au projet

Les règles de nommage ont deux objectifs principaux :

- Pour **garantir la pérennité des données**, il est conseillé d'utiliser des noms succincts et ne contenant pas d'espace ou de caractères spéciaux. En effet, un nom trop long ou contenant des caractères interdits peut compromettre l'accès ultérieur au fichier en cas de copie/transfert.

- Pour **permettre la compréhension des informations** contenues dans les fichiers, il est conseillé d'utiliser des noms de fichiers précis et explicites. Des règles communes peuvent ainsi être définies avec les collaborateurs. Par exemple, utiliser systématiquement des noms de fichier comprenant la date, le sujet, le type de données, la version...

### **Associer de la documentation ou des métadonnées pour décrire les données**

Il est recommandé d'associer aux données un fichier README décrivant les données, de façon à ce qu'elles soient compréhensibles et réutilisables dans l'avenir. Il est préférable de décrire les données au moment où elles sont créées, car c'est le moment où l'on connaît le mieux les données. L'entrepôt Recherche Data Gouv propose par exemple un modèle de README que vous pouvez télécharger et adapter.

### **Besoin d'un exemple concret ?**

Dans un tweet publié en mars 2021, Kira Höffler propose une infographie très utile (également disponible via ce lien) avec des exemples pour l'organisation des dossiers et le nommage des fichiers :

Every scientist should be interested in organising their data well. Here, I summarised a few tips [↓ #AcademicChatter @OpenAcademics pic.twitter.com/VBw7sblqU](#)

— Kira Höffler (@KiraHoeffler) March 5, 2021