

L'archivage de données

Version juillet 2023

1) Introduction

En débutant vos études graduées dans le laboratoire, vous prenez l'engagement d'adhérer à notre système d'archivage électronique de données¹. Puisque nos différents programmes de recherche impliquent des suivis à long terme, un système d'archivage de données bien structuré est nécessaire pour s'assurer que l'information demeure accessible pour tous les membres de l'équipe. Ceci nous permet, entre autres, d'utiliser des bases de données issues de projets antérieurs si celles-ci sont pertinentes à un projet en cours. De plus, le système d'archivage constitue un excellent moyen de protéger vos données contre toutes pertes potentielles.

2) Comment fonctionne le système d'archivage de données?

2.1 Archivage temporaire des données issues des projets en cours

2.1.1 Emplacement par programmes de recherche

Pour tous les étudiants, l'archivage temporaire se fait dans le dossier « Archives » sur les sites sharepoint de chaque programme de recherche, accessibles avec votre compte Office365² de l'Université Laval via les hyperliens suivant :

- [Caribou Ungava](#)
- [Partenariat de recherche Tique-Original-Climat](#)
- [Écologie et comportement de la chèvre de montagne en Alberta](#)
- [Laboratoire Jean-Pierre Tremblay](#) (pour les étudiants de Jean-Pierre Tremblay qui ne sont pas impliqués dans les autres projets de recherche)
- [Aménagement intégré cerfs-forêt à l'île d'Anticosti](#)

Dans ces espaces de stockage virtuel, chaque étudiant possède un dossier personnel à son nom. Seul l'étudiant concerné, les chercheurs responsables du programme de recherche et les professionnels de recherche du laboratoire ont accès à ce dossier. Les données sauvegardées dans ce dossier peuvent être modifiées à tout moment. Seules les données liées à votre projet de recherche doivent y être stockées (voir section 5 pour la liste des données à archiver).

¹ À noter que pour raccourcir le présent document, le terme « données » est utilisé au sens large et inclut tous les documents jugés nécessaires à la réplique des projets et à la conservation des résultats (ex. protocoles, bases de données, scripts, résultats et documents de rédaction finaux).

² L'adresse courriel à saisir pour ouvrir votre compte est IDUL@ulaval.ca et le mot de passe est celui de votre messagerie Uval.

Les sites sharepoint hébergent également des bases de données pouvant être pertinentes pour plusieurs projets. En général, ces bases de données peuvent être consultées par tous, mais peuvent seulement être modifiées par les professionnels ou les chercheurs. Des accès pourraient être accordés pour des besoins spécifiques, mais dans tous les cas vous devrez télécharger le document sur votre ordinateur pour vous en faire une copie personnelle si vous souhaitez l'utiliser.

2.1.3 Espace de stockage alloué à chacun

Puisque l'espace de sauvegarde de 1To alloué à chaque site sharepoint ou sur le serveur de l'université est partagé par tous les membres du programme de recherche, l'espace disponible pour chacun est limité. Les besoins de chacun peuvent varier selon la nature des projets, mais il faut en général se limiter à **un maximum de 150Go par personne**. Si vous avez besoin de plus d'espace, simplement aviser un des professionnels de recherche et nous trouverons d'autres options mieux adaptées.

En plus de ces dossiers sur sharepoint, chaque étudiant dispose d'un espace de stockage personnel de 1To sur l'application [OneDrive](#) de Microsoft 365 de l'université. Seul l'étudiant concerné y a accès, donc cet espace ne peut pas servir d'alternative pour l'archivage des données visé par ce document. Cet espace peut néanmoins être utilisé pour faire des sauvegardes plus fréquentes de vos données temporaires ou sauvegarder des données d'autres projets (par exemple, des données liées à vos cours) ou plus personnelles.

2.2 Archivage permanent des données issues des projets terminés

L'archivage permanent des données finales se fait une fois que les projets sont complètement terminés. Nous utilisons alors des disques durs « miroir » permettant une double copie de l'information sauvegardée. Ces disques durs sont accessibles seulement par les professionnels de recherche qui sont les seuls gestionnaires de leur contenu. Nous procéderons à la migration de vos archives sur les disques durs seulement une fois votre projet terminé et l'archivage de toutes vos données complété sur le sharepoint ou le serveur. Merci de nous tenir au courant lorsque votre archivage est complété sur le sharepoint et que vos dossiers peuvent être transférer.

3) À quelle fréquence archiver les données?

L'archivage périodique des données demeure à la discrétion des étudiants, mais nous vous suggérons de le faire une fois par mois environ. Il est par contre important de ne pas dépasser un délai de 4 mois (période qui correspond aussi au délai du dépôt des rapports d'étape à Steeve et Jean-Pierre) pour éviter des pertes trop grandes et assurer un suivi adéquat.

4) Quelles données doivent être archivées et comment?

Les données doivent être archivées **par chapitre**, mais faites également un dossier pour les aspects généraux (ex. proposé ou mémoire/thèse) ou qui touchent à plusieurs chapitres. La

structure présentée ci-dessous est celle qui sera exigée à la fin de votre projet. Il n'est donc pas obligatoire de suivre toutes les directives pour un archivage temporaire, mais cela est tout de même fortement suggéré, car cela vous évitera de modifier la structure de vos archives à la fin de votre projet. Voici donc la liste des données/documents à archiver :

4.1 Échantillonnage et données brutes

- Protocoles: Tout protocole d'échantillonnage et/ou de manipulations en laboratoire utilisé.
- Données brutes: Copie des données brutes telles que récoltées sur le terrain ou obtenues en laboratoire. Il peut s'agir de fichier excel, de versions numérisées de feuilles de terrain ou de résultats de laboratoire.
- Métadonnées : Les métadonnées servent à décrire les données de vos fichiers. Chaque fichier pour lequel l'information ci-dessous n'est pas incluse doit contenir un onglet de métadonnées (dans Excel ou Access) ou un fichier associé (.txt, .docx). Les métadonnées doivent contenir :
 - La signification des colonnes (variable, unité de mesure)
 - La définition de toutes les abréviations utilisées
 - La provenance des données, s'il y a lieu
 - La période couverte par vos données

4.2 Bases de données Excel / Access / .csv / .txt ...

- Données de base NETTOYÉES : Le ou les fichier(s) de base compilé(s) avec l'ensemble des données brutes, sans transformation, mais nettoyé(s). Il s'agit généralement de la compilation de l'ensemble des données récoltées avant toutes transformations mathématiques, mais après avoir retiré les valeurs erronées (ex. doublons ou donnée aberrante en raison d'erreur de mesure). Selon les projets, il est possible que cette base soit la même que celle mentionnée en 4.1
- Données TRANSFORMÉES : Le ou les fichier(s) obtenu(s) une fois la transformation de données terminée. Il s'agit des bases de données utilisées pour faire les analyses.
- Métadonnées : Voir section 4.1

4.3 Scripts R / SAS

- Scripts NETTOYÉS : Scripts finaux permettant de réaliser les analyses et d'obtenir les résultats finaux. Les essais et erreurs survenus lors de la création de chacun des scripts doivent donc être supprimés.
- Métadonnées : Une personne consultant les scripts devrait être en mesure de suivre la logique de votre démarche. Au tout début du script, insérez une courte description de l'objectif du script et le lien qu'il a avec vos résultats. Les symboles « # » dans R et « * » suivi d'un « ; » à la fin de l'énoncé dans SAS devraient être utilisés à même les scripts afin d'inclure le plus d'informations possible. D'autres options comme l'utilisation de Rmarkdown sont également possibles pourvu que la description du script permette de bien comprendre l'objectif et les étapes de celui-ci.

4.4 Fichiers GIS

- Fichiers géospatiaux (shapefiles, raster ou autres): Tous les fichiers essentiels pour obtenir vos résultats ou pour recréer les cartes et figures finales issues de votre projet. Assurez-vous de faire suivre toutes les extensions nécessaires.
- Métadonnées : Un fichier associé (.txt, .docx) portant le même nom que votre fichier géospatial suivi de « _meta » devrait être créé pour expliquer la nature de votre fichier. Assurez-vous d'indiquer les données sources, le système de référence géographique et la projection utilisés.

4.5 Figures et tableaux finaux

- Figures et tableaux finaux inclus dans votre mémoire, thèse ou articles. Fournir les figures en format « image » (ex : .jpg, .tif, .png ...) et ajouter un fichier « métadonnée » qui stipule le programme utilisé et la provenance des données illustrées. Assurez-vous d'archiver les bases de données ayant servi à créer ces figures et tableaux (voir section 4.2).

4.6 Documents écrits

- Fichier word (ou autre format similaire) : Versions soumises de vos principaux documents écrits : les versions initiales et révisées de votre proposition de projet et de votre mémoire ou thèse, ainsi que les versions soumises des articles acceptés pour publication.
- Fichier pdf : La version finale de votre mémoire ou thèse et la version publiée de vos articles.

5) Comment nommer les fichiers?

Assurez-vous de nommer vos fichiers de manière à ce que le contenu soit facilement identifiable avec deux informations essentielles :

- Un nom court représentant les données (ex : points_caribous)
- La date de la dernière mise à jour du fichier : aaaammjj (ex : 20150325)

Ex : points_caribous-20150325

Bon archivage!