**Plan de gestion des données (PGD)**

**Modèle du CHUSJ**

**Aussi disponible sur l’**[**Assistant PGD**](https://dmp-pgd.ca/)**.**

Bienvenue dans le modèle de PGD (plan de gestion des données de recherche) du CHU Sainte-Justine.

L’équipe de soutien en gestion des données de recherche (GDR) a conçu ce modèle de plan de gestion des données de recherche (PGD) pour faciliter le travail des équipes de recherche. L’utilisation de la version Word du modèle institutionnel ou de l’[**Assistant PGD**](https://dmp-pgd.ca/) démontre l’importance que nos équipes accordent à la GDR. Votre proactivité dans l’usage de ces outils est appréciée.

Prochainement, les [**organismes subventionnaires**](https://science.gc.ca/site/science/fr/financement-interorganismes-recherche/politiques-lignes-directrices/gestion-donnees-recherche) vont exiger que chaque demande de subvention soit accompagnée d’un PGD. Nous vous invitons à vous familiariser à cet outil afin de vous préparer au déploiement général de cette exigence. Le PGD est un **document évolutif** qui vous suivra au cours de votre étude. Continuez de le **mettre à jour** au fur et à mesure de votre projet.

Pour plus d’information et de ressources liées à la gestion des données de recherche, consultez le [**site web des bibliothèques du CHU Sainte-Justine**](https://enseignement.chusj.org/fr/bibliotheques/les-Ressources/Gestion-de-donnees-de-recherche).

Vous avez des questions? Besoin de soutien? Contactez-nous par courriel à l’adresse suivante : [**soutien.gdr.hsj@ssss.gouv.qc.ca**](mailto:soutien.gdr.hsj@ssss.gouv.qc.ca).

**N’oubliez pas de supprimer les directives, les ressources et les exemples une fois après avoir complété votre plan!**

Table des matières

[Détails du projet 1](#_Toc83637238)

[1. Collecte de données 2](#_Toc446907041)

[2. Documentation et métadonnées 5](#_Toc524779690)

[3. Stockage et sauvegarde 8](#_Toc1105325981)

[4. Préservation 10](#_Toc540974774)

[5. Partage et réutilisation 12](#_Toc1829178589)

[6. Responsabilités et ressources 16](#_Toc477263293)

[7. Conformité éthique et juridique 18](#_Toc1312299770)

# Détails du projet

**Titre du projet :**

**Résumé du projet :**

**Date de début et de fin du projet :**

**Contributeurs :**

# Collecte de données

Cette section traite des questions relatives à la collecte de données comme les types de données, les formats de fichier, les conventions d’appellation et l’organisation des données – des facteurs qui amélioreront la facilité d’utilisation de vos données et contribueront au succès de votre projet.

1.1 Décrivez les types de données, les sources de données potentielles que vous acquérez dans le cadre de cette étude, ainsi que les formats dans lesquels vos données seront collectées. Les formats que vous utiliserez permettent-ils une réutilisation future de vos données? Faudra-t-il un logiciel tiers pour y avoir accès?

Directives

Soyez aussi précis que possible quant aux types de données que vous collecterez (données numériques, images, audio, vidéo, texte, données tabulaires, données de modélisation, données géospatiales, données d’instrumentation, données cliniques, données administratives, dossiers de patients électroniques, données d’imagerie, données d’enquête, etc.) ainsi que les formats des fichiers/données.

Pour chaque type de données acquises ou recueillies, soyez aussi précis que possible en mentionnant les formats associés.

*Les formats propriétaires nécessitant l’utilisation de logiciels ou de matériels spécialisés ne sont pas recommandés, mais peuvent être nécessaires pour certaines méthodes d’analyse ou de collecte des données. Envisagez de transformer tout fichier propriétaire* **en** [**formats convenables à la préservation**](https://www.canada.ca/fr/reseau-information-patrimoine/services/preservation-numerique/recommandations-formats-fichier-preservation-numerique.html) *pour en faciliter l’accès et la réutilisation.*

*Si vous planifiez convertir vos données en format ouvert une fois votre projet terminé,* ***mentionnez-le et mentionnez la phase du projet où se produira la migration****. Les formats de collecte et les formats de préservation peuvent être différents (ex. données collectées dans Microsoft Excel mais stockées en format .csv).*

Ressources

[***Recommandations relatives aux formats de fichier pour la préservation numérique***](https://www.canada.ca/fr/reseau-information-patrimoine/services/preservation-numerique/recommandations-formats-fichier-preservation-numerique.html)*, guide du Ministère du Patrimoine canadien*

[***Guide concernant les formats recommandés***](https://numerique.banq.qc.ca/patrimoine/details/52327/4076856)*, guide de BAnQ*

*Pour plus d’information sur les formats ouverts et fermés, voir* [***DORANum***](https://doranum.fr/wp-content/uploads/FS2_liste_indicative_formats_V1.pdf)*.*

*Pour en apprendre plus sur les formats qui facilitent la réutilisation, le partage et l’accès à long-terme des données, voir* [***UK Data Service***](https://ukdataservice.ac.uk/learning-hub/research-data-management/)*.*

Exemple

*Les données recueillies au cours de cette étude comprennent :*

* *Enregistrement d’entrevues cliniques (vidéo, .mp4)*
* *Transcription d’entrevues cliniques (texte, .rtf)*
* *Résultats de sondages*
* *Des images en neuroimagerie (en .dcm)*
* *Des données cliniques (en .csv)*
* *Les signes vitaux*
* *La température du corps*
* *Groupe sanguin du patient*
* *La date d’admission du patient*

1.2 Comment prévoyez-vous acquérir les données utilisées dans le cadre de cette étude?

Données originales créées dans le cadre de l’étude

Données existantes provenant d’un ancien projet de recherche

Données provenant de partenaires externes (ex. ISQ, MSSS, etc.)

Données existantes provenant d’un dépôt en ligne (ex. Borealis)

Autre :

Directives

*Pour chaque méthode d’acquisition sélectionnée ci-haut, veuillez décrire les sources de données auxquelles vous aurez recours.*

*S’il s’agit de données existantes, décrivez leur provenance (nom, lieu, URL) ainsi que la manière dont vous les obtiendrez et les intégrerez à votre projet de recherche. Décrivez toute licence d’utilisateur final attribuée aux données existantes, ou les conditions d’utilisation que vous devez respecter.*

*Si vous comptez combiner des données de recherche originales avec des données existantes ou déjà utilisées, décrivez ici le processus d’appariement.*

*Ressources*

[***Données disponibles à l’Institut de la statistique du Québec (ISQ)***](https://statistique.quebec.ca/fr/institut/services-recherche/donnees)

[***Réutiliser des données, pourquoi et comment?***](https://bibliotheque.uqac.ca/c.php?g=723711&p=5180030)*, un guide de l’Université du Québec à Chicoutimi*

*Exemple*

*Données originales créées dans le cadre de l’étude*

*Les réponses aux sondages seront recueillies à l’aide de REDCap, puis exportés vers Excel (données tabulaires, .csv).*

1.3 Quelles conventions et procédures utiliserez-vous pour structurer, nommer et gérer les versions de vos fichiers pour vous aider et aider les autres à mieux comprendre comment vos données sont organisées?

Directives

*Des structures de fichier logiques, des conventions d’appellation informatives et des indications claires des versions de fichier aident à mieux utiliser vos données pendant et après votre projet de recherche. Ces pratiques permettront à votre équipe de recherche et vous de vous assurer que vous utilisez toujours la version appropriée de vos données, de réduire la confusion concernant les copies enregistrées sur différents ordinateurs ou médias et d’aider l’équipe à mieux structurer les fichiers.*

*Décrivez les procédures et conventions que votre équipe mettra en œuvre pour assurer une uniformité dans le nommage des fichiers et le contrôle des versions, notamment en donnant des exemples qui seront utilisés pour cette étude.*

*Ressources*

[***Nommage de fichiers et gestion de versions***](https://www.uottawa.ca/bibliotheque/gestion-donnees-recherche/designation-fichiers-disposition-donnees)*, guide des bibliothèques de l’Université d’Ottawa*

[***Nommage des fichiers et des dossiers***](https://cours.ebsi.umontreal.ca/sci6005/co/nommage.html)*, guide de l’EBSI à l’Université de Montréal*

*Exemple*

*Les titres des fichiers seront structurés de la façon suivante : Titreduprojet\_aaaammjj\_contenu\_AA\_v0*

*Où AA représente les initiales du membre de l’équipe responsable de sa création et v0 représente la version du fichier.*

*Ex: LEUCAProject\_20200418\_ethicsmeeting\_BF\_v3.rtf*

*IMProjet\_20210105\_resultats\_SP\_v1.txt*

# Documentation et métadonnées

Les données sont rarement explicites, c’est pourquoi toutes les données de recherche doivent être accompagnées de métadonnées (des renseignements qui décrivent les données selon les meilleures pratiques de la communauté).

Les normes de métadonnées varient d’une discipline à l’autre, mais elles indiquent généralement la personne qui a créé les données, quand et comment elles ont été créées, leur qualité, exactitude et précision ainsi que d’autres caractéristiques nécessaires pour faciliter la découverte, la compréhension et la réutilisation des données. Toute restriction concernant l’utilisation des données doit être expliquée dans les métadonnées et, dans la mesure du possible, des renseignements doivent être fournis sur la façon d’obtenir un accès approuvé aux données.

Fournissez des renseignements sur vos données, afin que d’autres puissent mieux comprendre, interpréter et éventuellement réutiliser vos données pour une analyse secondaire.

2.1 Quelle documentation sera nécessaire pour que les données soient lues et interprétées correctement à l’avenir?

*Directives*

*La documentation de vos données doit inclure toutes les informations pertinentes à la conduite de votre étude, ainsi que les informations requises pour assurer la réutilisabilité et l’interprétabilité de vos données par d’autres chercheurs.*

*Notamment, une bonne documentation comprend les renseignements sur l’étude et son contexte.*

*Il pourrait être pertinent de documenter d’autres éléments, tels que :*

* *le protocole de votre étude;*
* *le cahier de laboratoire;*
* *la méthodologie de recherche utilisée;*
* *le guide des entretiens;*
* *les copies des questionnaires vierges;*
* *les détails des logiciels utilisés;*
* *les définitions des variables;*
* *les vocabulaires propres à votre étude;*
* *les systèmes de classification;*
* *les unités de mesure;*
* *les hypothèses formulées;*
* *le type de format et de fichier des données;*
* *une description des données recueillies et des méthodes de collecte associées;*
* *l’explication de l’analyse et de la codification des données réalisées (y compris les fichiers de syntaxe);*
* *les renseignements sur les personnes ayant contribué au projet et les tâches réalisées par chaque contributeur;*
* *un dictionnaire de données;*
* *le niveau de sécurité que vos données impliquent.*

*Décrivez les types de documents qui seront créées, leurs formats et la manière dont ils contribueront à l’interprétation juste de vos données.*

*Ressources*

[***Comment rédiger un fichier « readme », description des données recueillies et des méthodes de collecte associées***](https://data.research.cornell.edu/data-management/sharing/readme/)*, guide de l’Université Cornell (en anglais seulement)*

[***Modèle de fichier « readme »***](https://uqtr.libguides.com/ld.php?content_id=35593298)*, Université du Québec à Trois-Rivières*

[***Comment rédiger un dictionnaire de données***](https://help.osf.io/article/217-how-to-make-a-data-dictionary)*, guide du Center for Open Science (en anglais seulement)*

[***Les niveaux de sécurité des données***](https://www.tpsgc-pwgsc.gc.ca/esc-src/protection-safeguarding/niveaux-levels-fra.html)*, guide du Ministère de la défense nationale*

[***eLabFTW***](https://crchusj.elab.one/login.php)*, la plateforme de cahier de laboratoire électronique du CRCHUSJ*

[***Utiliser un cahier de laboratoire numérique***](https://www.datacc.org/bonnes-pratiques/utiliser-un-cahier-de-laboratoire-numerique/)*, un guide de DATACC*

*Exemple*

*Voici les documents qui seront créés afin d’assurer la réutilisabilité et l’interopérabilité de nos données.*

1. *Document maître de l’étude, contenant le protocole de recherche, les outils de collecte/analyse (nom, numéro de version, etc.), les procédures d’analyses, les hypothèses et un suivi des changements majeurs apportés aux données de l’étude.*
2. *Document maître des questionnaires, incluant le guide d’entretien, les questionnaires vierges et les fichiers de syntaxe pour la codification des données collectées.*
3. *Dictionnaire de données exporté de REDCap contenant les codes et les variables des questionnaires.*
4. *Fichiers README.txt contenant la description des données recueillies selon la méthode et le jeu données, ainsi que les fichiers XML (Extensive Markup Language) contenant les métadonnées.*
5. *Le PGD final, comprenant le suivi des modifications.*

2.2 Comment ferez-vous en sorte que la documentation soit créée ou saisie de manière cohérente tout au long de votre projet?

*Directives*

*Prenez en compte la manière dont vous saisirez ces informations, l’emplacement où elles seront enregistrées et les personnes impliquées dans la saisie des données.*

*Idéalement, ces éléments sont réfléchis avant la collecte et l’analyse des données, mais il n’est jamais trop tard pour rectifier les habitudes et mettre en place des* ***pratiques de documentation uniformes****.*

*La collecte des renseignements sur les données doit être une partie intégrante des responsabilités des membres de l’équipe et des opérations. Il est donc recommandé de consulter régulièrement les membres de son équipe de recherche afin de connaître les changements possibles dans la collecte ou le traitement des données qui doivent être pris en considération lors de la saisie dans les documents. Répondez en précisant les personnes responsables de la documentation de vos données, les outils utilisés et les moments ou intervalles auxquels la documentation cohérente des données sera vérifiée.*

*Exemple*

*Des protocoles de documentation des métadonnées seront établis avant le début de la collecte et rédigés de manière collective. Ces protocoles contiendront les normes, attentes et processus relatifs à la collecte des données. Chaque membre de l’équipe sera impliqué dans la documentation des métadonnées, mais c’est à Jean Lavoie que revient la responsabilité de vérifier l’uniformité de la documentation. Cette vérification se fera mensuellement jusqu’à la finalisation du projet.*

2.3 Si vous utilisez une norme de métadonnées ou des outils pour documenter et décrire vos données, veuillez les énumérer ici.

*Directives*

*En plus des normes de métadonnées générales, il existe plusieurs normes propres à des domaines précis. Les schémas de métadonnées à privilégier sont les schémas standards, ouverts et lisibles par machine. Choisir un schéma qui répond à ces critères assure l’échange efficace d’information entre les utilisateurs et les systèmes.*

*La documentation de vos données peut également comprendre un vocabulaire contrôlé propre à votre domaine, c’est-à-dire une liste terminologique normalisée pour la description de l’information.*

*L’utilisation d’un outil informatique pour documenter vos données de recherche peut vous aider entre autres à garder une trace des différentes versions et à rassembler à un même endroit vos notes, vos codes et vos logiciels.*

*Si vous utilisez un ou des outils pour documenter et décrire vos données (ex. GitHub, OSF, Procols.io, etc.), mentionnez-le.*

*Ressources*

[***La documentation des données***](https://bib.umontreal.ca/gerer-diffuser/gestion-donnees-recherche?tab=5243848)*, guide des bibliothèques de l’Université de Montréal*

[***Les métadonnées***](https://ukdataservice.ac.uk/learning-hub/research-data-management/document-your-data/metadata/)*, guide du UK Data Service (en anglais seulement)*

*Exemple*

*Nous utiliserons la norme* [***BIDS***](https://bids-specification.readthedocs.io/en/stable/) *(Brain Imaging Data Structure) pour structurer et nommer nos données de neuro-imagerie. Ce faisant, les données provenant des scans seront converties en format JSON, organisées selon le schéma de répertoire BIDS.*

*Les termes MeSH seront privilégiés pour normaliser la description de l’information.*

# Stockage et sauvegarde

La sécurité et l’intégrité de vos données reposent sur la planification de la manière dont vos données de recherche seront stockées et sauvegardées pendant et après votre projet de recherche. Le stockage et la sauvegarde appropriés contribuent non seulement à protéger les données de recherche des pertes catastrophiques (en raison de défaillances du logiciel et du matériel, de virus, de pirates informatiques, de catastrophes naturelles, d’erreurs humaines, etc.), mais y facilitent également l’accès approprié par les chercheurs.

Cette section ne concerne que les données **actives** (en cours de projet). Pour la conservation à long terme, voir la section **Préservation**.

3.1 Quels sont les besoins prévus en matière de stockage pour votre projet (en mégaoctets, giga-octets, téraoctets, etc.)?

*Directives*

*L’estimation de la taille des données que vous allez produire est au cœur de votre stratégie en matière de sauvegarde et de stockage. Votre estimation concernant l’espace de stockage requis doit prendre en considération les exigences pour le versionnage de fichiers, les formats des données, les sauvegardes et la croissance du nombre de fichiers au fil du temps.*

*Exemple*

*Nous estimons que nous produirons environ 200 sondages, 50 entrevues (de 30 à 40 minutes chacune, qualité audio de 128 kb/s) et environ 300 fichiers textes.*

*La taille totale des données, en tenant compte des versions (brutes, maîtresses, analytiques), est estimée à moins de 125 Go.*

3.2 De quelle manière et à quel endroit vos données seront-elles stockées et sauvegardées pendant votre projet de recherche?

*Directives*

*Le risque de perdre des données en raison d’une erreur humaine, de catastrophes naturelles ou d’autres mésaventures peut être atténuées en respectant la règle du 3-2-1 :*

* *Posséder au moins* ***3*** *copies de vos données.*
* *Stocker les copies sur* ***2*** *types de médias différents.*
* *Conserver* ***1*** *copie de sauvegarde hors site.*

*Chaque méthode de stockage comporte des avantages et des inconvénients qui doivent être pris en considération au moment de déterminer la solution la plus appropriée aux particularités de votre projet de recherche.*

*Les médias amovibles (DVD, clés USB, etc.) ne sont pas recommandés. Les médias fixes (lecteurs de disque dur d’ordinateur de bureau ou d’ordinateur portatif) ne doivent être utilisés qu’en combinaison avec des médias en réseau (lecteurs en réseau ou serveurs basés sur l’infonuagique).*

*Décrivez, le plus précisément possible, les appareils/plateformes sur lesquels vos données seront stockées. Comment les endroits de stockage sélectionnés assurent-ils la sécurité de vos données? Qui y aura accès?*

*Si vous disposez de* ***copies papiers*** *de vos données, mentionnez l’endroit où elles seront entreposées et les mesures que vous prendrez pour garantir leur sécurité.*

*Exemple*

*Les données de recherche seront recueillies par les membres de l’équipe et stockées sur leurs disques durs réseau, puis versées sur le serveur de Calcul Québec chaque semaine. L’ensemble des données hébergées sur le serveur de Calcul Québec sera téléchargé sur le serveur réseau du CHUSJ pour les préserver une fois le projet terminé.*

3.3 De quelle manière l’équipe de recherche et les autres collaborateurs vont-ils accéder aux données, les modifier et y contribuer tout au long du projet?

*Directives*

*La solution idéale répond aux critères suivants :*

1. *Facilite la collaboration;*
2. *Assure la sécurité des données;*
3. *Requiert une formation minimale.*

*Décrivez les groupes d’utilisateurs et les accès qui leurs seront accordés. Décrivez également les mesures de sécurité associées à la plateforme de partage de données et de collaboration de votre choix. Le transfert de fichiers par courrier électronique n’est pas une solution robuste ou sécuritaire. De même, les services de partage de fichiers commerciaux (comme Google Drive et Dropbox) facilitent l’échange de fichiers, mais ne représentent pas une solution permanente, sécuritaire ou conforme aux exigences en matière d’hébergement en sol canadien.*

*Exemple*

*Chaque membre de l’équipe de recherche (chercheurs, stagiaires et membres du personnel autorisés) détient un compte protégé par mot de passe sur la plateforme de Calcul Québec. L’accès à notre espace de projet sécurisé est attribué par le chercheur principal. Cet accès permet de visualiser et d’éditer les données.*

# Préservation

Cette section vous aidera à élaborer votre stratégie de préservation afin de garantir que vos données seront disponibles et utilisables dans un avenir prévisible après la fin de votre projet. La stratégie de conservation des données dépendra

* de la valeur de réutilisation potentielle;
* des exigences en matière de conservation ou de destruction des données;
* des ressources requises pour la saine organisation des données et pour assurer qu’elles demeurent utilisables dans le futur.

Dans certains cas, il peut être souhaitable de conserver toutes les versions des données (par exemple, les données brutes, traitées, analysées, définitives), mais dans d’autres cas, il peut être préférable de conserver uniquement les données sélectionnées ou définitives (par exemple, les transcriptions au lieu d’entrevues sonores).

4.1 Où déposerez-vous vos données pour les conserver à long terme et y accéder après la fin de votre projet de recherche?

*Directives*

*Idéalement, la question de la préservation des données de recherche est examinée tôt dans votre projet. Il vaut mieux être préparé, afin de s’assurer que tous les membres de l’équipe sont sur la même page. Les décisions relatives à la conservation des données peuvent être dictées par des politiques internes, des politiques externes (par exemple : organismes de financement, éditeurs de revues) ou par une compréhension de la valeur durable d’un ensemble de données déterminé.*

*La nécessité de préserver les données à court terme (c.-à-d. à des fins de vérification par des pairs) ou à long terme (pour les données de valeur durable) influencera le choix de l’archive ou du dépôt de données. Une analogie utile consiste à penser à créer un « testament de vie » pour les données, c’est-à-dire un plan qui décrit la façon dont les futurs chercheurs continueront d’avoir accès aux données.*

*Vous vous questionnez sur la valeur de vos données ou sur les dépôts de données qui leur conviendraient? Communiquez avec le service de soutien en gestion des données de recherche à* [***soutien.gdr.hsj@ssss.gouv.qc.ca***](mailto:soutien.gdr.hsj@ssss.gouv.qc.ca)*.*

*Exemple*

*Après la fin du projet de recherche, les données collectées seront conservées sur le serveur interne du CHUSJ pendant une période d’un an pour qu’elles soient disponibles dans l’éventualité qu’une vérification de nos résultats soit nécessaire. Comme nos données ont une valeur durable potentielle, la fin du projet sera également le moment où notre équipe produira un article de données afin de faire connaître l’existence des données générées dans le cadre de notre étude.*

*Une fois l’année écoulée, le chercheur principal sera en charge de la compression des fichiers, à l’exception des fichiers vidéo parce qu’ils sont susceptibles d’identifier les participants, et de leur archivage dans Borealis.*

4.2 Indiquez comment vous vous assurerez que vos données sont prêtes pour la conservation.

*Directives*

*À prendre en considération : formats de fichier pour la conservation qui préservent l’intégrité des données; anonymisation et dépersonnalisation des fichiers, y compris les fichiers de documentation.*

*Certains formats de données sont idéals pour la* ***conservation à long terme des données****. Par exemple, les formats de fichier non-propriétaires, comme le format texte (.txt) et les valeurs séparées par des virgules (.csv), sont considérés comme des formats conçus pour la conservation. Le UK Data Archive (dépôt de données du Royaume-Uni) fournit un tableau utile des formats de fichier pour divers types de données. Il faut se rappeler que les fichiers qui sont convertis d’un format à l’autre peuvent perdre des renseignements (p. ex. en convertissant un fichier TIFF non compressé à un fichier JPG compressé). Ainsi, les modifications apportées aux formats de fichier doivent être documentées.*

*Déterminez les étapes à suivre à la fin d’un projet pour s’assurer que les données que vous choisissez de préserver ou de partager sont anonymes, exemptes d’erreur et converties dans les formats recommandés en assurant un risque minimal de perte de données.*

*Ressources*

[***L’anonymisation et la désidentification***](https://researchdata.library.ubc.ca/deposit/anonymize-and-de-identify/)*, guide des bibliothèques de l’Université de la Colombie-Britannique (en anglais seulement)*

[***L’anonymisation des données quantitatives***](https://ukdataservice.ac.uk/learning-hub/research-data-management/anonymisation/anonymising-quantitative-data/)*, guide du UK Data Service (en anglais seulement)*

[***L’anonymisation des données qualitatives***](https://ukdataservice.ac.uk/learning-hub/research-data-management/anonymisation/anonymising-qualitative-data/)*, guide du UK Data Service (en anglais seulement)*

*Exemple*

*Les données susceptibles d’identifier les participants seront anonymisées avant leur préservation. Notamment, les noms des participants seront remplacés par des identifiants uniques randomisés et les données cliniques seront anonymisées pour empêcher l’identification des participants.*

*Nous avons confiance que les formats .csv et .rtf ne sont pas un obstacle à la réutilisation des données, c’est pourquoi nous utiliserons ces formats pour la préservation.*

# Partage et réutilisation

La plupart des organismes subventionnaires canadiens ont maintenant des politiques exigeant que les données de recherche soient partagées au moment de la publication des résultats de la recherche ou dans un délai raisonnable. Il existe forcément un équilibre entre le souhait légitime des chercheurs de maximiser les résultats de leurs recherches par la publication et la contribution du partage des données à la visibilité et à l’impact de la recherche. De plus, il est tout aussi important de protéger la vie privée des répondants et de traiter de façon les données sensibles de manière appropriée.

Le comité d’éthique de la recherche du CHU est votre meilleur allié pour toute question liée à l’utilisation secondaire ou le partage de données cliniques. Assurez-vous de leur communiquer vos questions et vos préoccupations **avant le début du projet**.

5.1 Quelles données partagerez-vous et sous quelle forme les partagerez-vous (par exemple, données brutes, traitées, analysées, finales)?

*Directives*

*Les* ***données brutes*** *sont des données obtenues directement de l’instrument, de la simulation ou de l’enquête.*

*Les* ***données traitées*** *découlent d’une certaine manipulation des données brutes afin d’éliminer les erreurs ou les valeurs aberrantes, de préparer les données en vue de l’analyse, d’obtenir de nouvelles variables ou d’anonymiser les participants humains.*

*Les* ***données analysées*** *sont les résultats d’une analyse qualitative, statistique ou mathématique des données traitées. Elles peuvent être présentées sous forme de graphiques, de diagrammes ou de tableaux statistiques.*

*Les* ***données définitives*** *sont des données traitées qui ont été converties, au besoin, dans un format pouvant être préservé. Examinez les données qui peuvent devoir être partagées afin de respecter les exigences institutionnelles ou de financement et les données qui peuvent faire l’objet de restrictions en raison de préoccupations relatives à la confidentialité, au respect de la vie privée et à la propriété intellectuelle.*

*Exemple*

*Les résultats des sondages traités seront rendus disponibles. Le fichier .csv les contenant sera traité afin de retirer tout identifiant direct. Les transcriptions des entrevues (.rtf) seront également traitées afin de supprimer tout identifiant direct. Les données cliniques (.csv) ne seront disponibles que sur demande, suivant la procédure d’accès aux données de santé définie au CHUSJ.*

5.2 Avez-vous songé au type de licence d’utilisateur final à inclure avec vos données?

*Directives*

*Les licences déterminent les utilisations permises de vos données. Les organismes de financement et les dépôts de données peuvent avoir des* ***exigences relatives à la licence d’utilisation****. Si ce n’est pas le cas, ils peuvent vous guider dans la création d’une licence. Une fois la licence créée, veuillez inclure une copie de votre licence d’utilisation avec votre plan de gestion des données. Veuillez noter que seuls les titulaires des droits de propriété intellectuelle peuvent émettre une licence, il est donc essentiel de préciser à qui appartiennent ces droits lors de la création d’une licence.*

*Il existe plusieurs types de licences standards mises à la disposition des chercheurs. En fait, pour la plupart des ensembles de données, il est plus facile d’utiliser une licence standard plutôt que de concevoir une licence personnalisée. Veuillez noter que même si vous choisissez de rendre vos données publiques, il est préférable de l’indiquer de manière explicite en utilisant une licence comme celle CC0 de Creative Commons.*

*Si vous travaillez avec des données provenant de communautés autochtones ayant statué la propriété collective des données recueillies, les droits de propriété intellectuelle peuvent ne pas vous revenir. Le cas échéant, mentionnez à quelle communauté revient les droits sur les données recueillies.*

*Ressources*

[***Licences Open Data Commons***](https://opendatacommons.org/licenses/) *(en anglais seulement)*

[***Licences Creative Commons***](https://chooser-beta.creativecommons.org/) *(en anglais seulement)*

[***Comment obtenir une licence pour les données de recherche***](https://www.dcc.ac.uk/guidance/how-guides/license-research-data)*, guide du Digital Curation Centre (en anglais seulement)*

[***Comprendre les principes de PCAP des Premières Nations***](https://fnigc.ca/fr/les-principes-de-pcap-des-premieres-nations/)*, guide du CGIPN*

[***Boîte à outils des principes de la recherche en contexte autochtone***](https://cssspnql.com/produit/boite-a-outils-des-principes-de-la-recherche-en-contexte-autochtone-contributions-evaluees-par-les-pairs/)*, conçu par la Commission de la santé et des services sociaux des Premières Nations du Québec et du Labrador – CSSSPNQL*

*Exemple*

*Le chercheur principal (Jean Tremblay) sera le titulaire des droits de propriété intellectuelle des données et sera responsable de la procédure d’obtention d’une licence CC. La licence reste à établir, mais nous savons déjà qu’une réutilisation des données sera permise, tel que stipulé dans les formulaires de consentement signés par nos participants. De plus, nous nous sommes entendus pour ne permettre un partage secondaire des données réutilisées que dans les cas où ces données font l’objet d’une licence permettant la réutilisation.*

5.3 Comment rendrez-vous accessibles vos données à la communauté des chercheurs?

Dépôt spécifique au domaine d’étude (ex. Repéré sur re3data)

Dépôt généraliste (ex. Borealis)

Bouche-à-oreille

Publications (ex. article, article de données, etc.)

Conférences

Réseaux sociaux (ex. ORCID, Twitter, LinkedIn, ResearchGate, Google Scholar, etc.)

Autre :

*Directives*

*Précisez vos choix.*

*Comment pourra-t-on accéder aux données (service Web, ftp, etc.)? Dans la mesure du possible, choisissez un dépôt qui attribuera un identifiant constant (comme un DOI, Digital Object Identifier) à votre ensemble de données. Cela permettra ainsi d’assurer un accès stable à l’ensemble de données et il pourra être repérable par l’entremise de divers outils de recherche.*

*Une des meilleures façons d’orienter d’autres chercheurs vers vos ensembles de données déposés consiste à les citer de la même façon que d’autres types de publications (articles, livres, procédures). Veuillez noter que certains dépôts de données établissent également des liens entre les ensembles de données et les articles connexes, ce qui accroît ainsi la visibilité des publications.*

*Il existe aujourd’hui quelques dizaines de revues de données, soit des revues savantes composées d’articles généralement évalués par les pairs présentant des jeux de données. Ce type d’article est aussi de plus en plus courant dans des revues savantes traditionnelles. La publication d’un* ***article de données*** *permet notamment de publiciser l’existence de vos données sans publier publiquement votre jeu de données.*

[***Borealis***](https://borealisdata.ca/fr/)*, aussi appelé le Dataverse canadien, est un dépôt de données généraliste où vous pouvez verser vos données de recherche* ***si elles ne contiennent aucune information susceptible d’identifier vos participants****. En naviguant ce dépôt, vous serez en mesure d’établir si l’institution à laquelle vous êtes affiliés détient un espace de dépôt sur Borealis. Vous pouvez également consulter* [***Re3data***](https://www.re3data.org/) *pour découvrir des dépôts de données spécifiques à votre discipline.*

*Ressources*

[***Comment citer un jeu de données***](https://www.dcc.ac.uk/guidance/how-guides/cite-datasets)*, un guide du Digital Curation Centre (en anglais)*

[***Qu’est-ce qu’un article de données?***](https://coop-ist.cirad.fr/gerer-des-donnees/publier-un-data-paper/1-qu-est-ce-qu-un-data-paper)*, un guide du CIRAD*

[***Identifiants pérennes***](https://doranum.fr/identifiants-perennes-pid/)*, un guide de DORANum*

*Exemple*

Dépôt spécifique au domaine d’étude (ex. Repéré sur re3data)

Publications (ex. article, article de données, etc.)

Réseaux sociaux (ex. ORCID, Twitter, LinkedIn, ResearchGate, Google Scholar, etc.)

*Un Digital Persistant Identifier (DOI) a été créé pour cet ensemble de données avec l’aide de DataCite Canada. Tel que mentionné à la question 4.1, nous planifions publier un article de données pour publiciser l’existence de nos données.*

*Une partie des données (les réponses aux sondages et les transcriptions des entrevues anonymisées) seront publiées dans un dépôt disciplinaire.*

*Les données seront également publicisées sur les comptes ORCID des chercheurs impliqués lorsque nos résultats auront été publiés.*

# Responsabilités et ressources

La notion de gestion des données fait référence au « quoi » et au « comment » des opérations de gestion liées aux données tout au long du cycle de vie du projet.

La notion d’intendance des données met plutôt l’accent sur « qui » est responsable de s’assurer que la gestion des données soit faite.

Un gros projet de recherche aura par exemple plusieurs personnes responsables. Le chercheur principal devra déterminer en début du projet quelles personnes de l’équipe auront des responsabilités en matière de gestion des données pendant et après le projet.

6.1 Désignez qui sera responsable de la gestion des données pendant et après le projet et les principales tâches de gestion des données dont cette personne sera responsable.

*Directives*

*Votre plan de gestion des données vous a permis d’identifier les activités importantes relatives à la gestion des données dans le cadre de votre projet.*

*Indiquez les responsables, personnes ou organisations, de la mise en place et de la mise à jour du plan de gestion des données, ainsi que leurs affiliations et coordonnées s’il y a lieu.*

*Mentionnez également tout outil ou formation nécessaires afin de préparer le personnel à assumer ces fonctions.*

*Exemple*

*Jean Tremblay, PhD, de l’Université de Montréal, sera la personne responsable de la mise en place et de la mise à jour de ce plan de gestion de données. Il sera également le point de contact pour toute question liée aux données de recherche. En tant que chercheur principal, il aura la responsabilité d’évaluer les besoins en formation des chercheurs impliqués. Vous pouvez le contacter à* [*jean.tremblay.hsj@ssss.gouv.qc.ca*](mailto:jean.tremblay.hsj@ssss.gouv.qc.ca)*.*

6.2 De quelle façon allez-vous administrer les responsabilités quant aux activités de gestion des données s’il y a des changements importants de personnel, notamment un changement de chercheur principal?

*Directives*

*Comme les changements au sein des équipes de recherche sont communs, assurez-vous que les membres de votre équipe soient prêts à prendre la responsabilité du plan de gestion de données et à s’assurer qu’il est mis à jour régulièrement.*

*Indiquez une stratégie de planification de la relève pour ces données dans l’éventualité où une ou plusieurs personnes responsables des données quittent le projet (par exemple, un étudiant aux cycles supérieurs qui quitte après l’obtention de son diplôme).*

*Décrivez le processus à suivre dans l’éventualité où le chercheur principal quitte le projet. Dans certains cas, un co-chercheur, le département ou la division qui supervise cette recherche en assumera la responsabilité.*

*Exemple*

*Si Dr Tremblay devait quitter le projet, ses responsabilités concernant le plan de gestion de données seraient confiées à Guillaume Dion.*

*Si jamais ces deux personnes ne sont pas capables de remplir leurs responsabilités, un autre membre de l’équipe sera chargé de superviser la mise en place du plan de gestion de données, avec l’accord de l’équipe de recherche.*

6.3 De quelles ressources aurez-vous besoin pour mettre en œuvre votre plan de gestion des données? À combien estimez-vous le coût global de la gestion des données?

*Directives*

*Cette évaluation doit comprendre les coûts de gestion des données encourus au cours du projet ainsi que les coûts requis pour le soutien à plus long terme des données après la fin du projet. Les points à considérer dans la dernière catégorie de dépenses comprennent les coûts d’entretien et de prestation d’un accès à long terme aux données.*

*Certains organismes de financement indiquent de façon explicite le soutien qu’ils fourniront afin de respecter les coûts relatifs à la préparation des données à déposer. Cela peut comprendre les aspects techniques de la gestion de données, les exigences en matière de formation, le stockage et la sauvegarde des fichiers et le travail du personnel ne faisant pas partie du projet.*

*Exemple*

*Les coûts associés à la gestion des données de recherche devraient être de l’ordre de 1500 à 2000 dollars canadiens par an (ces coûts comprennent le temps nécessaire pour s’assurer que le PGD soit respecté, le coût du stockage local, etc.) En outre, l’archivage des données à la fin du projet devrait représenter un coût unique de 500 $ CAD.*

# Conformité éthique et juridique

Les chercheurs et chercheuses, ainsi que leurs équipes, doivent connaître les politiques et les processus, éthiques et juridiques, auxquels leur gestion des données de recherche doit se conformer.

Dans votre plan de gestion des données, vous devez indiquer la façon dont vous préparerez, stockerez, partagerez et archiverez vos données de façon à vous assurer que les renseignements sur les participants et participantes soient protégés tout au long du cycle de vie de la recherche contre la divulgation, l’utilisation préjudiciable ou les liens inappropriés avec d’autres données personnelles.

Il peut arriver des situations où certaines données et métadonnées **ne peuvent pas être rendues publiques** en raison de politiques ou de considérations légales. Toutefois, dans un esprit de sciences ouvertes, la position par défaut devrait être que toutes les données et métadonnées de recherche soient publiques.

7.1 Si votre projet comprend des données sensibles, comment vous assurerez-vous qu’il est géré de manière sécuritaire et que les données sont accessibles uniquement aux membres approuvés du projet?

*Directives*

*Examinez où et comment les données sensibles ayant une valeur à long terme reconnue doivent être rendues disponibles et à qui, de même que pendant combien de temps elles doivent être archivées. Ces décisions doivent respecter les exigences du comité d’éthique de la recherche. Les méthodes utilisées pour partager les données dépendront d’un certain nombre de facteurs comme le type, la taille, la complexité et le degré de sensibilité des données. Décrivez les problèmes prévus dans le partage des données, ainsi que les causes et les mesures possibles pour les atténuer. Les problèmes peuvent comprendre, entre autres, la confidentialité, l’absence de consentement ou les préoccupations au sujet des droits de propriété intellectuelle. Dans certains cas, une période d’embargo peut être* *justifiée; ces cas peuvent être définis dans la politique sur les données de recherche d’un organisme de financement.*

*Des restrictions peuvent être imposées en limitant l’accès physique aux dispositifs de stockage, en enregistrant les données sur des ordinateurs qui ne possèdent pas d’accès au réseau externe (c.-à-d. accès à Internet), grâce à une protection par mot de passe, et en codant les fichiers. Les données sensibles ne doivent jamais être partagées par courriel ou au moyen de services de stockage infonuagique comme Dropbox.*

*Exemple*

*L’accès aux données de recherche stockées sur le serveur du CHUSJ nécessitera une vérification à deux facteurs avec l’application Microsoft Verify.*

*Seuls les utilisateurs disposant d’une adresse courriel institutionnelle du CHUSJ, qui ont été vérifiés par le coordinateur du projet, seront autorisés à y accéder.*

7.2 Quelles stratégies allez-vous mettre en œuvre pour gérer les utilisations secondaires des données sensibles, le cas échéant?

*Directives*

*L’obtention du consentement approprié des participants de recherche constitue une étape importante pour assurer au comité d’éthique de la recherche que les données de recherche peuvent être partagées avec des chercheurs ne faisant pas partie de votre projet.*

*La déclaration de consentement peut indiquer certaines conditions précisant l’utilisation des données par d’autres chercheurs. Par exemple, elle peut stipuler que les données seront uniquement partagées à des fins de recherche sans but lucratif ou que les données ne seront pas croisées avec des données personnelles provenant d’autres sources.*

*Exemple*

*Le jeu de données nécessaire pour vérifier les résultats sera indexé sur la plateforme de partage de données Zenodo. Cet élément d’indexation contiendra des informations bibliographiques, quantitatives et qualitatives de base sur le projet de recherche.*

*Pour accéder aux données réelles, les utilisateurs devront contacter le coordinateur de la recherche et donner une description de la manière dont ils prévoient de réutiliser les données et signer un accord afin de garantir que les données sont utilisées de manière éthique. Les demandes de réutilisation des données seront évaluées individuellement.*

7.3 Comment allez-vous gérer les questions juridiques, éthiques et de propriété intellectuelle?

*Directives*

*Le respect des règlements et des lois sur la protection de la vie privée qui peuvent imposer des restrictions de contenu dans les données devrait être discuté avec le bureau de la recherche ou le responsable de la protection de la vie privée de votre établissement. L’apport du comité d’éthique de la recherche est aussi essentiel dans ce processus.*

*Décrivez ici ce qui en est de la propriété des données, de l’octroi de licences et de la propriété intellectuelle. Les conditions énoncées pour la réutilisation des données doivent être clairement mentionnées et respecter les exigences juridiques et éthiques pertinentes, le cas échéant (p. ex. consentement du sujet, permissions, restrictions, contrats avec des partenaires du secteur privé, etc.).*

*Si vos données* *seront collectées auprès des communautés autochtones, mentionnez la manière dont vous vous assurerez de respecter les* [***principes CARE***](https://www.gida-global.org/care) *et les* [***principes PCAP***](https://fnigc.ca/fr/les-principes-de-pcap-des-premieres-nations/)*.*

*Exemple*

*Le projet a reçu l’autorisation du comité d’éthique du CHUSJ avant de commencer. Diverses inquiétudes liées à la gestion des données ont été portées à l’attention du comité. Toute question pouvant représenter un obstacle éthique ou juridique sera portée à l’attention du comité afin de résoudre la situation dans les plus brefs délais.*

*Le dossier de ce projet sur Nagano sera également mis à jour régulièrement, afin que l’administration du CHUSJ soit au courant de tout changement concernant le stockage, l’archivage et la sécurité des données de recherche de notre projet.*