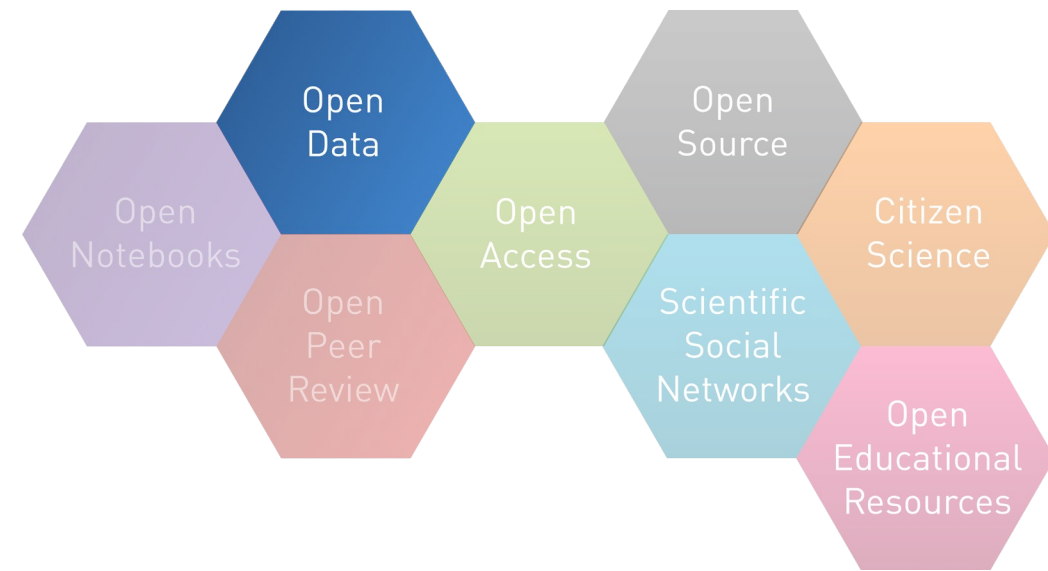




Élaborer une stratégie de Research Data Management

Une stratégie de Research Data Management ?

S'accorder sur la manière dont on gère les données

Documenter les pratiques de gestion des données à un niveau supérieur que celui des projets de recherche :
groupe de recherche, laboratoire, institut, GTHC, GTRC, institution

Bénéfices

Traitement concerté des données de recherche

Simplification de la transmission des données et de la documentation

Facilitation de la collaboration à l'intérieur et à l'extérieur du groupe

Guide d'introduction pour les nouvelles et nouveaux membres du groupe

Extraction d'informations pour un DMP de projet

À garder en tête

Il faut du **temps** pour établir une stratégie en Research Data Management

Elle doit être régulièrement **adaptée**

Il faut **tenir compte du cadre** institutionnel et légal, suisse et européen

Comment la mettre en œuvre ?

Drafting a Data Management Strategy –

Créée par Bucher Andres David, dernière modification par Schmid Fabian le sept. 29, 2023

 [Data Management Strategy Instructions for ETH Zurich Research Groups](#)

[\[About this page\]](#)

Overview

- [Abbreviations](#)
- [Exemplary Template](#)
- [What is a Research Data Management Strategy \(DMS\)?](#)
- [Implementation steps](#)
 - [1. Responsibilities](#)
 - [1.1 List of RDM tasks and responsibilities in your research g](#)
 - [1.2 Individual RDM responsibilities](#)
 - [1.3 Processes that need approval by the group leader](#)
 - [2. Data collection and documentation](#)
 - [2.1 List of scientific data and metadata](#)
 - [2.2 Research data ownership in contract agreements with th](#)
 - [2.3 File types and formats](#)
 - [2.4 File naming convention](#)
 - [2.5 Defined vocabularies](#)
 - [2.6 File folder organization](#)
 - [2.7 Data and code version control](#)

Comment la mettre en œuvre ?

FICHE PRATIQUE : faire un PLAN de GESTION de DONNÉES (PGD ou DMP) à l'échelle d'une structure (établissement, laboratoire, équipe de recherche)

3 QUESTIONS pour éclairer le sujet :

Qu'est-ce qu'un plan de gestion de données « structure » ?

i C'est un document de référence pour la structure (établissement, laboratoire, équipe) sur la gestion des données qu'elle produit dans le cadre de son activité de recherche.
Il est élaboré de manière collective, dans une démarche de **définition d'une politique de gestion des données de recherche** pour la structure.
Le « PGD structure » a un périmètre plus large que le « PGD projet », une durée de vie plus longue.

Que disent les textes officiels ?

” Les établissements publics et fondations reconnues d'utilité publique [...] définissent une politique de conservation, de communication et de réutilisation des résultats bruts des travaux scientifiques menés en son sein. A cet effet, ils veillent à la mise en œuvre par leur personnel de plans de gestion de données et contribuent aux infrastructures qui permettent la conservation, la communication et la réutilisation des données et des codes sources. ”

Décret n°2021-1572 du 03/12/2021 relatif au respect des exigences de l'intégrité scientifique, article 6

Pourquoi élaborer un plan de gestion de données structure ?

- 🌀 Pour harmoniser les pratiques de gestion/ouverture des données au sein de la structure
- 🌀 Pour monter en qualité sur la gestion des données
- 🌀 Pour faciliter la rédaction des « PGD projets » pour les chercheur·se·s et personnels de recherche, et respecter les exigences des financeurs de la recherche publique
- 🌀 Pour préserver le patrimoine scientifique de la structure :
 - conserver et valoriser les données produites au cours du temps
 - documenter les données produites
 - ne pas « perdre » les données produites par les doctorant·e·s à leur départ
- 🌀 Pour adopter une démarche « science ouverte » et respecter les principes FAIR
- 🌀 Pour faciliter la réutilisation des données de recherche : le 1^{er} ré-utilisateur potentiel est le·la chercheur·se à l'origine des données, ou bien ses collaborateur·rice·s

Par où COMMENCER ?

Identifier les données générées par la structure de recherche...

- Données d'observation ? (capturées en temps réel : imagerie médicale, données d'enquête, relevés de capteurs...)
- Données expérimentales ? (obtenues à partir d'équipements de laboratoire : chromatogrammes, séquençage ADN...)
- Données computationnelles ? (générées par des modèles informatiques ou de simulation)
- Données dérivées ou compilées ? (issues du traitement de données brutes : fouille de texte, bases de données compilées...)
- Données de référence ? (collection de jeux de données revus par les pairs et mis à disposition : GenBank, base de données de l'IGN, archives d'images historiques...)

... et en déduire les éventuels points de vigilance :

- Volumétrie importante ?
- Accès fréquents et/ou traitements intensifs sur les données ?
- Données à caractère personnel ? Voir données sensibles ? (données de santé, relatives à l'orientation sexuelle, politique ou syndicale, relatives à l'appartenance à un groupe ethnique, relatives à des infractions ou

Quelles données ?

Données d'observation - capturées en temps réel : imagerie, données d'enquête, relevés de capteurs...

Données expérimentales - obtenues à partir d'équipements de laboratoire : chromatogrammes, séquençage ADN...

Données computationnelles - générées par des modèles informatiques ou de simulation

Données dérivées ou compilées - issues du traitement de données brutes : fouille de texte, bases de données compilées...

Données de référence - collection de jeux de données revus par les pairs et mis à disposition : GenBank, base de données de l'IGN, archives d'images historiques...

Données jamais sans métadonnées

Table 2. List of scientific data and related metadata created by scientific methods that are frequently employed in our research group [example only, modify and replace as needed]

Research data type	Data type ID	Generation method	Method ID	File format	Collected metadata	Metadata standard	Data quality control
Interview data		Recordings			Date, participant's name or gathered Metadata is described in README file: [name]		
Seismic events		Seismographic recordings					
Microscopic 3D cell images		Laser scanning confocal microscopy					
...							

Convention de nommage

Identification plus aisée
des fichiers et/ou des données

Partage facilité

Maîtrise du contenu

3.4. Les éléments de nommage des dossiers

- Le classement des dossiers suit la structure du plan de classement.
- La numérotation des 1^{ers} niveaux de dossier est définie par le plan de classement et indiquée dans le référentiel de conservation. Les 4 premiers niveaux de classement sont fixes, les modifications sont effectuées seulement en collaboration avec le-la RGDA¹.
- Dans les sous-niveaux libres, l'intitulé identifie clairement l'affaire dont il s'agit.
- Ne pas reprendre les termes de la rubrique à laquelle le dossier est rattaché.
- Utiliser seulement l'année AAAA si mention de la date.
- L'année de clôture du dossier peut également être indiquée.
Ex : 2010-2018_[NomDossier]
- Les dossiers peuvent être ordonnancés par ordre alphabétique ou à l'aide d'une numérotation.
- A la clôture du dossier, on peut identifier ce dernier par un _C (= clos), ajouté au début ou à la fin du nom.
Ex : 2010-2018_[NomDossier]_C ou C_2010-2018_[NomDossier]

3.5. Les éléments de nommage des fichiers

Elément	Règle	Exemple
Domaine administratif	Recommandé ○ Utiliser une abréviation des organes, entités, etc. qui doit être reprise et insérée en tête des noms de fichiers, sans les caractères spéciaux (<u>Dénominations et abréviations des principaux organes, entités, instances, fonctions et rôles</u>) ○ Cas des documents reçus : il est recommandé de renommer un fichier provenant de l'extérieur de la HEIG-VD et de préciser l'origine du fichier : soit en précédant le nommage de l'abréviation « EXT », soit en indiquant le service créateur grâce à son acronyme	DIR (Direction) CORE (Conseil représentatif) GTHC (Groupe thématique de compétences) EXT_HESSO_PV_20240118.docx HESSO_PV_20240118.docx
Sujet	Obligatoire ○ Sujet principal traité au sein du document, sans accent	Formation Projet
Type de document	Obligatoire ○ Qualifie la nature du document : utiliser des abréviations reconnues pour réduire le nombre de caractères (cf. Point 6)	CdC (cahier des charges) PV (procès-verbal) NTSEAN (Notes de séance)
Date (fichier)	Obligatoire ○ Date de création du document ou de l'événement : AAAA au minimum	2024 20240118 2024-01-18 202401_03 (de janvier à mars 2024)

Ouvrir les données – Open research data

Définir **quelle part** des données de recherche doit être partagée

Suivre les **conventions ou recommandations** spécifiques à votre discipline ou aux exigences des bailleurs de fonds

Tenir compte des **exigences légales** et des niveaux de **confidentialité**

Et plus encore...

Comment les dossiers sont organisés ?

Qui a quelle responsabilité par-rapport à la gestion des données ?

Qui a accès aux données ?

etc.

Vous souhaitez

Une aide personnalisée pour votre publication / DMP

Une formation spécifique à un ou plusieurs outils

...

biblio@heig-vd.ch
marylene.grzesiak@heig-vd.ch

Site web de la bibliothèque
[Soutien à la recherche](#)

[Empruntez votre bibliothécaire](#)

Lundi au jeudi 8h - 18h30

Vendredi 8h - 17h

Formations complémentaires

Recherche et intelligence
artificielle

Utiliser les licences Creative
Common

Introduction à
l'Open Access

Et plus encore...

Utiliser ORCID

Déposer un set de données sur
Zenodo

Licence de la présentation

A l'exception des slides qui mentionnent d'autres auteurs·trices

Bibliothèque de la HEIG-VD, 6 juin 2024

Ce document est sous licence Creative Commons Attribution-ShareAlike 4.0 International <https://creativecommons.org/licenses/by-sa/4.0/>



Cette formation est librement adaptée de :
DEDERKE, J., & SCHMID, F., 2023. *Getting started with a Data Management Strategy for your Research Group* [en ligne]. Consulté le 6 juin 2024.
Disponible à l'adresse : <https://doi.org/10.5281/zenodo.10402122>

Icônes by [Freepik – Flaticon](#)