

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Résumé

Ce modèle de PGD a été créé par le groupe de recherche en écohydrologie de l'université de Waterloo, dans le but de fournir un gabarit de PGD standardisé pour ses nombreux projets de recherche pluridisciplinaires et transversaux. Les processus fondamentaux du groupe de recherche en écohydrologie ont été intégrés comme point de départ pour le corps professoral, le personnel et le personnel hautement qualifié (PHQ) pour l'élaboration de projets de recherche. Cet exemple démontre l'utilité d'un modèle de PDG utilisé par un groupe ou un collectif de recherche pour maintenir les meilleures pratiques en matière de gestion des données.

### Détails administratifs

#### Nom du projet :

Plan de gestion des données du groupe de recherche en écohydrologie

#### Chercheurs principaux :

Philippe Van Cappellen (Orcid ID: <https://orcid.org/0000-0001-5476-0820> )

Fereidoun Reza Nezhad (Orcid ID: <https://orcid.org/0000-0002-9608-8005>)

#### Description:

Le [groupe de recherche en écohydrologie](#) (Ecohydrology Research Group – ERG) de l'Université de Waterloo mène des recherches fondamentales sur l'utilisation rationnelle des ressources en eau, c'est-à-dire de l'équilibre entre les besoins en eau de la société et ceux des écosystèmes naturels. Les activités de recherche de l'ERG couvrent une grande diversité d'échelles spatiales et temporelles, allant des études au niveau moléculaire sur les processus déterminant les formes chimiques et la biodisponibilité des nutriments et des polluants aux évaluations à l'échelle mondiale des perturbations anthropiques des cycles hydrologiques et biogéochimiques.

Les professeurs, le personnel de l'établissement et l'ensemble du personnel hautement qualifié (PHQ – soit étudiants, postdoctorants, chercheurs et associés) de l'ERG s'engagent à appliquer des pratiques de gestion des données afin de garantir que *les données, logiciels, codes, entrées et sorties (« les données »)* de l'ERG sont stockées en toute sécurité, conservées et

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

facilement accessibles pour une réutilisation ultérieure. L'ERG aligne son plan de gestion des données (PGD) sur l'approche commune adoptée par l'Université de Waterloo.

En consultation avec les superviseurs et les techniciens de laboratoire, le personnel hautement qualifié (PHQ) est censé adopter ce PGD dans la conduite du projet de recherche. Les membres du personnel peuvent également être invités à modifier certaines parties de leur projet de recherche en utilisant les lignes directrices fournies par le Réseau Portage ou pour répondre aux exigences des organismes de financement ou des contrats. Si, pour des membres du PHQ en particulier, on s'attend à des écarts importants par rapport au PGD de l'ERG, ces membres du PHQ intégreront un PGD individuel dans leur proposition de recherche, qui sera réexaminé et mis à jour au fur et à mesure de l'avancement de son projet de recherche, au besoin.

Si vous êtes membre d'un groupe de recherche utilisant un PGD partagé et standardisé, il est important de le mentionner dans le préambule et de fournir des détails sur les écarts que votre projet individuel a pris par rapport au plan partagé.

*Si vous êtes membre d'un groupe de recherche utilisant un PGD partagé et standardisé, il est important de le mentionner dans le préambule et de fournir des détails sur les écarts que votre projet individuel a pris par rapport au plan partagé..*

### **Établissement :**

Institut de l'eau de l'Université de Waterloo

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Collecte de données

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

Le Groupe de recherche en écohydrologie (ERG) génère des données à partir de documents existants, de travaux de terrain, d'expériences de laboratoire, de sources de données secondaires (y compris la télédétection) et de modélisation numérique. Les types de données comprennent des données de laboratoire, de modélisation, d'hydrologie, de chimie et de biologie (à l'heure, au jour, à la semaine, à la saison, à l'année) en rapport avec les objectifs de recherche de l'ERG. Les données supplémentaires comprennent les observations ou des images des sites, le lieu d'échantillonnage grâce aux coordonnées ou aux cartes du système de positionnement mondial (GPS), les données climatiques et les fichiers d'entrée de la modélisation, les paramètres et les fichiers de sortie. Les données générées sont utilisées dans la création, l'étalonnage et la validation des modèles. Les tâches de modélisation sont effectuées en même temps que la génération des données issues des expériences en laboratoire et des observations sur le terrain au sein de l'ERG. Cette approche nous permet d'écarter les données inacceptables et de répéter ou de redéfinir les expériences, au besoin.

#### **Selon quels formats de fichiers les données seront-elles recueillies? Ces formats permettront-ils la réutilisation, le partage et l'accès à long terme aux données?**

Le Groupe de recherche en écohydrologie (ERG) génère plusieurs types de données primaires en utilisant divers formats de fichiers propriétaires à partir des instruments ou des modèles numériques. Le personnel hautement qualifié (PHQ) est censé convertir ces fichiers dans des formats non propriétaires, tels que les valeurs séparées par des virgules (CSV) ou le texte (TXT), dans la mesure du possible, afin de faciliter leur réutilisation future.

*La définition des acronymes dans chaque nouvelle section fait en sorte que les lecteurs non familiarisés avec la terminologie n'aient pas à consulter les sections précédentes.*

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

**Quelles conventions et procédures utiliserez-vous pour structurer et nommer vos fichiers, et pour gérer leur version, afin de vous aider, ainsi que les autres, à mieux comprendre comment vos données sont organisées?**

Les conventions de nomenclature des répertoires et des fichiers sont établies au début d'un projet et peuvent intégrer des conventions spécifiques à la discipline pour les variables collectées. Des conventions de gestion de versions seront également établies. Les noms de fichiers comprennent le nom du projet (sous forme abrégée) et une brève description du contenu du fichier. Par exemple :

PX\_WP1\_WaterQuality\_Thames\_River

Les versions des documents doivent être nommées de manière séquentielle (avec des noms de fichiers se terminant par v1, v2, etc.). En voici un exemple :

WSP\_WaterQuality\_Thames\_River\_raw\_200617\_v1

Toutes les dates doivent respecter le [format de date ISO 8601](#) (AAAA-MM-JJ) et les noms des sites d'étude doivent respecter la [base de données toponymiques du Canada](#), le cas échéant.

*Les fichiers contrôlés permettent de montrer plus précisément comment vous allez organiser et comprendre vos données.*

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Quelles mesures d'assurance et de contrôle de la qualité des données utiliserez-vous?

*Cette section ne fait pas partie du modèle par défaut du PGD de Portage. Bien que les modèles ne soient pas adaptables dans l'Assistant PGD, le téléchargement de votre PGD sous forme de document Word permettra d'apporter des modifications supplémentaires pour répondre aux besoins de votre projet. Pour plus d'informations sur les modèles de l'Assistant PGD, veuillez contacter*

Tous les nouveaux chercheurs du Groupe de recherche en écohydrologie (ERG) sont formés quant aux protocoles normalisés d'assurance et de contrôle de la qualité (AQ/CQ) par des techniciens et du personnel expérimentés. Ils sont censés mettre en œuvre l'AQ/CQ pendant l'échantillonnage et l'analyse des échantillons. Pour les analyses chimiques et biologiques, le contrôle de la qualité comprendra des répliques, des blancs, des matériaux de référence certifiés et des pics. Dans la mesure du possible, une tâche spécifique de laboratoire doit être effectuée par le même étudiant, technicien ou chercheur pour l'assurance qualité. Les cahiers de données finales seront stockés dans un serveur de fichiers du dépôt du projet et contiendront des fiches techniques correspondant aux échantillons et aux normes, aux matériaux de référence certifiés (CRM), à la préparation, à la sortie brute des instruments, à la sortie manipulée (c'est-à-dire corrigée par digestion ou dilution) des instruments, aux résultats et aux limites de détection et de quantification communiquées.

### Documentation et métadonnées

#### Quels seront les documents nécessaires pour que les données soient lues et interprétées correctement à l'avenir?

Pour que les données soient lues et interprétées correctement, une documentation contextuelle est fournie, notamment des procédures d'exploitation standard (SOP) de laboratoire, des manuels d'utilisation des instruments et manuels d'utilisation des logiciels. La documentation sur les données est conforme aux meilleures pratiques nationales et internationales.

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Comment vous assurez-vous que la documentation est créée ou saisie de façon uniforme tout au long de votre projet?

Tous les chercheurs du Groupe de recherche en écohydrologie (ERG) collecteront des métadonnées et élaboreront une documentation appropriée dès le début du projet, avec des mises à jour régulières tout au long de la vie de leurs projets de recherche conformément aux pratiques développées par l'équipe de gestion des données de l'ERG. Par exemple, lors de la collecte des données de laboratoire et de terrain, toute la nomenclature chimique utilisée suivra les directives de l'[Union internationale de chimie pure et appliquée](#). On demande aux chercheurs de tenir des registres méticuleux des échantillonnages et des détails expérimentaux tels que la préparation des échantillons, les procédures d'échantillonnage, le temps, etc..

*L'intégration de liens dans votre PGD orientera les lecteurs vers des ressources externes qui peuvent fournir des informations plus détaillées sur des sujets liés à la discipline.*

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

Les métadonnées associées à tous les projets sont collectées pour faciliter la découverte, l'accès et la réutilisation futurs. Les données sont décrites selon le schéma de métadonnées [ISO 19115](#); des métadonnées supplémentaires sont collectées en fonction des besoins spécifiques du projet de recherche du personnel hautement qualifié (PHQ).

## Stockage et sauvegarde

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

Chaque membre du personnel hautement qualifiée (PHQ) devrait générer en moyenne 500 gigaoctets (Go) de données pendant le cycle de vie de la gestion des données; ainsi, toutes les données générées pour leur projet individuel doivent être stockées jusqu'à la publication de la thèse ou de l'article et l'attribution d'un identifiant d'objet numérique (DOI) à l'ensemble de données pour permettre la découverte et la reproduction de la recherche.

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Comment et où vos données seront-elles stockées et sauvegardées pendant votre projet de recherche?

Chaque membre du personnel hautement qualifié (PHQ) est censé suivre la méthode recommandée de sauvegarde et de stockage 3-2-1. Cela signifie que les membres du Groupe de recherche en écohydrologie (ERG) créeront trois copies de tous les fichiers de données, qui seront stockées sur deux types de supports différents, dont une copie sera conservée dans un lieu hors site.

*Lorsque l'on suit une méthode ou un protocole spécifique (comme la sauvegarde et le stockage 3-2-1), il est utile de fournir une brève explication sur la manière dont la méthode sera appliquée. Les lecteurs ne connaissent pas nécessairement toutes les pratiques ou lignes directrices de sauvegarde.*

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.

Le personnel hautement qualifié du Groupe de recherche en écohydrologie (ERG) décidera de la meilleure façon de sauvegarder ses données, pour autant qu'il respecte la règle 3-2-1 précédente. Toutefois, tous les membres du personnel hautement qualifié de l'ERG sont encouragés à fournir des copies de leurs fichiers pour les sauvegarder sur le serveur de l'ERG sur un nœud du service d'informatique de l'Université de Waterloo. Ce serveur est entretenu et sauvegardé régulièrement par le personnel de science informatique. Ces sauvegardes suivent un calendrier de rétention de 52 semaines qui est sauvegardé chaque semaine. Des options de stockage supplémentaires sont présentées sur le site web de l'[IST de l'Université de Waterloo](#).

### Comment l'équipe de recherche et les autres collaborateurs pourront-ils accéder aux données, les modifier et y contribuer tout au long du projet?

Les données de recherche active sont stockées par chaque membre du personnel hautement qualifié (PHQ) responsable de leur collecte jusqu'à ce qu'elles soient prêtes à être publiées/diffusées. Si les données doivent être partagées entre les chercheurs et les collaborateurs, l'équipe de gestion du Groupe de recherche en écohydrologie (ERG), avec l'approbation du chercheur principal, gèrera l'accès aux données et leur utilisation pendant la



## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

phase active du travail du projet. Les membres du personnel hautement qualifié utilisent la plateforme [Microsoft Teams](#) pour accéder aux données et les partager avec les collaborateurs. Microsoft Teams est une application de collaboration et de communication faisant partie de la suite Microsoft Office. Les fichiers de données volumineux peuvent être partagés via un lecteur portable.

### Conservation

#### Où déposerez-vous vos données pour la conservation à long terme et l'accès à la fin de votre projet de recherche?

Les données sont rendues disponibles au cours du mois suivant leur publication dans un dépôt fiable, tel que le [Dépôt fédéré de données de recherche](#) (DFDR), offrant des identifiants persistants tels que les identifiants d'objet numérique (DOI) ou les pseudonymes.

Conformément à la [Politique 73 de l'Université de Waterloo](#), tout le personnel hautement qualifié (PHQ) du Groupe de recherche en écohydrologie (ERG) doit également fournir une copie numérique de toutes les données de recherche à leurs superviseurs et collaborateurs respectifs à la fin de leur projet, thèse ou publication. Les données seront publiées sur le [Groupe de stockage de l'Institut de l'eau](#) de l'Université de Waterloo dans le DFDR et sur [DataStream](#) de la Fondation Gordon, le cas échéant.

#### Indiquez comment vous vous assurerez que vos données soient prêtes à être conservées. Envisagez des formats de fichiers favorables à la conservation, en assurant l'intégrité des fichiers et la dépersonnalisation, et l'inclusion des documents justificatifs.

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.



## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Partage et réutilisation

**Quelles données partagerez-vous et sous quelle forme les partagerez-vous (p. ex. brute, traitée, analysée ou définitive)?**

Les données finales utilisées dans les publications et les thèses seront partagées pour démontrer la reproductibilité des recherches du Groupe de recherche en écohydrologie (ERG).

**Avez-vous envisagé le type de licence d'utilisateur final à inclure dans vos données?**

Les données de recherche traitées qui donnent lieu à des publications seront rendues accessibles librement via une licence générique [licence Creative Commons Attribution 1.0](#) (CC BY 1.0 License). Si une licence plus restrictive est exigée, les chercheurs consulteront soit l'équipe de gestion du Groupe de recherche en écohydrologie (ERG) soit les [services de données de recherche de la bibliothèque de l'Université de Waterloo](#).

**Quelles mesures seront prises pour aide à faire connaître vos données dans le milieu de la recherche?**

Les données déposées dans le dépôt fédéré de données de recherche (DFDR) et le dépôt DataStream ont des identifiants d'objets numériques (DOI) uniques, des mots clés interrogeables et d'autres métadonnées interrogeables. Ces métadonnées avec liens de données seront rendues disponibles via le site web du Groupe de recherche en écohydrologie (ERG), les médias sociaux et l'[Index des métadonnées de Waterloo](#). Le personnel hautement qualifié (PHQ) de l'ERG citera également les données dans ses articles revus par les pairs. Des annonces par courrier électronique seront envoyées aux serveurs de listes de diffusion appropriés dès leur publication.

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

### Responsabilités et ressources

**Déterminez 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.**

Chaque membre du personnel hautement qualifié (PHQ) du Groupe de recherche en écohydrologie (ERG) est responsable de la gestion quotidienne des données de son projet. Leur gestion globale des données doit s'aligner sur les meilleures pratiques de gestion des données et sur le PGD de l'ERG. Cela comprend, notamment, les procédures standard d'assurance et de contrôle de la qualité (QA/QC), la compilation, le partage et la sauvegarde régulière des données et des métadonnées sur des serveurs sécurisés. Si plusieurs membres de l'équipe travaillent conjointement sur un projet de recherche spécifique, ceux-ci détermineront au début de leurs travaux quel membre est responsable de la mise en œuvre des pratiques de gestion des données pour leur projet. Les chercheurs principaux sont chargés de veiller à ce que les membres de l'équipe suivent le plan de gestion des données de l'ERG.

**Comment les responsabilités relatives à la gestion des activités de données seront-elles gérées si des changements importants surviennent dans le personnel qui supervise les données du projet, y compris un changement de chercheur principal?**

Les pratiques de gestion des données seront révisées chaque semestre par au moins deux membres de l'équipe de direction du Groupe de recherche en écohydrologie (ERG). Différents canaux de communication tels que la communication par courrier électronique, les réunions hebdomadaires de l'ERG et un canal de communication de [Microsoft Teams](#) serviront à rappeler ce plan à tous les chercheurs.

**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?**

Étant donné la taille du Groupe de recherche en écohydrologie (ERG), il serait nécessaire d'engager un gestionnaire de données à plein temps pour soutenir la gestion des données de recherche de l'ERG ; il faut espérer que ce type de soutien à la gestion des données de recherche (GDR) pourra être intégré dans les futures demandes de subvention. Pour l'instant, le personnel hautement qualifié (PHQ) de l'ERG devrait tirer parti des ressources en ligne

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

(telles que les ressources de [Mantra](#) et du [Réseau Portage](#)) ou demander conseil au gestionnaire de données du Global Water Future Program et aux [services de données de recherche de la bibliothèque de l'Université de Waterloo](#) afin de garantir le respect des meilleures pratiques en matière de gestion des données.

*Comme il s'agit d'un PGD conçu pour un groupe de recherche, il faut fournir des détails aux chercheurs sur ce qu'on attend d'eux pour gérer les données tout au long de leur projet.*

### Conformité aux lois et à l'éthique

**Si votre projet de recherche comprend des données sensibles, comment veillerez-vous à ce qu'elles soient gérées de façon sécuritaire et accessible uniquement aux membres approuvés du projet?**

Si les ensembles de données sont considérés comme sensibles, la gestion des données sera alors sous la responsabilité du chercheur principal. Les données seront soumises à des restrictions conformément aux meilleures pratiques et aux protocoles définis par l'[Office d'éthique de la recherche de l'Université de Waterloo](#). Seules les métadonnées appropriées et approuvées seront rendues publiques. On conseille fortement au personnel hautement qualifié (PHQ) possédant des données sensibles de faire en sorte que les fichiers soient cryptés..

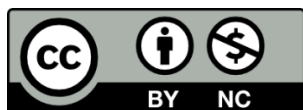
**Quelles stratégies entreprendrez-vous pour aborder les utilisations secondaires des données sensibles, le cas échéant?**

Les données de recherche active, ainsi que les données sensibles, peuvent être soumises à des restrictions conformément aux meilleures pratiques et aux protocoles décrits dans la [Politique 73 de l'Université de Waterloo](#), l'[Énoncé de politique des trois conseils : Éthique de la recherche avec des êtres humains](#) (EPTC 2) et les [Principes de PCAP® des Premières Nations](#). La disponibilité des données collectées auprès des intervenants externes sera soumise à des ententes conclues avec ces intervenants et le chercheur principal ou le superviseur du PHQ du projet.

## Plan de gestion des données du groupe de recherche en écohydrologie (v. 1.0)

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

Toutes les questions juridiques, éthiques et de propriété intellectuelle seront assujetties aux politiques pertinentes de l'Université de Waterloo et aux lois du Canada.



Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)

Portage Network | [portage@carl-abrc.ca](mailto:portage@carl-abrc.ca) | [portagenetwork.ca/fr](http://portagenetwork.ca/fr)