

Guide de rédaction du plan de gestion des données (PGD) de l'Université de Sherbrooke

Ce guide est issu de la version en ligne de l'[Assistant PGD](#) de l'Alliance de recherche numérique du Canada (L'Alliance). Les exemples et conseils fournis sont tirés du [matériel produit par L'Alliance \(CC-BY\)](#) et du [PGD produit par l'Université du Québec à Trois-Rivières](#).

Pour rédiger votre plan, un gabarit en format Word est disponible à partir du [guide en ligne](#). Pour toute question, contactez l'équipe de gestion des données de recherche (GDR) : gdr@usherbrooke.ca.

1. Vos données

1.1 Types et formats de données

Durant la collecte et l'analyse, quels types de données projetez-vous recueillir, créer, utiliser ou consigner pour l'ensemble de votre projet ? Pour chaque type de données, quel(s) seront le(s) format(s) de fichiers ? Ces formats permettront-ils la réutilisation, le partage et l'accès à long terme aux données ?

Directives :

Répondez à cette question en précisant le processus de collecte et d'analyse et le **type des données** recueillies (numériques, textuelles, audio, vidéo, géospatiales, etc.), et ce, pour l'ensemble des données (primaires, secondaires, dérivées, traitées, analysées, etc.).

Il est préférable d'utiliser des **formats** de fichier ouverts ou largement utilisés par une communauté. Réfléchissez au format initial des données et au format qui pourrait être éventuellement utilisé pour préserver l'accès et la conservation des données à moyen ou long terme. À noter qu'il est possible de convertir les fichiers dans un format plus pérenne, mais il peut en résulter des pertes d'information.

Exemple :

Nous recueillerons des données de sondages qui seront ensuite exportées en format tabulaire (XLS). Nous mènerons également des entrevues semi-structurées et des groupes de discussion, ce qui produira des données audio et textuelles (MP3 et DOC). Les formats choisis pour la période de collecte et d'analyse sont ceux couramment utilisés par notre équipe. Au moment de déposer et de partager nos données, nous les convertirons en formats ouverts (DOC vers RTF et .XLS vers CSV).

Ressource :

- [Section « Données : types et formats » du guide GDR](#)

1.2 Convention de nommage des fichiers

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 :

Gardez une trace des différentes copies et **versions** des fichiers, des fichiers détenus dans différents formats ou emplacements, et de toute information croisée entre les fichiers. L'arborescence de classement doit être pensée en amont de même que les conventions de nommage des fichiers et dossiers.

Exemple :

Dans la base de données, à mesure que les enregistrements sont créés ou modifiés, ils seront horodatés avec leurs modifications. Le nom de la personne qui apporte les modifications est également demandé. Un fichier journal est conservé pour consigner l'information sur les modifications apportées à la base de données. Le titre des fichiers respectera [les recommandations du Service des bibliothèques et archives de l'UdeS](#) et il influera sur le nom de la variable étudiée, par exemple : « `echantillon_orchidees_v02_2022-04-15` ».

Les membres de l'équipe de recherche suivront une formation à ce sujet. La documentation accompagnant les jeux de données décrira l'arborescence ainsi que la convention de nommage de sorte que tous les membres de l'équipe pourront continuer de s'y référer pendant tout le projet.

Ressource :

- [Section « Organisation et nommage des dossiers et des fichiers » du guide GDR](#)

2. Documentation et métadonnées

2.1 Documentation de référence

Quels documents seront nécessaires pour que les données soient lues et interprétées correctement ? Dans quels documents seront consignées les informations relatives aux étapes du projet de recherche ?

Directives :

Une bonne documentation comprend des renseignements sur l'étude, des descriptions des éléments tels que les **variables** et d'autres renseignements contextuels requis pour que d'autres chercheuses et chercheurs puissent comprendre et utiliser les données.

Exemples d'éléments à documenter : les **méthodes** utilisées, les définitions des variables et des concepts clés, les vocabulaires, les systèmes de classification, les **unités de mesure**, les **hypothèses** formulées, la description des données recueillies et des méthodes de collecte et d'analyse, les grilles de codes ou les abréviations, les fichiers de **traitement**, les **outils** ou **logiciels** utilisés et les renseignements sur les personnes ayant travaillé au projet et réalisé chacune des tâches, etc.

La documentation des jeux de données peut donc également inclure un vocabulaire contrôlé, qui est une liste terminologique normalisée pour la description de l'information. Parmi les exemples de vocabulaires contrôlés, citons le [Répertoire de vedettes-matières de l'Université Laval](#) ainsi que les [Library of Congress Subject Headings](#) (LCSH).

Exemple :

Il faudra quatre documents différents.

1. Une **description du projet**, qui comprendra une description du processus entrepris pour déterminer les divers fichiers de données historiques du recensement ;
2. Une **description de la structure** de champ dans la base de données utilisée (p. ex., si un champ est requis, utilise un vocabulaire contrôlé, est reproductible, etc.) ;
3. Les **instructions de saisie des données** suivies pour remplir la base de données ;
4. Les **listes montrant les valeurs possibles** des divers champs de vocabulaire contrôlé et les listes de substitution utilisées pour traduire automatiquement l'anglais en français et vice versa dans les champs de vocabulaire contrôlé appariés.

Ressource :

- [Section « Documenter son projet et ses processus » du guide GDR](#)

2.2 Mise à jour de la documentation

Comment s'assurer que la documentation soit créée ou saisie de manière cohérente tout au long de votre projet ?

Directives :

Examinez les façons dont vous consignerez ces informations et l'emplacement où elles seront enregistrées, idéalement avant de procéder à l'analyse et à la collecte de données, afin d'assurer l'exactitude, l'uniformité, l'exhaustivité et le repérage des documents.

Éléments à considérer : consultations régulières de l'équipe pour intégrer des changements, rédaction et utilisation de protocoles préétablis (incluant des processus et règles à suivre), utilisation de normes ou nomenclatures définies, tenues de registres, etc.

Exemple :

Nous élaborerons et mettrons en œuvre des protocoles clairs de documentation des dossiers et des fichiers avant le début de la collecte des données. Ces protocoles expliqueront les attentes, les normes et les processus pour la saisie et la mise en œuvre des documents à l'appui du projet de recherche. Parmi les exemples de sujets couverts par ces protocoles, on compte les conventions d'appellation des fichiers, le contrôle des versions des fichiers et la structure des dossiers.

Les membres de l'équipe auront l'occasion de contribuer à ces protocoles et, une fois qu'ils seront définitifs, ils seront mis en œuvre pour soutenir les phases actives du projet de recherche. Les documents seront conservés dans des dossiers clairement nommés dans l'espace du dossier du projet de recherche et seront accessibles à tous les membres de l'équipe de recherche.

Ressource :

- [Section « Documenter son projet et ses processus » du guide GDR](#)

2.3 Fichier de métadonnées

Si vous utilisez une norme de métadonnées ou des outils pour décrire vos données, énumérez-les ici.

Directives :

Les métadonnées sont littéralement « des données sur les données », c'est-à-dire qu'elles décrivent et définissent les caractéristiques d'autres données ([CODATA](#)). Cette question traite des métadonnées qui décriront vos fichiers lors de leur dépôt. La qualité des informations et leur format d'encodage seront donc importants pour que vos données soient facilement repérables et, éventuellement, réutilisables. Certains dépôts offrent un formulaire que vous n'aurez qu'à remplir. Enrichissez vos données de descripteurs riches concernant les sujets couverts, la méthodologie ou l'échantillon, par exemple.

Exemples :

Exemple 1

Nous utiliserons les normes proposées par le dépôt choisi, c'est-à-dire le Dataverse de l'Université de Sherbrooke. Le personnel responsable de la gestion de ce dépôt institutionnel a créé un formulaire que nous pourrions simplement remplir et un service de révision est offert pour assurer la qualité de ce que nous inscrivons comme information.

Exemple 2

Nous avons conclu que la norme de métadonnées de la Data Documentation Initiative (DDI) était optimale pour nos données de sondage, d'autant plus que notre plateforme de collecte de données utilise également des normes de métadonnées DDI.

Les normes de métadonnées qui seront utilisées pour appuyer la recherche qualitative n'ont pas encore été définies. Nous consulterons le personnel des bibliothèques, d'autres chercheuses et chercheurs effectuant de la recherche qualitative, ainsi que le Qualitative Data Repository (QDR) situé à la *Maxwell School of Citizenship and Public Affairs de la Syracuse University*.

Ressource :

- [Section « Métadonnées » du guide GDR](#)

3. Stockage et sauvegarde

3.1 Espace de stockage

Quels sont les besoins de stockage (en mégaoctets, giga-octets, téraoctets, etc.) et de durée pendant les phases actives de votre projet ?

Directives :

Les estimations concernant l'**espace de stockage** doivent prendre en considération les exigences du **versionnage** de fichiers, les **sauvegardes** de sécurité et la croissance du nombre de fichiers au fil du temps. Le versionnage des fichiers consiste à garder des copies des versions plus anciennes des documents, cela permet de retracer facilement le processus de travail et d'identifier des erreurs au besoin. La méthode 3-2-1 est recommandée pour la sauvegarde (voir la ressource). Ces deux facteurs font en sorte que la quantité de fichiers augmentera substantiellement au cours du projet. Prenez en considération les éléments suivants :

1. la phase de collecte et la phase d'analyse des données,
2. les copies de sauvegarde,
3. le versionnage de fichiers,
4. la croissance du nombre de fichiers au fil du temps.

Exemple :

Nous estimons une cueillette d'environ 800 sondages, 20 entrevues (environ 30 minutes chacune) et deux séances de groupe de discussion (environ 90 minutes chacun). La taille totale des données, en tenant compte des versions (brutes, maîtresses, analytiques), est estimée à moins de 30 Go. Ces exigences en matière de stockage ont été prises en compte par nos ressources informatiques, ce qui garantit que les besoins d'espace à venir seront satisfaits pour toute la durée du projet.

Ressource :

- [Section « Stockage et accès » du guide GDR](#)

3.2 Méthode de stockage et de sauvegarde

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 atténué en respectant la **règle du 3-2-1** pour les données non sensibles :

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

Décrivez les procédures qui permettront de protéger les **données sensibles** recueillies au cours de votre étude. Cela peut inclure le stockage des données identificatoires séparément des données anonymisées, ainsi que les mesures de protection des fichiers.

Les **plateformes de stockage** de vos données, incluant vos données sensibles, doivent se conformer aux **directives de la Direction de la sécurité de l'information** de l'Université de Sherbrooke (consulter la section 5.5 du document : enregistrement de données ou de documents).

Si vous comptez utiliser les services de plateformes numériques comme Calcul Canada ou Calcul Québec, faites-en la mention en expliquant brièvement leur méthode de fonctionnement.

Établissez un **horaire de sauvegarde** des données, la configuration automatique étant idéale. Les données peuvent être stockées au moyen d'un média optique ou magnétique qui peut être amovible (p. ex. DVD et clés USB), fixe (p. ex. lecteurs de disque dur d'ordinateur de bureau ou d'ordinateur portatif) ou en réseau (p. ex. lecteurs en réseau ou serveurs basés sur l'infonuagique). 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.

Exemples :

Exemple 1

Pendant le projet de recherche, les données seront stockées sur un serveur de Calcul Canada, qui possède sa propre politique de fréquence de sauvegarde. Un espace de stockage de 1 To sur le serveur Graham est d'ailleurs déjà alloué au projet à la suite d'une demande au service d'accès rapide. Une copie de chaque fichier sera conservée sur le serveur de Calcul Canada, une autre copie sera également conservée sur un disque dur externe dans le bureau de la chercheuse principale, une troisième copie sera conservée dans l'espace institutionnel de stockage infonuagique OneDrive de la chercheuse principale.

Exemple 2

Les données numériques sont stockées dans des tableurs, des bases de données, des fichiers de traitement de texte et des carnets de laboratoire. Toutes les données enregistrées dans les carnets de terrain ou de laboratoire sont transcrites numériquement, et les copies en papier seront archivées pendant au moins cinq (5) ans après la fin du projet. Les données numériques seront conservées dans l'espace de stockage réservé à notre équipe sur les serveurs institutionnels et une sauvegarde hebdomadaire sur un disque dur externe et sur l'ordinateur de la chercheuse principale assurera de respecter la règle 3-2-1.

Ressource :

- [Section « Stockage et accès » du guide GDR](#)

3.3 Accès aux données pendant le projet

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

Directives :

Assurer la **sécurité des données** lors de la transmission de données entre les espaces de travail ou au sein d'équipes de recherche n'est pas toujours simple selon l'infrastructure de gestion des données. Le courrier électronique pour le transfert de données n'est pas une solution robuste ou sécuritaire. Les services de partage de fichiers commerciaux de tiers (comme Google Drive et Dropbox) facilitent l'échange de fichiers, mais ne sont pas nécessairement permanents ou sécuritaires et sont souvent situés à l'extérieur du Canada.

De plus, si vos données contiennent des informations confidentielles, vous devez décrire comment votre méthode de stockage assurera la protection de ces données (accès restreint, mots de passe, etc.)*. Les partenaires externes à l'Université doivent également se conformer à la **Politique de sécurité de l'information et la Directive relative à l'utilisation, à la gestion et à la sécurité des actifs informationnels**.

*Consulter les sections 6.4.4- 6.5- 6.6 de la **Directive relative à l'utilisation, à la gestion et à la sécurité des actifs informationnels** du document (*comptes invités, comptes à privilèges spéciaux, mobilité*)

Exemples :

Exemple 1

Toutes les données seront stockées sécuritairement dans notre espace de projet de recherche infonuagique fourni par Calcul Canada et ne seront accessibles qu'aux membres de l'équipe de recherche qui ont les autorisations nécessaires. L'accès à la plateforme est protégé par un mot de passe sécurisé, et les droits d'accès sont approuvés par le chercheur principal et attribués par Calcul Canada. Les données sensibles ne sont accessibles qu'aux chercheuses et chercheurs responsables de l'analyse.

Exemple 2

L'accès aux données de recherche sur OneDrive sera limité aux membres de l'équipe de recherche. L'utilisation de OneDrive minimise les efforts nécessaires au transfert des données. Toutes les manipulations dans les fichiers seront faites à même ceux qui sont sauvegardés dans le OneDrive pour éviter la multiplication des versions sur les espaces de travail des membres de l'équipe.

Ressource :

- [Section « Stockage et accès » du guide GDR](#)

4. Conservation

4.1 Lieux de conservation des données

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

Directives :

Les données que vous créez ne doivent pas nécessairement être conservées à long terme. Tenez compte de leur valeur, des exigences de financement, etc., puis décidez lesquelles, le cas échéant, devront être préservées.

Les décisions relatives à la **rétenion des données** peuvent être dictées par des politiques externes (p. ex. organismes de financement, éditeurs de revues, lois), par une compréhension de la **valeur durable** d'un ensemble de données déterminé et/ou dans le formulaire de consentement pour la participation à la recherche comme recommandé par le comité éthique.

Exemples :

Exemple 1

Pour l'instant, nous pensons utiliser le dépôt de données institutionnel (Dataverse de l'Université de Sherbrooke) pour déposer les données retenues et en assurer la conservation à long terme, la découverte et l'accès. Ce dépôt a plusieurs avantages importants pour nous :

- Dépôt sans frais pour les membres de l'institution,
- Attribution d'identifiants d'objets numériques (DOI) uniques et permanents,
- Capacité de restreindre l'accès aux données au niveau des fichiers,
- Mesures d'utilisation des données,
- Contrôle des versions des fichiers.

Exemple 2

Les données retenues seront rendues disponibles au cours du mois suivant leur publication dans le [Dépôt fédéré de données de recherche \(DFDR\)](#), celui-ci offrant un grand espace de stockage qui conviendra à nos gros jeux de données. Il permet aussi d'obtenir des identifiants persistants tels que les identifiants d'objets numériques (DOI) et est répertorié dans Lunarix ce qui augmentera la découverte de nos données.

Ressources :

Un soutien est offert à l'UdeS pour le dépôt [Dataverse de l'Université de Sherbrooke](#), qui répond aux exigences des organismes subventionnaires canadiens. Pour plus d'informations sur le choix du dépôt, contactez l'équipe GDR (gdr@usherbrooke.ca) ou consultez le [Guide GDR](#).

4.2 Préparation des données pour la conservation

Indiquez comment vous vous assurez que vos données sont prêtes pour la conservation. À prendre en considération : formats de fichier appropriés pour la conservation et qui préservent l'intégrité des données ; anonymisation et dépersonnalisation des fichiers, y compris les fichiers de documentation.

Directives :

Certains formats de données sont idéaux pour la conservation à long terme des données. Par exemple, les formats de fichier non propriétaires (ouverts), comme le format texte (TXT) et les valeurs séparées par des virgules (CSV), sont considérés comme des formats conçus pour la conservation.

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 vous assurer que les données que vous choisissez de conserver ou de partager sont anonymes, exemptes d'erreur et converties dans les formats recommandés en assurant un risque minimal de perte de données.

Exemple :

Pour faciliter l'interopérabilité, les données seront sauvegardées dans des formats logiciels en accès libre. Il s'agit par exemple de format en valeurs séparées par des virgules (CSV) pour les feuilles de calcul, de format de fichier d'image étiqueté (TIFF) pour les images, de format de fichier d'image étiqueté pour géoréférencement (GEOTIFF) pour les données spatiales et de format (TXT) pour le texte.

Ressources :

- [Section « Éléments d'évaluation et de préparation » du guide GDR](#)
- [Section « Données sensibles ou confidentielles » du guide GDR](#)

5. Partage et réutilisation

5.1 Choix des données et méthode de partage

Prévoyez-vous partager certaines de vos données? Si oui, quelles données partagerez-vous et sous quelle(s) forme(s)? Vos organismes subventionnaires ou vos éditeurs exigent-ils le partage de vos données ? Avez-vous besoin d'un embargo ou d'un délai avant de partager vos données ou métadonnées ? Comment allez-vous anonymiser vos données recueillies par sondages et entrevues ?

Directives :

Indiquez sous quelle forme (**brutes, traitées, analysées** ou **finale**s) les données seront partagées. Identifiez aussi les données qui ne peuvent pas être partagées et pourquoi (en raison de préoccupations relatives à la **confidentialité**, la **propriété intellectuelle**, etc.).

Exemples :

Exemple 1

Seules les métadonnées pourront être partagées en raison de l'identification possible des participantes et des participants. Les données brutes en format MP3 seront accessibles sur demande. Les personnes qui souhaiteront accéder aux données devront contacter la responsable du projet de recherche.

Exemple 2

Nous obtiendrons, avant le début de la collecte, le consentement des participantes et des participants pour partager les données récoltées lors des sondages et des entrevues. Nous déposerons et rendrons disponible un exemple de notre lettre d'information, du formulaire de consentement, ainsi que de notre demande de certification éthique.

1. Données de sondage : Nous prévoyons partager les données de sondage traitées en format CSV. Cela comprendra les réponses, mais tous les identifiants auront été supprimés préalablement.
2. Entrevues : Nous prévoyons partager les transcriptions traitées et anonymisées en format TXT.

Ressource :

- [Section « Déposer, partager et réutiliser des données » du guide GDR](#)

5.2 Licence d'utilisation

Sous quel type de licence souhaitez-vous permettre l'utilisation de vos données partagées ?

Directives :

Les **licences** déterminent les utilisations permises de vos données. Identifiez la licence qui correspond le mieux à votre recherche. Celle-ci devrait refléter votre volonté, les exigences des organismes externes, des revues et des dépôts de données, tout en favorisant autant que possible la libre réutilisation.

1. Il convient de noter que seulement les titulaires des droits de propriété intellectuelle peuvent émettre une licence, il est donc essentiel de préciser à qui appartiennent ces droits.
2. Il existe plusieurs types de licences standards mises à la disposition des chercheuses et des chercheurs comme les licences [Creative Commons](#) et les licences [Open Data Commons](#). Pour la plupart des ensembles de données, il est plus facile d'utiliser une licence standard plutôt que de concevoir une licence personnalisée.
3. 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.

Exemple :

Toutes les données utilisent les licences de données ouvertes CC-BY. Le chercheur principal s'assurera de bien avoir les droits de propriété des données diffusées.

Ressource :

- [Section « Critères et licences » du guide GDR](#)

5.3 Diffusion du partage

Quelles mesures prendrez-vous pour faire savoir aux autres chercheuses et chercheurs que vos données existent ?

Directives :

Décrivez la façon dont les jeux de données seront publicisés (ex. dépôt des données, enregistrement des données, leur **indexation**, bouche-à-oreille, ou des publications comme un **data paper**).

1. Le choix d'un **dépôt** qui attribuera un **identifiant constant** (comme un DOI) à votre ensemble de données permettra d'assurer un accès stable à l'ensemble de données, qui seront accessibles grâce à divers outils de recherche.
2. **Citer** vos ensembles de données de la même façon que d'autres types de publications (articles, livres, procédures, etc.) permettra leur **repérage**. 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.
3. Enfin, il existe aujourd'hui quelques dizaines de **data journals**, soit des revues savantes composées d'articles généralement évalués par les pairs présentant des jeux de données. Ce type d'articles est aussi de plus en plus courant dans des revues savantes traditionnelles.

Exemples :

Exemple 1

Les données déposées dans le dépôt fédéré de données de recherche (DFDR) ont des identifiants d'objets numériques (DOI) uniques, des mots clés et d'autres métadonnées interrogeables. Tous ces éléments seront rendus disponibles via le site Web du Groupe de recherche et les médias sociaux. Les membres de l'équipe de recherche citeront également les données dans leurs articles revus par les pairs.

Exemple 2

Nous présenterons le projet lors de conférences, publierons les résultats dans une revue révisée par des pairs avec des liens vers les données, et communiquerons l'existence des données par le biais des médias sociaux et du bouche-à-oreille. Nous espérons qu'un identifiant d'objet numérique (DOI) sera attribué aux données.

La citation des données déposées sera incluse dans la publication de la revue, et la citation de la publication sera incluse dans les données déposées afin que les informations puissent être découvertes des deux manières.

Ressources :

- [Section « Déposer, partager et réutiliser des données » du guide GDR](#)
- [Guide de Polytechnique Montréal à propos des *data papers* et *data journals*](#)

6. Responsabilités et ressources

6.1 Responsable de la GDR pour le projet

Qui sera responsable des tâches de GDR identifiées dans les sections précédentes ?

Directives :

Indiquez les éléments suivants pour chacune des étapes du projet :

1. Nom et affiliation
2. Rôle dans le projet
3. Tâches de GDR
4. Plan de relève en cas de retrait

Dans cette section, le niveau de détail dépendra de l'ampleur du projet. Il peut être possible de créer un tableau définissant en détail les tâches et le plan de relève ou de décrire le tout de façon concise pour respecter les limites de longueur des organismes subventionnaires.

Exemples :

Exemple 1 (projet complexe)

Bobby Smith, le chercheur principal du projet, aura la charge de la supervision du bon déroulement de la gestion des données de recherche. Il sera aussi le responsable de la collecte et Jeanne Côté, la co-chercheuse prendra le relai en cas de pépin. Les tâches d'organisation, de nomenclature et de documentation seront la responsabilité d'une personne assistante de recherche qui sera embauchée au moment de démarrer le projet. Comme nous engagerons plus d'une personne, nous établirons un système de relève au sein de cette équipe. C'est Céline Richard, technicienne en informatique pour notre laboratoire, qui s'occupera des tâches reliées au stockage, à la sauvegarde et à la protection des données sensibles.

La relève pourra être assuré par Jean Bourbon, technicien en informatique pour la faculté. Pour ce qui est de l'analyse, nous ferons affaires avec le Centre d'analyse de données en humanité numérique, la relève

sera ainsi assurée au sein de leur équipe. Un peu de la même façon nous ferons appel au Service des bibliothèques et archives pour tout ce qui touche à la conservation et au partage des données, une fois que le choix des données à conserver et/ou partager aura été fait par Jeanne Côté, qui sera appuyée par Michel Tremblay-Blais, étudiant au doctorant. Un tableau détaillant les rôles et les responsabilités de chacune des personnes sera intégré à la documentation du projet comme référence.

Exemple 2 (projet simple)

La chercheuse principale et son assistant seront les deux seules personnes qui auront accès, par mot de passe, aux questionnaires remplis par les participantes et participants, via la plateforme OneDrive, approuvée par la Direction de la sécurité de l'information (DSI). En outre, les réponses seront anonymisées dès leur réception par l'assistante de recherche, et l'ensemble du PGD et du projet feront l'objet d'une approbation par le Comité d'éthique à la recherche.

Ressource :

- [Section « Planifier » du guide GDR](#)

6.2 Évaluation des coûts

Quelles ressources vous faut-il pour mettre en œuvre le plan de gestion des données ? Combien cela pourrait-il coûter, incluant les coûts encourus après la fin du projet ?

Directives :

- a. Prévoir les coûts de gestion des données encourus au cours du projet
 1. licences des logiciels de collecte et d'analyse;
 2. des espaces de stockage payants;
 3. salaires des ressources humaines, incluant la formation du personnel;
 4. dépôt payant et services de curation, etc.
- b. Prévoir les coûts requis pour le soutien à plus long terme des données après la fin du projet
 1. préparation des données à déposer;
 2. 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.

Exemples de réponse :

Exemple 1

Le stockage des données (OneDrive) pendant la phase active du projet est fourni par l'établissement ; il n'y a pas de coûts supplémentaires. Comme nous envisageons de déposer les données dans un dépôt de données ouvert, nous ne prévoyons pas non plus de coûts directs pour la conservation des données à long terme.

Exemple 2

Les besoins en ressources que nous avons cernés sont les suivants : un gestionnaire de données de recherche spécialisé pour une période de 12 mois à raison de 3 jours semaine. Nous incluons ce poste dans notre demande de financement et notre budget, pour un montant global de 42 000 \$. Espace de l'environnement de recherche virtuel en nuage. Ce service sera fourni par Calcul Canada.

Ressource :

- Guide de OpenAIRE : [How to identify and assess Research Data Management \(RDM\) costs](#)

7. Conformité éthique, juridique et partenariale

7.1 Gestion des exigences éthiques et partenariales

Comment allez-vous gérer les exigences éthiques, juridiques et partenariales liées à vos données ?

Directives :

Inclure les éléments suivants dans votre réponse :

1. Qui est le chercheur ou la chercheuse responsable des données associées à ce projet de recherche ?
2. Quelles licences, conditions d'utilisation ou autres exigences y sont associées ?
3. Comment entendez-vous vous y conformer ?

Si vos recherches impliquent la participation de personnes des Premiers Peuples, assurez-vous de respecter les différents principes en place pour reconnaître et respecter leur autodétermination.

Exemple de réponse :

Exemple 1

Certains ensembles de données expérimentales issus du projet de recherche sont à l'origine d'une formule chimique faisant l'objet d'une demande de brevet. En l'attente de l'obtention du brevet, ces données feront l'objet d'un embargo de six mois.

Ressources :

- [Section « Données sensibles ou confidentielles » du guide GDR](#)
- [EPTC 2 \(2022\) – Chapitre 9 : Recherche impliquant les Premières Nations, les Inuits ou les Métis du Canada](#)
- [CARE Principles for Indigenous Data Governance](#)
- [Les principes de PCAP® des Premières Nations](#)

7.2 Gestion des données sensibles

Votre projet de recherche comprendra-t-il des données sensibles ? Si oui, comment allez-vous gérer ces données ?

Directives :

Si votre projet ne comprend pas de données sensibles, vous pouvez simplement l'indiquer et ignorer cette question. Si oui, intégrez les éléments suivants dans votre réponse :

- a. Méthode pour obtenir le **consentement éclairé** des participantes et des participants pour :
 1. la divulgation des résultats dépersonnalisés;
 2. la **réutilisation** des données (ex., par des personnes au doctorat) pour des recherches futures;
 3. tous ces mêmes éléments pour des études ultérieures.
- b. Moyens pour protéger des données sensibles durant leur vie active (accès par mot de passe, liste des personnes autorisées, stockage sécuritaire...), incluant les documents visuels, audio, imprimés, etc.;
- c. Techniques de **dépersonnalisation** ou d'anonymisation des données avant le dépôt final ;
- d. Date de **destruction des données** non anonymisées/dépersonnalisées ;
- e. Le dépôt de vos données fera-t-il l'objet d'un **embargo** pour des raisons éthiques ou commerciales ?

Exemple de réponse :

Dès le début du projet, l'équipe fera approuver son plan de projet par le comité d'éthique à la recherche.

Avant la collecte de données auprès des participantes et participants, il leur faudra signer un formulaire de consentement prévoyant le dépôt, le partage et la réutilisation par d'autres projets de recherche des données anonymisées résultant de leurs réponses au questionnaire. Durant la vie active du projet, une équipe restreinte (la chercheuse et son assistant) aura accès aux données non anonymisées, qui seront stockées de façon sécuritaire sur le OneDrive institutionnel et protégées par mot de passe.

Les renseignements identificatoires sur les participantes et les participants seront conservés dès le départ dans un dossier distinct de celui où seront conservées les réponses au formulaire. Avant le dépôt des données dans Dataverse, celles-ci seront anonymisées, et les données non anonymisées seront détruites.

Ressource :

- [Section « Données sensibles ou confidentielles » du guide GDR](#)