

Structure des dossiers numériques de recherche au Centre de recherche Azrieli du CHU Sainte-Justine

Objectifs de ce guide

Une bonne organisation des dossiers de recherche simplifie la conduite de vos projets et réduit le risque de perte de données. Le présent guide aidera la communauté du Centre de recherche Azrieli du CHU Sainte-Justine (CRACHUSJ) à mieux collaborer, à garder des traces des travaux de recherche effectués et à intégrer les bonnes pratiques en gestion des données de recherche.

Ce guide est à son plus utile lorsqu'il est utilisé en tandem avec le [guide de nommage des fichiers numériques](#).

Pour toute question sur le contenu de ce guide ou sur la gestion des données de recherche, contactez l'équipe de soutien GDR à soutien.gdr.hsj@ssss.gouv.qc.ca.

Comment organiser vos dossiers numériques

Il est préférable de réfléchir à la structure de vos dossiers numériques au tout début de votre projet. Cette structure sera informée par la façon de faire de votre équipe, ainsi que par les pratiques de vos collaborateurs. Considérez les quatre éléments suivants pour mieux organiser vos dossiers numériques :

1. **Arborescence**
2. **Nommage/versionnage des dossiers**
3. **LISEZMOI (README)**
4. **Sensibilité des données et des documents**

1. Arborescence

L'arborescence de dossiers fait référence à la manière dont vos répertoires de dossiers sont organisés, comme les branches d'un arbre. Une bonne hiérarchisation des dossiers et une structure bien définie créent des parcours d'utilisateur clairs et efficaces. Si les membres de votre équipe perdent du temps à parcourir vos répertoires pour trouver un fichier particulier, pensez à réorganiser vos dossiers.

Il existe diverses façons d'organiser vos projets, mais la règle générale est de placer le dossier principal à la racine de votre répertoire. À partir de la racine, les autres dossiers, sous-dossiers et sous-sous-dossiers se déploient. Pensez à la profondeur et à l'étendue des dossiers, en essayant de limiter le nombre de niveaux (ex. sous-sous-sous-sous-sous-dossier). Une trop grande profondeur peut causer un répertoire complexe et difficile à naviguer.

Organiser les dossiers contenant des données

Selon votre préférence, l'organisation des dossiers contenant vos données pourrait se faire de deux façons : en créant des arborescences selon le type de fichier (A) ou selon le stade de transformation des données (B). Comme vos données prendront une forme hétérogène en raison des différents formats utilisés (images,

questionnaires, protocoles, etc.), il est important de distinguer ces différents types et de les organiser en conséquence.

A) Organisé par type de fichier/format

Cette option est préférable si votre équipe produit des données hétérogènes aux types et formats variés (ex. Images, vidéos, questionnaires, etc.).

Voici un exemple d'arborescence basée sur les types de fichiers :

```
Projet_ABC/  
| LISEZMOI.txt  
|  
├─ Images/  
|   └─ LISEZMOI.txt  
|   └─ Brutes/  
|   └─ Redimensionnees/  
|   └─ Annotees/  
|  
├─ Questionnaires/  
|   └─ LISEZMOI.txt  
|   └─ Questionnaires_vierges/  
|   └─ Resultats/  
|  
├─ Audio/  
|   └─ LISEZMOI.txt  
|   └─ Guides_entrevues/  
|   └─ Brutes/  
|   └─ Transcriptions/  
|  
└─ Documentation/  
    └─ LISEZMOI.txt  
    └─ Dictionnaire_de_donnees.xlsx  
    └─ Protocoles/
```

B) Organisé par version/stade de transformation des données

Certaines équipes préfèrent organiser leurs données par version, en distinguant les données brutes, propres, analysées et finales. L'avantage de cette organisation est qu'elle permet de créer des chemins d'accès simples qui priorisent le niveau d'analyse effectué. Dans les cas où les données brutes contiennent des renseignements personnels de participants à la recherche, cette option facilite la gestion des accès à ces renseignements confidentiels et permet un usage approprié des données.

Voici un exemple d'arborescence basée sur le versionnage des données :

```
Projet_XYZ/  
|  
├── 00_Donnees_brutes/  
|   ├── Sondages/  
|   ├── Export_DSP/  
|   └── Logs_capteurs/  
|  
├── 01_Donnees_nettoyees/  
|   ├── Fichiers_csv/  
|   └── Scripts_nettoyage/  
|  
├── 02_Donnees_analyses/  
|   ├── Statistiques_intermediaires/  
|   ├── Graphiques_preliminaires/  
|   └── Scripts_analyses/  
|  
├── 03_Donnees_finales/  
|   ├── Tableaux_publication/  
|   └── Graphiques_final/  
|  
└── Documentation/  
    ├── Dictionnaire_donnees.xlsx  
    ├── LISEZMOI.txt  
    └── Protocoles/
```

Pour plus d'exemples, consultez le guide de [Dryad sur les bonnes pratiques d'organisation](#).

2. Nommage et versionnage des dossiers

Lorsqu'une équipe de recherche utilise des conventions de nommage cohérentes, tant pour les fichiers que pour les dossiers, la structure de votre répertoire devient prévisible et facile à naviguer. Mettez sur la cohérence et la standardisation des pratiques au sein de votre équipe. Le nommage des dossiers suit les mêmes principes que le nommage des fichiers (voir [guide de nommage des fichiers numériques](#)).

Les noms de dossiers clairs permettent de comprendre le contenu d'un dossier d'un coup d'œil, sans avoir à l'ouvrir. Ces noms devraient être lisibles par un humain comme par une machine. Autrement dit, tout système informatique doit être compatible avec les caractères utilisés, facilement identifiables et ne pas inclure de caractères spéciaux. Pour délimiter les éléments du nom de votre dossier ou fichier, privilégiez l'usage du tiret bas (_). Pour séparer les mots au sein d'un élément d'un nom de fichier, vous pouvez utiliser le tiret (-) ou la majuscule.

Un autre point à considérer est la longueur du nom du dossier : il est préférable qu'il soit court et clair, avec trois mots-clés. Vous pouvez consulter [le guide de nommage des fichiers numériques](#) pour observer le format que le service de soutien en GDR du CRACHUSJ recommande.

Pensez également au versionnage de vos fichiers et aux méthodes d'organisation des différentes versions de vos fichiers. Le versionnage permet de conserver l'historique des modifications apportées aux fichiers et de savoir rapidement quelle version est la plus récente. C'est un élément qui peut être inclus dans le nom de vos dossiers et de vos fichiers de plusieurs façons :

- Dater les dossiers et fichiers lorsque pertinent (ex. AAAA-MM-JJ);
- Conserver les anciennes versions de fichiers plutôt que de les écraser (*overwrite* en anglais);
- Numéroter les fichiers de manière claire et uniforme pour distinguer les versions :

À éviter	À implémenter
Les noms ambigus. rapport_final.docx rapport_final2.docx rapport_final_vraiment_final.docx rapport_nouveau.docx	Les zéros initiaux dans les numéros de version permettent un classement automatique des fichiers. rapport_v01.docx rapport_v02.docx rapport_v03.docx

Il est conseillé de considérer deux éléments principaux pour le versionnage des dossiers : l'organisation des dossiers selon les jalons importants du projet, et l'archivage des versions antérieures de vos fichiers. Cela vous permettra d'éviter la multiplication de copies presque identiques. Par exemple, votre arborescence pourrait suivre le format suivant :

```
projet/  
├─ 01_donnees-brutes/  
├─ 02_donnees-traitees/  
├─ 03_scripts/  
├─ 04_analyses/  
├─ 05_figures/  
├─ 06_donnees-finales/  
├─ 07_publications/  
├─ 08_archives/  
└─ 09_documentation/
```

3. LISEZMOI (README)

Un autre élément à prendre en compte lors de l'organisation de vos fichiers de projet de recherche est la création de fichiers LISEZMOI. Ces fichiers sont généralement en format .txt et servent à guider les usagers dans votre répertoire de fichiers. Ils aident vos collègues à comprendre la structure, le contenu et la provenance des fichiers déposés dans un dossier. Idéalement, chaque niveau de votre répertoire devrait contenir un fichier LISEZMOI décrivant son contenu.

Pour les données, il est recommandé qu'elles soient accompagnées d'un LISEZMOI décrivant leur stade et les instruments utilisés. Vos données bénéficieront également d'être accompagnées d'un plan de gestion des données décrivant la façon dont elles ont été collectées et la manière dont vous planifiez les gérer. Si vos données ou vos logiciels comportent une licence, un fichier LISEZMOI est l'outil parfait pour recenser ces informations. Le CRACHUSJ met à votre disposition un [guide sur les fichiers LISEZMOI](#).

D'autres documents peuvent être utiles à votre équipe, tels qu'un sommaire des travaux réalisés jusqu'à présent, un glossaire, un dictionnaire des variables ou une liste de liens vers les autres données pertinentes à votre projet. Votre documentation devra également comprendre tout autre détail pouvant faciliter la compréhension et le partage de votre travail, surtout si vous travaillez avec des collaborateurs externes qui ne connaissent pas nécessairement les procédures du CRACHUSJ. Toute cette documentation supplémentaire qui ne tient pas dans un fichier LISEZMOI peut être déposée dans un dossier "documentation".

4. Sensibilité des données et des documents

La structure des dossiers d'un projet de recherche devrait faciliter non seulement le repérage des fichiers, mais aussi la gestion des accès. En regroupant les documents selon leur niveau de sensibilité et les besoins d'accès des membres de l'équipe, il devient plus simple de protéger les renseignements confidentiels et d'attribuer les permissions appropriées. Voici quelques pratiques recommandées :

- Respectez le principe du moindre privilège : Chaque membre de l'équipe devrait avoir accès uniquement aux informations dont il a besoin pour accomplir ses tâches. Il n'est généralement pas nécessaire que tous les membres aient accès à l'ensemble des données du projet.
- Regroupez les fichiers selon leur fonction et leur niveau de sensibilité. Par exemple :
 - Les protocoles, guides et documents administratifs sont accessibles à toute l'équipe.
 - Les données contenant des renseignements permettant d'identifier des participants ne doivent être accessibles qu'aux personnes autorisées.
 - Les collaborateurs externes ne doivent avoir accès qu'aux données dépersonnalisées ou aux livrables finaux, dépendamment des autorisations obtenues auprès du Comité d'éthique de la recherche (CÉR) et le Bureau des ententes de recherche (BER).
 - Les étudiants et assistants de recherche ont accès aux données d'analyse, mais pas aux informations nominatives.
- Attribuez les permissions au niveau des dossiers plutôt qu'au niveau des fichiers individuels : Cela vous évitera d'avoir à gérer des permissions différentes dans un même dossier.
- Réduisez le plus possible les exceptions et les permissions particulières lorsque vous accordez les accès dans votre répertoire.
- Séparez les données sensibles des données dépersonnalisées le plus tôt possible.

- Concevez la structure des dossiers en fonction des groupes d'utilisateurs : Avant de créer les dossiers, identifiez les différents groupes qui pourraient avoir besoin d'y accéder. Idéalement, la structure devrait permettre d'accorder les accès à des groupes de personnes plutôt qu'à des individus, ce qui simplifie l'administration au fil du temps. Bien que la composition des équipes de recherche change fréquemment, tentez de prévoir l'évolution de votre équipe afin de faciliter l'intégration de nouveaux membres et le retrait des accès lorsqu'une personne quitte le projet.
- Désignez une personne responsable d'accorder et de retirer les accès des autres membres de l'équipe.

Références

Brigham, D., & Barsky, E. (14 janvier 2026). *PUBLIC UBC RDM documents*. The Open Science Framework.
<https://osf.io/nht4x/files/5br6a>.

Dryad | Good data practices. https://datadryad.org/help/guides/best_practices#organize-files-in-a-logical-schema.

Goldman, J. (2020). *Directory structure*. Data Management Services, Harvard.
<https://datamanagement.hms.harvard.edu/plan-design/directory-structure>.

Pour toute question sur le contenu de ce guide ou sur la gestion des données de recherche, contactez l'équipe de Soutien GDR à soutien.gdr.hsj@ssss.gouv.qc.ca.