# Plan de gestion des données de recherche

# Modèle de l’École de technologie supérieure

Disponible sur l’[Assistant PGD](https://assistant.portagenetwork.ca/)

Bienvenue dans le plan de gestion de données (PGD) de l’École de technologie supérieure, fondé sur le modèle générique de PGD pour Portage.

Conformément aux exigences de l’IRSC, du CRSNG et du CRSH, cet outil est mis à votre disposition pour vous aider à expliquer, en amont de votre demande de financement, comment vous prévoyez gérer les données de recherches qui seront générées et utilisées dans le cadre de votre projet. Le PGD s’applique aussi aux projets non subventionnés.

Votre plan de gestion de données est un document évolutif qui devrait être mis à jour au fur et à mesure que des modifications seront apportées au projet.

Le [Décanat de la recherche](https://www.etsmtl.ca/recherche/partenariats/Decanat-recherche-equipe), le [service des technologies de l'information (STI)](https://www.etsmtl.ca/services/sti/accueil), le [Comité d’éthique de la recherche](https://www.etsmtl.ca/recherche/soutien-aux-chercheurs/comite-ethique-ets) et la [bibliothèque](https://bibliotheque.etsmtl.ca/Services/Communication-scientifique/Gestion-des-donnees-de-recherche) sont des sources d’information pertinentes. N’hésitez pas à les consulter au besoin.

Pour toute assistance technique concernant l’utilisation de l’outil (affichage des questions, téléchargement du plan, etc.) veuillez consulter la bibliothèque à l’adresse suivante : biblio.GDR@etsmtl.ca

## Détails du projet

### Titre du projet :

 Cliquez ou appuyez ici pour entrer du texte.

### Résumé du projet :

Cliquez ou appuyez ici pour entrer du texte.

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

Cliquez ou appuyez ici pour entrer du texte.

### Contributeurs :

Cliquez ou appuyez ici pour entrer du texte.

## Collecte de données

### Quels types de données recueillerez-vous, créerez-vous, couplerez-vous, acquérez-vous ou consignerez-vous?

Cliquez ou appuyez ici pour entrer du texte.

*Exemples :* données numériques, images, audio, vidéo, texte, données tabulaires, données de modélisation, données spatiales, données d'instrumentation

### Dans quels formats de fichiers vos données seront-elles collectées? Ces formats permettront-ils la réutilisation, le partage et l’accès à long terme aux données?

Cliquez ou appuyez ici pour entrer du texte.

*Directives*: 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. L'utilisation de formats de fichier ouverts ou de formats conformes aux normes de l'industrie (p. ex. les formats largement utilisés par une communauté donnée) est préférable dans la mesure du possible.

*Pour en savoir plus sur les formats de fichier, consultez les sites suivants :* [*UBC Library*](http://researchdata.library.ubc.ca/plan/format-your-data/) *ou* [*UK Data Archive*](https://www.ukdataservice.ac.uk/manage-data/format/file-formats.aspx)*.*

### 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?

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Il est important de garder la trace des multiples copies ou versions des fichiers, des différents fichiers conservés dans plus d'un format ou à divers emplacements et des renseignements qui font l'objet de renvois croisés entre les fichiers. Ce processus s'appelle « contrôle de versions ».

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 et de réduire la confusion concernant les copies enregistrées sur différents ordinateurs ou médias. Pour en savoir plus sur l'appellation des fichiers et le contrôle des versions, consultez les sites suivants : [UBC Library](http://researchdata.library.ubc.ca/plan/organize-your-data/) ou [UK Data Archive](https://www.ukdataservice.ac.uk/manage-data/format/versioning.aspx).

## Documentation et métadonnées

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Normalement, 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 chercheurs puissent utiliser les données. Parmi les autres éléments à documenter, le cas échéant, mentionnons la méthodologie de recherche utilisée, les définitions des variables, les vocabulaires, 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, l'explication de l'analyse et de la codification des données réalisées (y compris les fichiers de syntaxe) et les renseignements sur les personnes ayant travaillé au projet et réalisé chacune des tâches, etc.

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Examinez la façon dont vous saisirez ces renseignements et l'emplacement où ils seront enregistrés, idéalement avant de procéder à l'analyse et à la collecte de données, afin d'assurer l'exactitude, l'uniformité et l'exhaustivité des documents. Souvent, les ressources que vous avez déjà créées peuvent contribuer à cela (p. ex. publications, sites Web, rapports d'étape, etc.). Il est utile de consulter régulièrement les membres de l'équipe de recherche afin de connaître les changements possibles dans la collecte ou le traitement des données qui doivent être pris en considération dans les documents. La collecte des renseignements sur les données doit être une partie intégrante des responsabilités des membres de l'équipe et des opérations.

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Il existe plusieurs normes de métadonnées générales et propres à un domaine. Les informations sur l'ensemble de données doivent être fournies dans un de ces formats standards, ouverts et lisibles par machine afin de permettre l'échange efficace d'information entre les utilisateurs et les systèmes. Ces normes s'appuient souvent sur des formats de données comme XML, RDF et JSON qui ne sont pas liés à un langage de programmation. Il y a plusieurs normes de métadonnées fondées sur ces formats, y compris des normes propres à une discipline.

La documentation des jeux de données peut également inclure un vocabulaire contrôlé, qui est une liste terminologique normalisée pour la description de l’information. Parmi les exemples de vocabulaires contrôlés, citons les [Library of Congress Subject Headings (LCSH)](https://id.loc.gov/authorities/subjects.html) ou [NASA’s Global Change Master Directory (GCMD) Keywords](https://earthdata.nasa.gov/earth-observation-data/find-data/idn/gcmd-keywords).

Pour en savoir plus sur les normes de métadonnées, consultez le site suivant : [UK Digital Curation Centre's Disciplinary Metadata](http://dcc.ac.uk/resources/metadata-standards).

## Stockage et sauvegarde

### Quels sont les besoins prévus en matière de stockage pour votre projet (en mégaoctets, gigaoctets, téraoctets, etc.) et quelle sera la durée de sauvegarde?

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Les estimations concernant l'espace de stockage doivent prendre en considération les exigences pour le versionnage de fichiers, les sauvegardes et la croissance du nombre de fichiers au fil du temps.

Si vous recueillez des données sur une longue période (p. ex. plusieurs mois ou années), votre stratégie en matière de sauvegarde et de stockage des données doit tenir compte de la croissance des données. De même, un plan de stockage à long terme est nécessaire si vous avez l'intention de conserver vos données après le projet de recherche

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

Cliquez ou appuyez ici pour entrer du texte.

 *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 les [3-2-1 règles de sauvegarde](http://dataabinitio.com/?p=501) :

* Posséder au moins trois copies de vos données.
* Stocker les copies sur deux types de médias différents.
* Conserver une copie de sauvegarde hors site

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. De plus amples renseignements sur les pratiques de stockage et de sauvegarde sont disponibles sur les sites suivants : [University of Sheffield Library](https://www.sheffield.ac.uk/library/rdm/storage) and the [UK Data Service](https://www.ukdataservice.ac.uk/manage-data/store.aspx).

### 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?

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Une solution idéale est une solution qui facilite la collaboration, assure la sécurité des données et qui peut être adoptée par les utilisateurs en suivant une formation minimale. La transmission de données entre les différents espaces de travail ou au sein d'équipes de recherche n'est pas toujours simple selon l'infrastructure de gestion des données. Se fier au 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. Communiquez avec votre bibliothèque pour mettre au point la meilleure solution pour votre projet de recherche.

## Préservation

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : La question de la rétention des données doit être examinée tôt dans le cycle de vie de la recherche. Les décisions relatives à la rétention des données peuvent être dictées par des politiques externes (p. ex. 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.

Si vous avez besoin d'aide pour trouver une archive ou un dépôt de données approprié, communiquez avec votre bibliothèque. [re3data.org](http://www.re3data.org/) est un répertoire de dépôts de données ouvertes. Vérifiez si le dépôt de données pourra vous fournir une confirmation que les termes du dépôt énoncés dans votre plan de gestion des données seront respectés.

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : 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. Pour en savoir plus sur l'anonymisation, consultez les sites suivants :[UBC Library](https://researchdata.library.ubc.ca/share/anonymize-and-de-identify/) ou [UK Data Service](https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation.aspx).

## Partage et réutilisation

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

Cliquez ou appuyez ici pour entrer du texte.

*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.

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

Cliquez ou appuyez ici pour entrer du texte.

*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 l'établissement d'une licence. Une fois la licence créée, veuillez envisager d'inclure une copie de votre licence d'utilisation avec votre plan de gestion des données. Il convient de noter que seuls les titulaires des droits de propriété intellectuelle peuvent émettre une licence, il est dont essentiel de préciser à qui appartient ces droits.

Il existe plusieurs types de licences standards mises à la disposition des chercheurs comme les [licences Creative Commons](https://creativecommons.org/choose/?lang=fr) et les [licences Open Data Commons](http://opendatacommons.org/licenses/). 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. Pour en savoir plus sur les licences d'exploitation des données, consultez le site suivant : [UK Digital Curation Centre](http://www.dcc.ac.uk/resources/how-guides/license-research-data).

### Quelles mesures prendrez-vous pour faire savoir à la communauté des chercheurs que vos données existent?

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Les possibilités comprennent : registres de données, dépôts, index, bouche-à-oreille, publications.

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) à votre ensemble de données. Cela permettra ainsi d'assurer un accès stable à l'ensemble de données et il pourra être accessible grâce à 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). Le Digital Curation Centre fournit un [guide](http://www.dcc.ac.uk/resources/how-guides/cite-datasets) détaillé sur la citation des données. 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.

Communiquez avec votre bibliothèque pour obtenir de l'aide pour donner de la visibilité à votre ensemble de données et le rendre facilement accessible.

Réutilisé des NIH. (2009). [Key Elements to Consider in Preparing a Data Sharing Plan Under NIH Extramural Support (Principaux éléments à prendre en considération dans l'établissement d'un plan de partage des données en vertu du soutien extra-muros des NIH](https://grants.nih.gov/grants/sharing_key_elements_data_sharing_plan.pdf)). National Institutes of Health.

## Responsabilités et ressources

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives* : Votre plan de gestion des données a identifié les activités importantes relatives à la gestion des données dans le cadre de votre projet. Indiquez les responsables -- des personnes ou des organisations -- de l'exécution de ces parties de votre plan de gestion des données. Cela pourrait également comprendre le calendrier associé à ces responsabilités et toute formation nécessaire afin de préparer le personnel à assumer ces fonctions.

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

Cliquez ou appuyez ici pour entrer du texte.

*Directives :* 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 (p. ex. 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.

### 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?

Cliquez ou appuyez ici pour entrer du texte.

*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.

## Conformité éthique et juridique

### 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?

Cliquez ou appuyez ici pour entrer du texte.

*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.

Réutilisé de : DCC. (2013). [Checklist for a Data Management Plan (Liste de vérification pour un plan de gestion des données).](http://www.dcc.ac.uk/resources/data-management-plans) v.4.0. Edinburgh: Digital Curation Centre

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 en nuage comme Dropbox.

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

Cliquez ou appuyez ici pour entrer du texte.

*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.

Pour en savoir plus sur la sécurité des données, consultez le site suivant : [UK Data Service](https://www.ukdataservice.ac.uk/manage-data/store/security.aspx).

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

Cliquez ou appuyez ici pour entrer du texte.

*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. Le 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 mentionnés et respecter les exigences juridiques et éthiques pertinentes, le cas échéant (p. ex. consentement du sujet, permissions, restrictions, etc.).