

# Gérer ses données et son matériel de recherche

# Les histoires d'horreur de la recherche

## Millions de sites affectés

L'incendie d'origine accidentelle, qui a détruit l'un de ses quatre centres de données à Strasbourg, a eu des conséquences majeures. OVHcloud estime qu'entre 12 000 et 16 000 de ses clients ont été touchés. Mais comme certains clients de la firme française sont eux-mêmes des hébergeurs, le nombre de victimes a été beaucoup plus grand: environ 3,6 millions de sites web sont devenus inaccessibles, selon la société de recherche anglaise Netcraft, soit 18% des adresses IP attribuées à OVHcloud. Le portail du Centre Pompidou ou la plateforme d'accès aux données publiques data.gouv.fr ont été brièvement déconnectés. Mais certains clients d'OVHcloud, tels que le studio britannique de jeux vidéo Facepunch ou l'application française Meteociel, ont affirmé la perte définitive de données.

Les flammes durant la nuit  
du 10 mars.  
— © REUTERS



# Les histoires d'horreur de la recherche

Courrier International, 2024. Sciences : en Suède, des décennies de recherche perdues à cause d'une panne de congélateur. In : *Courrier International* [en ligne]. 6 février 2024. [Consulté le 12 février 2024]. Disponible à l'adresse : <https://www.courrierinternational.com/article/sciences-en-suede-des-decennies-de-recherches-perdues-a-cause-d-une-panne-de-congelateur>

SCIENCES ET ENVIRONNEMENT • SUÈDE • SCIENCES & TECHNO

## Sciences. En Suède, des décennies de recherches perdues à cause d'une panne de congélateur

Lorsqu'un congélateur est tombé en panne au sein de l'institut suédois Karolinska, un nombre important d'échantillons de recherche recueillis au fil des décennies ont été perdus. Le "Corriere della Sera" se demande si ce type d'incidents est amené à devenir plus fréquent.



SOURCE  
Courrier International

Publié le 06 février 2024 à 18h24 • Lecture 1 min.



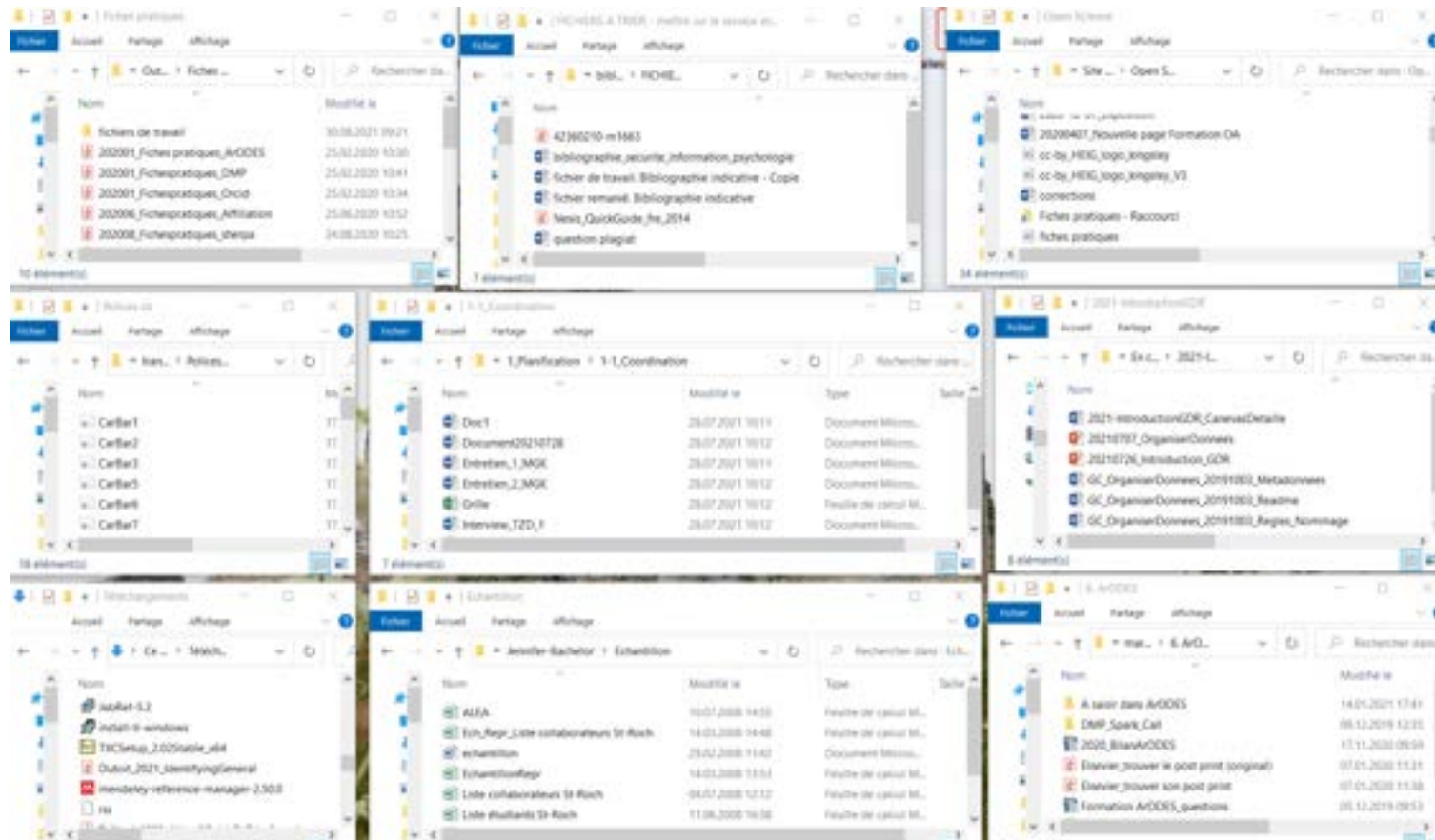
### Les plus lus

- 1 **Verbatim.** Trump : "J'encouragerais" la Russie "à faire ce qui lui chante" contre un allié de l'Otan ne payant pas assez
- 2 **Décryptage.** En Russie, le succès à double tranchant du film "Le Maître et Marguerite"
- 3 **Vu d'Espagne.** L'Europe s'attaque aux imposteurs "verts"
- 4 **Vatican.** Javier Milei et François : des injures à l'accolade avec le pape argentin

# Les histoires d'horreur de la recherche



# Les histoires d'horreur de la recherche



CYBERCRIMINALITÉ

## Une cyberattaque met à genoux l'Université de Neuchâtel



L'institution a été victime jeudi soir de ce qui semble être un rançongiciel. Ses services informatiques travaillent d'arrache-pied pour restaurer ses systèmes avant la rentrée de lundi

## Les histoires d'horreur de la recherche



Le bâtiment principal de l'Université de Neuchâtel est photographié avant une conférence de presse sur les résultats de deux enquêtes administratives ouvertes par le Conseil d'Etat de NE pour soupçon de plagiat et soupçon de mobbing au sein de la Faculté -- © JEAN-CHRISTOPHE BOTT / KEYSTONE

STEINER, Alexandre, 2022. Une cyberattaque met à genoux l'Université de Neuchâtel. In : *Le temps* [en ligne]. 19 février 2022. [Consulté le 3 mars 2022]. Disponible à l'adresse : <https://www.letemps.ch/suisse/une-cyberattaque-met-genoux-luniversite-neuchatel>

À lire aussi : ZAÏBI, Sami, 2022. Cyberattaque : l'Université de Neuchâtel toujours paralysée. In : *Le temps* [en ligne]. 21 février 2022. [Consulté le 3 mars 2022]. Disponible à l'adresse :

<https://www.letemps.ch/suisse/cyberattaque-luniversite-neuchatel-toujours-paralysee>

# Le pire est survenu: les données volées à l'Université de Neuchâtel ont été publiées



Salaires, photos d'étudiants ou encore contrats de travail ont été mis en ligne par les pirates sur le darknet après la cyberattaque intervenue mi-février

## Les histoires d'horreur de la recherche



Lors d'une rentrée de l'Université de Neuchâtel, le lundi 14 septembre 2020. — © LAURENT GILLIERON / keystone-sda.ch

## Les histoires d'horreur de la recherche

### EXCLUSIF – Le piratage de l'Université de Neuchâtel est tentaculaire



D'après nos recherches, l'étendue des données volées à l'Université de Neuchâtel est pire que révélé précédemment et concerne aussi Genève, la Confédération, Fedpol ou encore des multinationales. Les universités protègent-elles assez leurs informations sensibles?

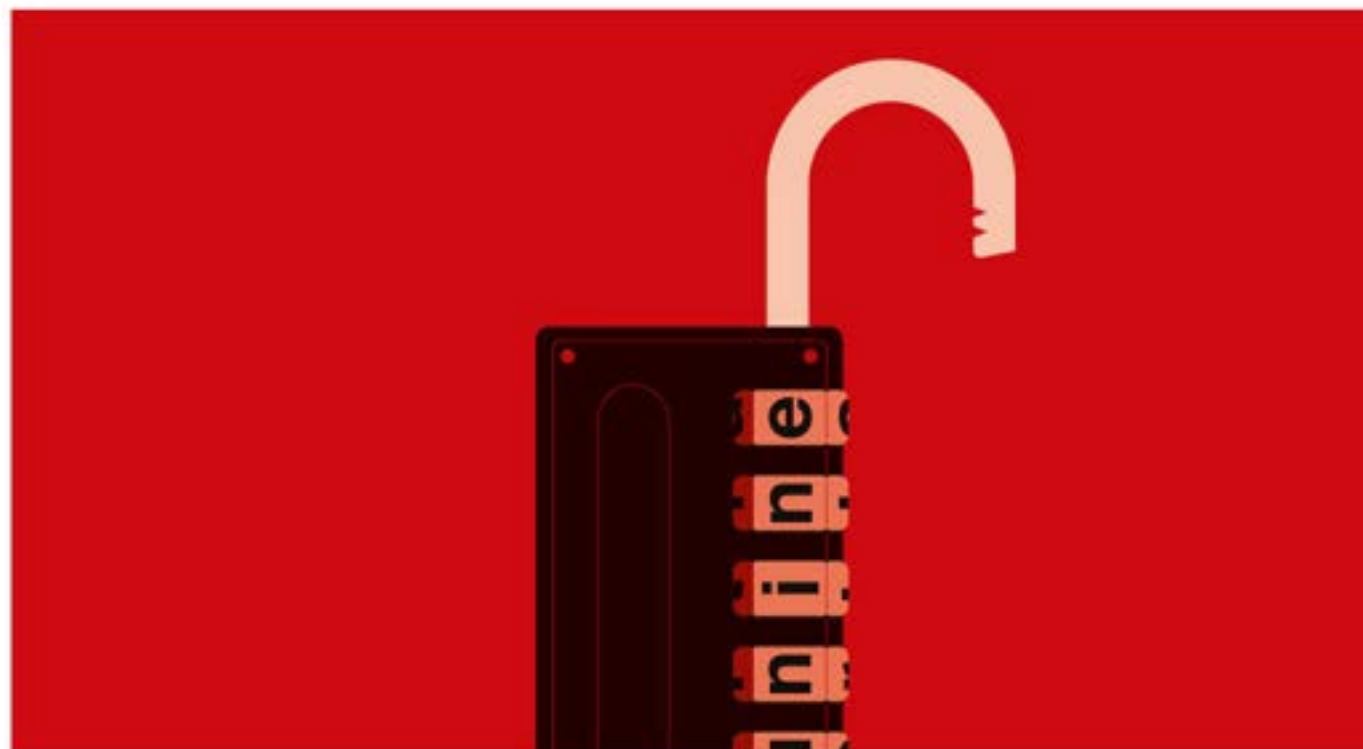


Illustration: Olivier Ploux pour Le Temps

SEYDTAGHIA, Anouch, 2022. Le piratage de l'Université de Neuchâtel est tentaculaire. In : *Le temps* [en ligne]. 9 mars 2022. [Consulté le 27 février 2023]. Disponible à l'adresse :

<https://www.letemps.ch/economie/cyber/exclusif-piratage-luniversite-neuchatel-tentaculaire>



# Les histoires d'horreur de la recherche

PERIAGE

## L'université de Zurich visée par une cyberattaque «sérieuse»

L'université a indiqué jeudi avoir renforcé ses mesures de protection. Aucune donnée sensible n'aurait été dérobée



Le bâtiment principal de l'Université de Zurich, en juin 2018 - © CHRISTIAN H

## La Haute Ecole Arc victime d'une cyberattaque

L'établissement annonce lundi être la cible de pirates informatiques. Des mesures ont été prises pour protéger ses infrastructures. Si des demandes de rançon ont été formulées, il n'est pas certain que des données aient été chiffrées, selon la directrice générale Brigitte Bachelard



Le bâtiment du campus recherche/lets de la Haute Ecole Arc (HE-Arc) en 2018 - © VALENTIN ELACRACD / KEYSTONE

## Les CFF ont été victimes d'une cyberattaque



CFF ont été victimes d'une cyberattaque / Le Journal horaire / 17 sec. / le 8 février 2023

# De l'art du rangement



# De l'art du rangement



# Avantages de la gestion des données



Organiser et faciliter le travail de recherche



Assurer la continuité des projets



Garantir leur cohérence et leur intégrité



Faciliter le partage et la réutilisation des données



Diffuser les résultats et faire connaître votre travail

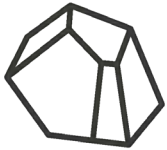


Permettre leur archivage

## Les données de la recherche sont...

« définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche. »

## Données brutes vs données traitées



Données brutes : données recueillies qui n'ont pas été modifiées, organisées ou mises en forme depuis l'acquisition

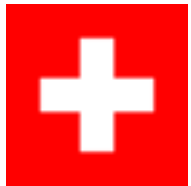
→ verbatims d'entretien



Données traitées : données brutes éditées, nettoyées ou modifiées, corrigées, réorganisées

→ verbatims après anonymisation

## Exigences des bailleurs de fonds



[FNS](#) « estime que le partage des données de recherche apporte une contribution essentielle à la recherche scientifique en termes d'impact, de transparence et de reproductibilité. En plus d'une préparation et d'un archivage soigneux, le FNS est d'avis que **les données de recherche doivent être partagées de manière aussi ouverte que possible.** »



Horizon Europe « les bénéficiaires rendent leurs données de recherche trouvables, accessibles, interopérables et réutilisables (FAIR) afin de garantir leur bonne gestion. Ce n'est pas un objectif en soi, mais plutôt un moyen essentiel permettant la découverte des connaissances et l'innovation, puis l'intégration et la réutilisation des données et des connaissances. »

[H2020 online manual - Data management](#)

# Politique des éditeurs scientifiques



## [PLOS Data Availability Policy](#)

**PLOS journals require authors to make all data necessary to replicate their study's findings publicly available without restriction at the time of publication. When specific legal or ethical restrictions prohibit public sharing of a data set, authors must indicate how others may obtain access to the data.**

When submitting a manuscript, authors must provide a Data Availability Statement describing compliance with PLOS' data policy. If the article is accepted for publication, the Data Availability Statement will be published as part of the article.

Acceptable data sharing methods are listed below, accompanied by guidance for authors as to what must be included in their Data Availability Statement and how to follow [best practices in research reporting](#).

PLOS believes that sharing data fosters scientific progress. Data availability allows and facilitates:

- Validation, replication, reanalysis, new analysis, reinterpretation or inclusion into meta-analyses;
- Reproducibility of research;
- Efforts to ensure data are archived, increasing the value of the investment made in funding scientific research;
- Reduction of the burden on authors in preserving and finding old data, and managing data access requests;
- Citation and linking of research data and their associated articles, enhancing visibility and ensuring recognition for authors, data producers and curators.



# Cadre légal et réglementaire

Loi 419.01 sur les hautes écoles vaudoises de type HES (LHEV)

[Art. 62 « Propriété intellectuelle »](#)

<sup>1</sup> A l'exception des droits d'auteur et des droits voisins, chaque haute école est titulaire des droits de propriété intellectuelle portant sur toute création intellectuelle technique ainsi que sur des résultats de recherche obtenus par les membres de son personnel dans l'exercice de leurs activités au service de la haute école. Sont réservés les accords comportant des clauses de cession ou de licence en faveur de tiers ayant financé partiellement ou totalement les recherches.

<sup>2</sup> La gestion et l'éventuelle cession des brevets sont assurées par chaque haute école.

<sup>3</sup> Les droits exclusifs d'utilisation des programmes informatiques créés par le personnel de chaque haute école dans l'exercice de ses activités au sein de la haute école reviennent à cette dernière.

# Cadre légal et réglementaire

Canton de Vaud

[Loi sur la protection des données personnelles](#) (LPrD)

Suisse

[Loi fédérale sur la protection des données](#) (LPD)

[Loi fédérale relative à la recherche sur l'être humain](#) (LRH)

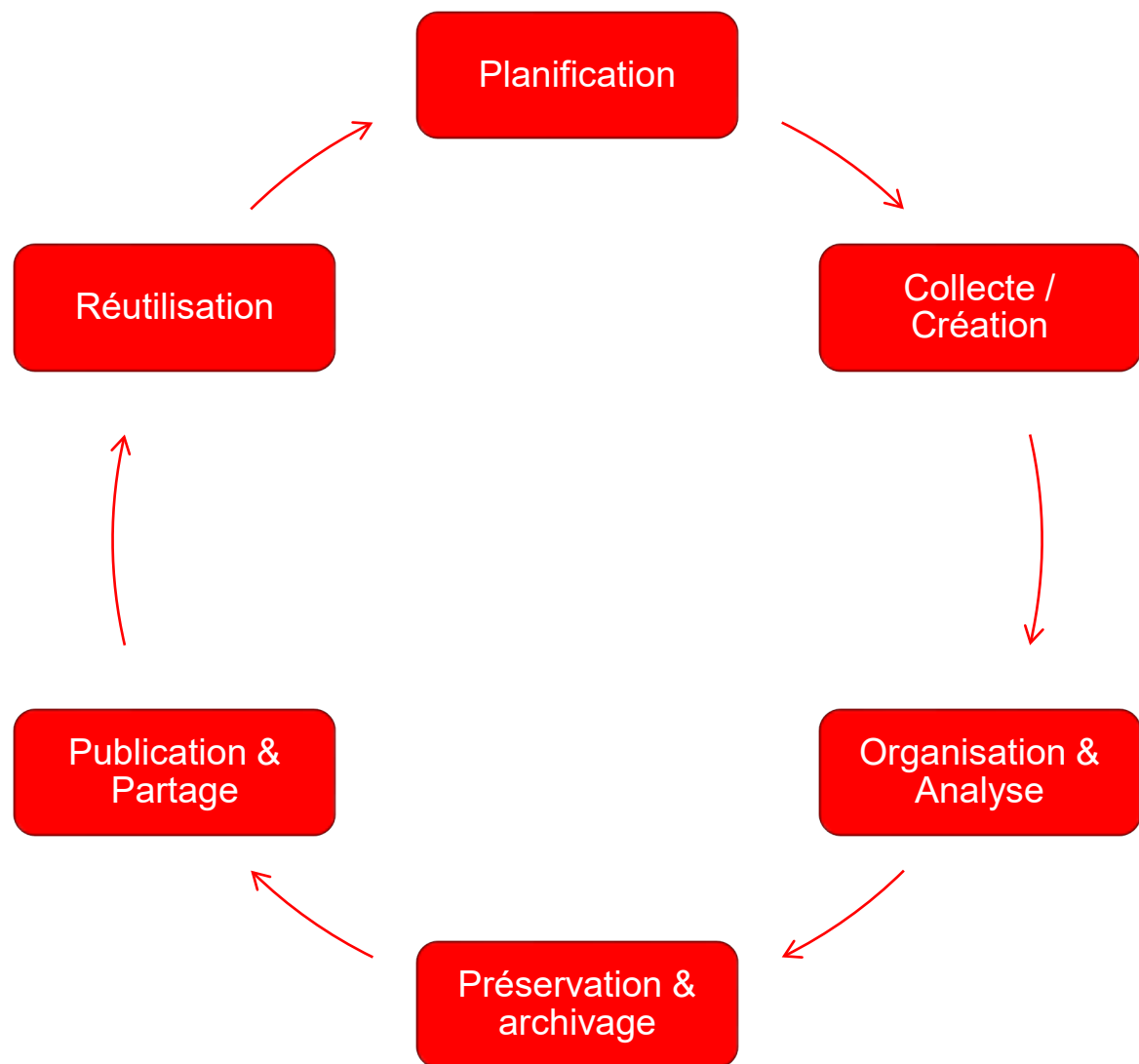
Europe (RGPD)

[Règlement européen sur la protection des données](#)

→ Définitions :

Données personnelles	Données personnelles liées à la santé
Données sensibles	...
Traitement des données	

# Le cycle de vie des données de recherche



# Planification

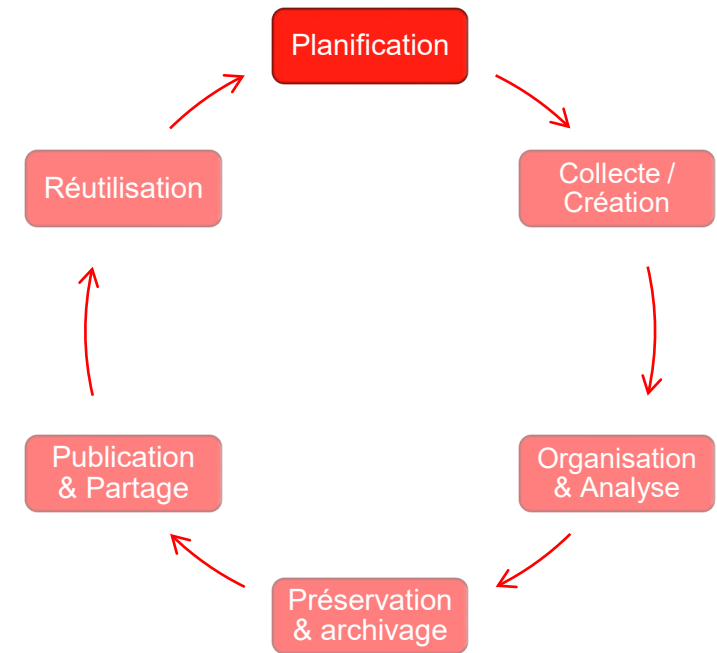
Comment planifiez-vous la gestion de votre projet ?

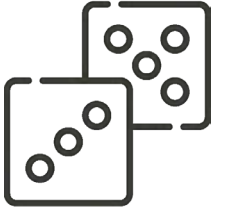
Quelle documentation existe pour le projet ?

Mentionnez-vous les données ?

Avez-vous nommé un responsable des données ?

Vous arrive-t-il de dire que vous auriez dû « y penser avant » ?





**Comment se nomme le document qui prépare l'organisation, le partage, la réutilisation et la pérennisation des données ?**

- A. Consensus de partage des données
- B. Data management plan ou Plan de gestion des données
- C. Aucune de ces réponses

# Data management plan (DMP) ou plan de gestion des données

## Everyone needs a data-management plan

They sound dull, but data-management plans are essential, and funders must explain why.



Keep your research data organized with a management plan. Credit: Jasper Juinen/Bloomberg/Getty

# Avantages

Amorcer la réflexion sur ses pratiques de recherche

Planifier la gestion de ses données de recherche

Anticiper certaines problématiques spécifiques

Évaluer les coûts et ressources nécessaires

Préparer son projet en conséquence





# Data Management Plan du FNS

## 1. Collecte des données et documentation

- ☑ **I** 1.1 Quelles données allez-vous collecter, étudier, générer ou réutiliser ?
- ☑ **I** 1.2 Comment les données seront-elles recueillies, étudiées ou générées ?
- ☑ **I** 1.3 Quelle documentation et quelles métadonnées allez-vous fournir avec les données ?

## 2. Questions éthiques, légales et de sécurité

- ☑ **I** 2.1 Comment les questions éthiques seront-elles abordées et traitées ?
- ☑ **I** 2.2 Comment seront gérés l'accès aux données et la sécurité ?
- ☑ **I** 2.3 Comment allez-vous gérer les questions de droits d'auteur et de propriété intellectuelle ?

## 3. Stockage et préservation des données

- ☑ **I** 3.1 De quelle manière vos données seront-elles stockées et sauvegardées au cours de la recherche ?
- ☑ **I** 3.2 Quel est votre plan en matière de conservation des données ?

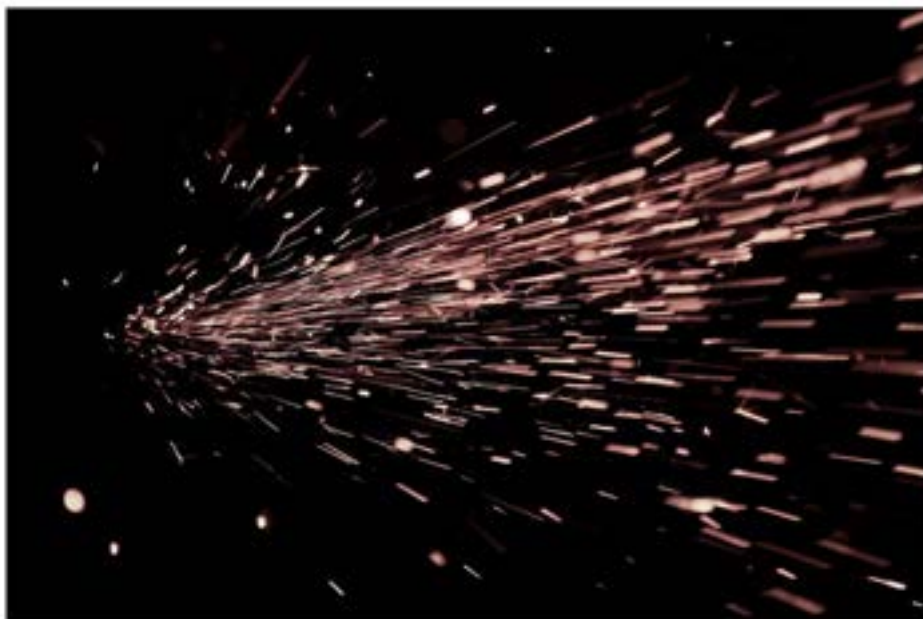
## 4. Partage et réutilisation des données

- ☑ **I** 4.1 De quelle manière et où seront partagées les données ?
- ☑ **I** 4.2 Y a-t-il des restrictions nécessaires pour protéger les données sensibles ?
- ☑ **I** 4.3 Je confirme que je choisirai exclusivement des bases de données (repositories) numériques conformes aux FAIR Data Principles
- ☑ **I** 4.4 Les bases de données (repositories) choisies pour le dépôt des données sont gérées par une organisation à but non lucratif.

# Open Research Data



## Spark



Financement rapide d'idées novatrices et originales

qu'aucune clause juridique, éthique, concernant le droit d'auteur ou autre ne s'y oppose. Un plan de gestion de données (Data Management Plan ou DMP) est demandé pour les projets approuvés.

## BRIDGE Discovery



aient être libres d'a  
ur le public.

Délai de soumission: TBD

Innosuisse et le FNS lancent la cinquième mise au concours de l'instrument BRIDGE Discovery.

en outre à ce que les données produites au cours des travaux  
dépôts de données numériques en libre accès, pour autant



Politique

Efficacité

Approvisionnement

Mesures  
d'encouragement

Recherche et  
cleantech

Actualités et médias

L'OFEN

[Page d'accueil](#) > [Recherche et cleantech](#) > [Programme d'encouragement SWEET](#)

[← Recherche et cleantech](#)

## Programme d'encouragement SWEET

[Aperçu de SWEET](#)

[Aperçu des appels d'offres](#)

[Actualités](#)



## Programme d'encouragement SWEET

SWEET – «SWiss Energy research for the Energy Transition» – est un programme d'encouragement de l'Office fédéral de l'énergie (OFEN). Il vise à promouvoir les innovations nécessaires à la mise en œuvre de la Stratégie énergétique 2050 et à la réalisation des objectifs climatiques de la Suisse.

## Domaine I&A

Dans le modèle « demande de financement »

« Avez vous prévu de réaliser un Data Management Plan (DMP) [...] Dans le cas inverse, merci d'indiquer ci dessous les raisons ayant mené à ce refus. »

# Collecte / création

Réutilisez-vous des données ?

Avez-vous contrôlé l'intégrité des données ?

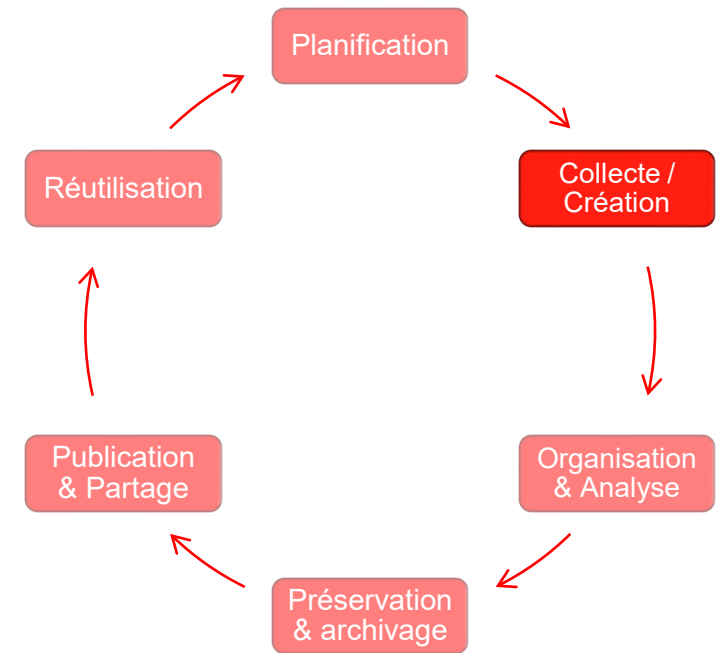
Les fichiers sont-ils lisibles ?

Est-ce qu'il y a de la documentation pour les comprendre ?

Comment collectez-vous les données ?

Comment sont-elles transcrites ?

Quel contrôle d'intégrité est mis en place ?



## Collecte / création

S'assurer de l'intégrité de ses données permet d'obtenir des résultats reproductibles

Avoir des protocoles de test, des tests de validité, etc.

S'assurer de la cohérence pour la saisie entre plusieurs personnes

Et plus encore

## Collecte & création

Choisir des formats de fichiers appropriés

Ouverts → non propriétaire

Standards, normalisés

Non compressés

Non chiffrés

# Formats de fichier

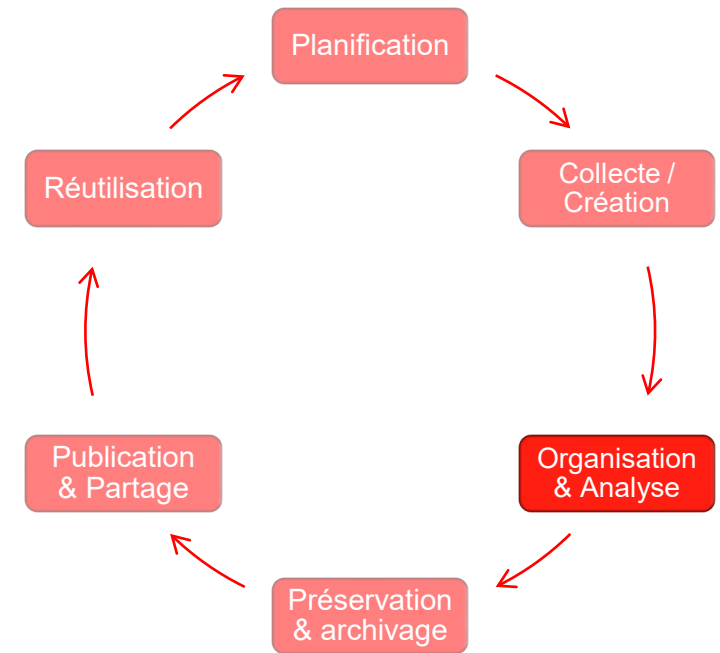
Types de données	Formats recommandés	Format à éviter
Texte	.txt	.docx, .doc, .rtf
Données tabulaires	.csv, .tab	.xlsx, .xls
Images	.tiff, .png	.jpg, .psd
Audio	.wav	.mp3/4
Vidéo	.mp4, .avi	.wmv, .mov



## Organisation & documentation

Est-ce que je dois souvent chercher un fichier, ou m'assurer que j'ai bien la dernière version ?

Est-ce que la structure de rangement de mes données et la dénomination de mes fichiers seront comprises par un·e collègue dans une année ?



# Organisation & documentation

What your research supposedly looks like:

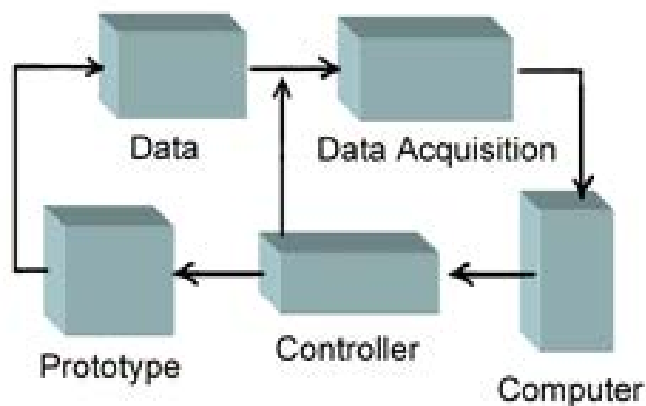


Figure 1. Experimental Diagram

What your research *actually* looks like:

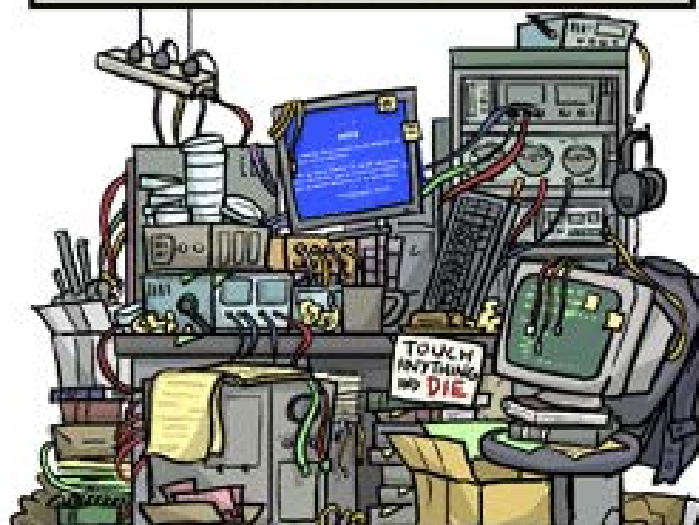


Figure 2. Experimental Mess

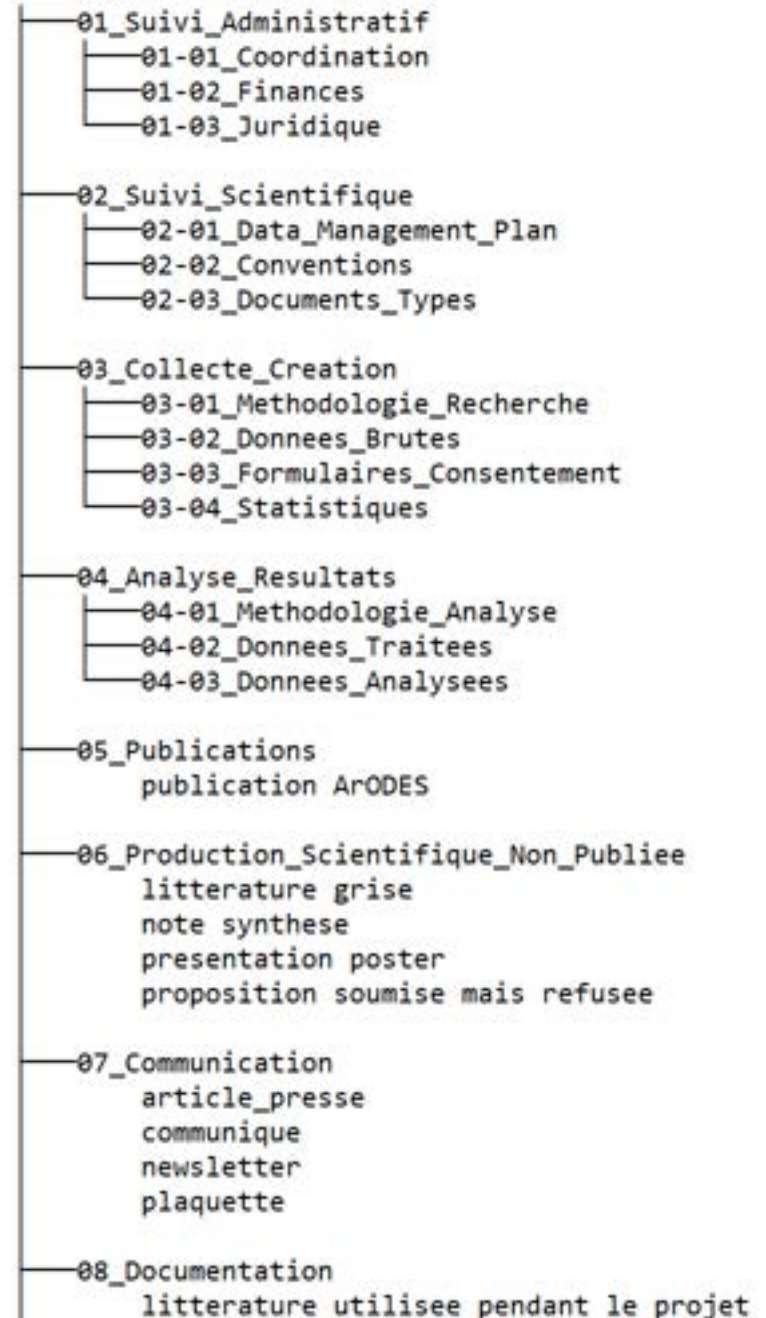
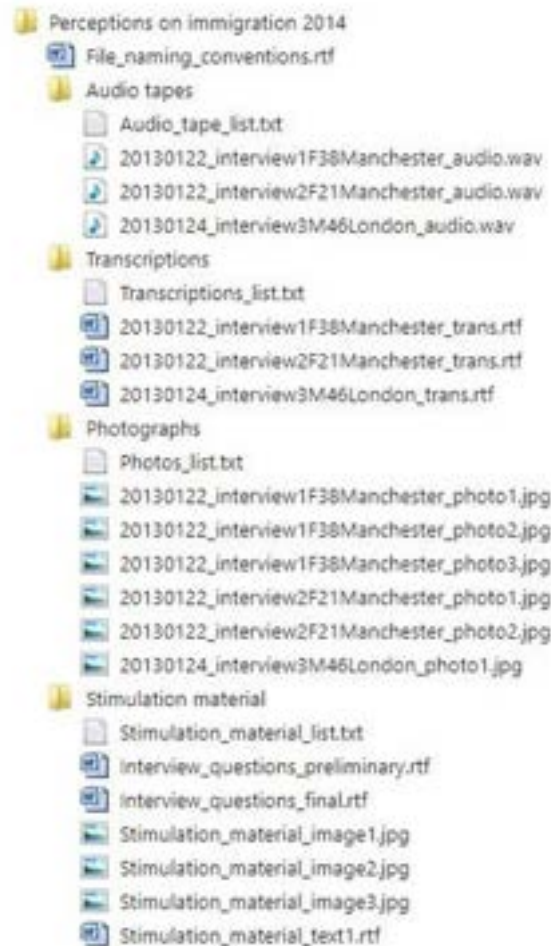
WWW.PHDCOMICS.COM JORGE CHAM © 2008

## Organisation & documentation

- ➔ Investir du temps au début pour en gagner plus tard
  
- ➔ Se concentrer sur la recherche et pas sur l'administration

# Organisation & documentation

Utiliser une arborescence classificatoire pour les dossiers et fichiers



La bibliothèque de la HEIG-VD propose une arborescence pour classer les données de recherche : elle peut être adaptée à chaque projet selon les besoins et les nécessités qui lui sont propres.

## 1\_Planification

### 1-1\_Coordination

- DMP
- Règles de nommage des dossiers et fichiers
- Journal de bord

### 1-2\_Documents-types

- Recrutement d'enquêtés (affiche, flyer, lettre, etc.)
- Modèle courrier d'invitation pour un entretien, questionnaire de sélection (vide)
- Modèle formulaire de consentement
- Modèle présentations PPT, poster, courrier, etc.

## 2\_Collecte

### 2-1\_Methodologie\_Recherche

- Méthodologie d'enquête (scénario)
- Protocoles de collecte
- Base de données : explicitation des différents champs avec conventions de remplissage
- Matériel : objets présentés aux enquêtés

### 2-2\_Donnees\_Brutes

*Cette partie peut contenir des données sensibles et/ou confidentielles.*

### 2-3\_Formulaires\_Consentement

- Questionnaires de sélection
- Consentements Focus groups
- Consentements Interviews
- Consentements Récolte de données
- Liste des répondants

### 2-4\_Statistiques

- Bases de données
- Questionnaires remplis
- Questionnaires complémentaires

### 2-5\_Notes

- Debriefing d'entretiens collectifs, de réunions, d'évènements (en lien avec la collecte)
- Observations faites par l'équipe de recherche



Arborescence à [télécharger](#)  
sous forme de dossier  
(en cours de révision)

## Organisation & documentation

- ➔ Nommer ses fichiers/dossiers de façon appropriée
- ➔ Garder la trace des différentes versions
- ➔ Des outils pour aider à (re)nommer rapidement

# Règles de nommage



Doc1



Document20210728



Entretien\_1\_MGK



Entretien2-mgk



Grille



Interview\_TZD\_1



Qui lit vraiment les présentations

# Organisation & documentation

Pas de caractères accentués, spéciaux ou des signes de ponctuation

- & , \* % # ( ) @ ^ ~ ' { } [ ] < > à é è ä ö ü ç ! ? , ; : + =

Pas de mots vides ou d'articles

- Le, la, les, un, une, des, et, ou, etc.

Pas de mots vagues, comme : divers, autres, à classer, archives, etc.

Pas d'espaces

- Utiliser des tirets ou tirets bas (underscore "\_") ou des majuscules pour séparer les mots



# Organisation & documentation

Nommer ses fichiers de façon approprié

- sujet\_type\_date\_version.extension
- type\_sujet\_date\_version.extension

Date de création du document ou de l'événement

- AAAA au minimum  
AAAAMMJJ ou AAAA-MM-JJ (ISO 8601)
- Pour nommer une période : utiliser un « \_ » au lieu du tiret « - ».
- À la fin du nom ou au début pour obtenir un classement chronologique

# Organisation & documentation

Utiliser un classement séquentiel avec des numéros

- Mettre des zéros à gauche pour garantir un affichage correct lors de l'affichage et du tri

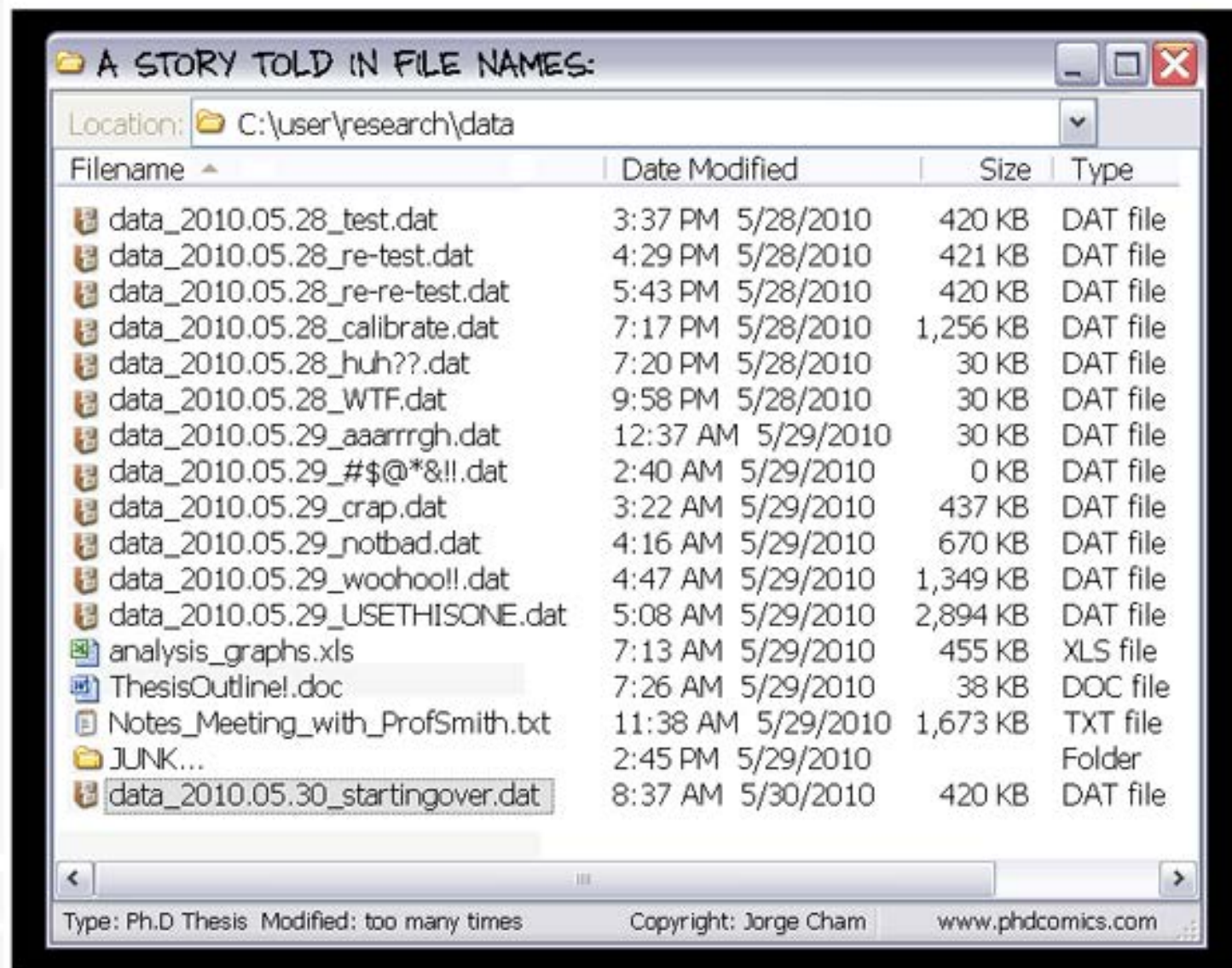
De 1 à 10 : 01, 02...010

De 1 à 100 : 001, 002...100

- Initiales des modificateurs
- Langue : ENG : en anglais / FR : en français / DE : en allemand

# Règles de nommage

Garder trace des  
différentes versions



# Règles de nommage

Garder trace des différentes versions

vp	Version provisoire
vf	Version finale
vv	Version validée
vdef	Version définitive
vf_conf	Version confidentielle
_pub	Version publiée (par ex. un article de presse)

# Règles de nommage

Utiliser un numéro de version

v0-0	Première version
v0-0 v1-0	Révision majeure
v0-1 v0-2	Révision mineure
Exemple : v2-3 → version majeure 2 avec 3 révisions	

## Règles de nommage


WSP\_2012Survey\_Apurimac\_20150718\_GP.csv

LabMeeting\_20180812\_RDM.docx

19630318\_letter\_LBJ-JFK\_p01.tiff

FR3S\_140623\_129C\_2653\_W.png

# Règles de nommage



## CONVENTION DE NOMMAGE

### 1. Introduction

En support de l'application de la Directive relative à la politique de gestion documentaire et d'archivage, la présente convention de nommage a été établie, ayant pour objectif d'uniformiser l'intitulé des dossiers et fichiers au sein de la HEIG VD.

Elle s'adresse à l'ensemble des collaboratrices et collaborateurs, et est connue de toutes les personnes impliquées dans la gestion des dossiers afin de permettre une collaboration efficace.

Il est important d'utiliser des règles de nommage claires et cohérentes : l'intitulé d'un dossier doit refléter l'affaire ou le processus auquel il se rapporte ; le contenu d'un fichier doit être compréhensible sans avoir à l'ouvrir.

### 2. Recommandations

- o Suivre les règles de nommage dès la création d'un dossier ou d'un fichier
- o Garder la trace des différentes versions de fichiers jusqu'à la clôture du dossier
- o Ne pas utiliser de nom trop long : de la racine jusqu'à l'extension du fichier (la partie rouge de l'exemple ci-dessous), il ne doit pas y avoir plus de 255 caractères dans un système d'exploitation Microsoft.

Exemple: 2008\_Soutien-aux-chercheurs08-08\_Outilsexemple.docx

### 3. Principes de nommage des dossiers et fichiers

#### 3.1 Le nom

- doit être court et significatif ;
- comporte 32 caractères au maximum pour un fichier ;
- comporte les éléments du tableau ci-dessous (voir 3.5 les éléments de nommage).

#### 3.2 Les caractères et pratiques à éviter

- Ne pas utiliser de caractères spéciaux, accentués ou des signes de ponctuation :
   
 & , \* % # [ ] @ ~ ' { } | < > & # \$ % & # 0 1 7 , : ; \* ' " \* \*
- Éviter les mots vides ou articles : le, la, les, un, une, des, et, ou ... ;
- Ne pas utiliser de mots vagues comme : divers, autres, à classer, archives, etc. ;
- Éviter les noms génériques qui pourraient entrer en conflit s'ils sont dupliqués dans un autre dossier (ex : « général » et « général (copie) »)

\* Si l'adresse de votre fichier est trop longue, vous ne pourrez plus accéder à votre document malgré son bon enregistrement, et le copier sur un autre support ou dans un autre dossier.

# Quelques outils pour aider à (re)nommer rapidement

Gestion des fichiers

😊 [Advanced Renamer](#) (Windows)

- [Ant Renamer](#) (Windows)
- [Bulk Rename Utility](#) (Windows)
- [Massive File Renamer](#) (Linux, Mac OS X, Windows)
- [Renamer 7](#) (macOS Ventura 13 or higher)



# Advanced Renamer

Advanced Renamer 3.92 - NON ENREGISTRÉ

**Méthodes de renommage**

Préréglages: [dropdown]

1 : Nouveau nom

Nouveau nom:  
Photo\_Inauguration\_20231001\_<Inc Nr:1>

Par Défaut | Avancé | Temps | Image | Vidéo

- <Inc Nr> - Incrementing number
- <Inc NrDir> - Incrementing number per folder
- <Inc Alpha> - Incrementing letter
- <Name> - File name without extension
- <Ext> - File extension without the dot
- <DirName> - Folder name
- <Num Files> - Total number of files in the folder
- <Num Dirs> - Total number of folders in the folder
- <Num Items> - Total number of files in the list
- <Word> - Indexed word of the filename

Tag : aide en ligne

Appliquer à: Nom

**Ajouter une méthode**

Nouveau nom	Changer la casse
Script	Date/Heure
Attributs	Rogner
Inversion	Liste de substitution
Liste	Insérer
Remplacer	Renommer

Batch mode: Renommer

Start batch

Files: Load, Save, Clear  
Folders: Add, List

Columns: Thumbnails, GPS Values  
Display: [icons]

Options: Pair renaming, Settings, Folder panel, Auto test

Echec

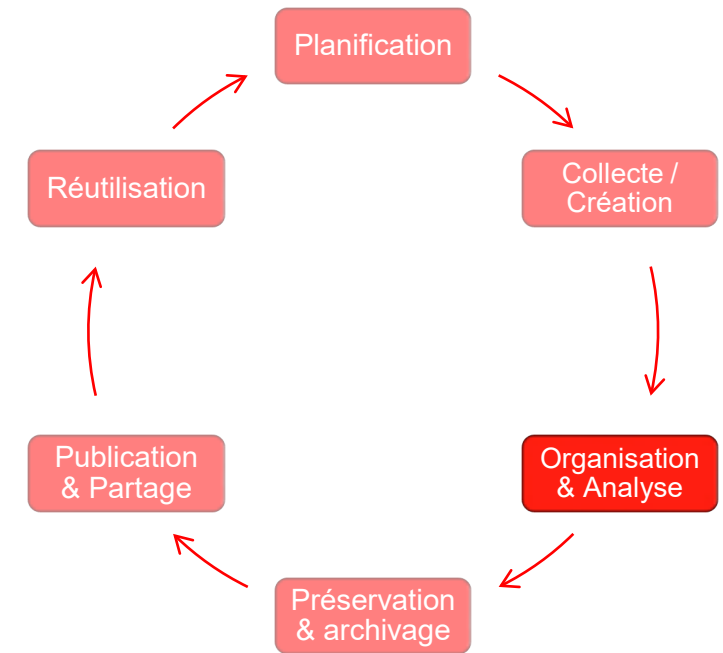
Renommer Fichiers | Renommer Dossiers

Nom de fichier	Nouveau nom	Chemin
2022-07-09 (1).jpg	Photo_Inauguration_20231001_001.jpg	D:\Data
2022-07-09 (2).jpg	Photo_Inauguration_20231001_002.jpg	D:\Data
2022-07-09 (3).jpg	Photo_Inauguration_20231001_003.jpg	D:\Data
6134210.pdf	Photo_Inauguration_20231001_004.pdf	D:\Data
6146757.pdf	Photo_Inauguration_20231001_005.pdf	D:\Data
6146805.pdf	Photo_Inauguration_20231001_006.pdf	D:\Data
6147216.pdf	Photo_Inauguration_20231001_007.pdf	D:\Data
IMG_20220724_152225.jpg	Photo_Inauguration_20231001_008.jpg	D:\Data
IMG_20220724_152229.jpg	Photo_Inauguration_20231001_009.jpg	D:\Data
IMG_20220724_152235.jpg	Photo_Inauguration_20231001_010.jpg	D:\Data
IMG_20220724_152238.jpg	Photo_Inauguration_20231001_011.jpg	D:\Data
IMG_20220724_152252.jpg	Photo_Inauguration_20231001_012.jpg	D:\Data

Nom de fichier:  
Dossier:  
Type de fichier:  
Taille:  
Date de Création:  
Date de Modification:  
Dernier Accès:  
Attributs:  
ExitTool...

267 Objets | 0 Erreurs | Etat: OK | Enregistrez-vous

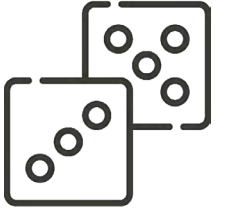
# Organisation & documentation



# Documenter ?

## UNFINISHED WORK





## Quel document est utile pour la documentation des données de recherche ?

- A. Dictionnaire de données
- B. Métadonnées
- C. Fichier README
- D. Toutes ces réponses

## Dictionnaires de données / Livre de code (codebook)

Document qui comprend la description des variables, etc.

Chaque variable devrait s'expliquer d'elle-même

Utilisation recommandée de vocabulaires contrôlés

Pour assurer la cohérence interne des données

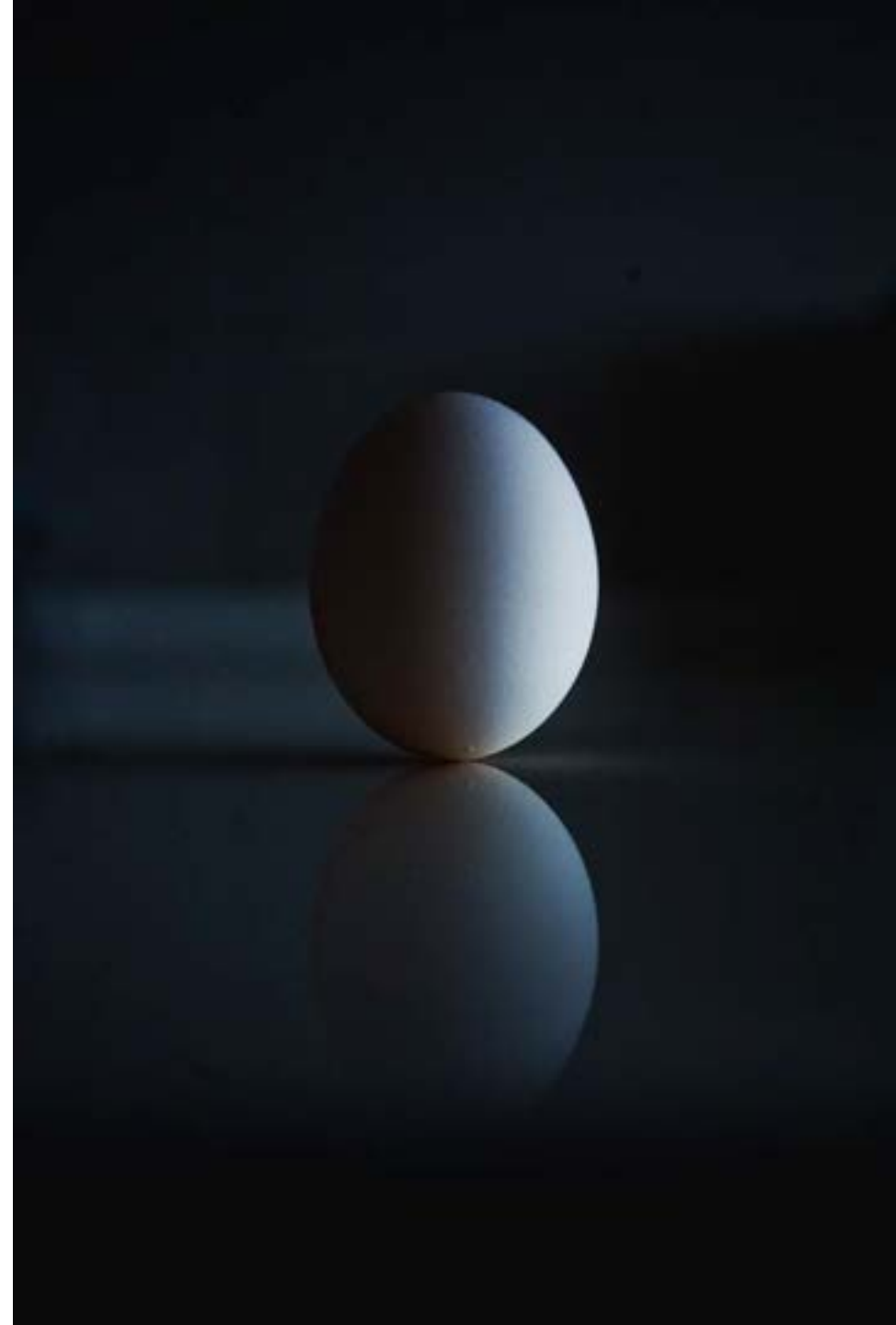
Pour faciliter leur réutilisation

# Donnée et métadonnée

Code œuf : 1FR WFA07

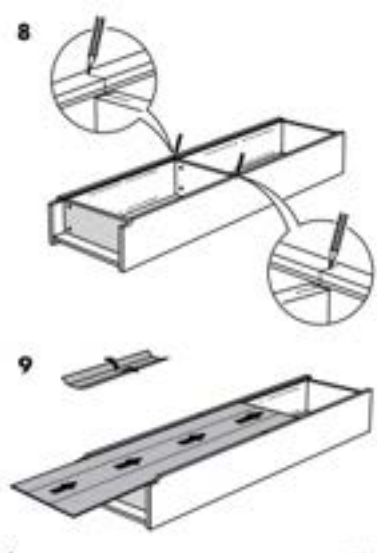
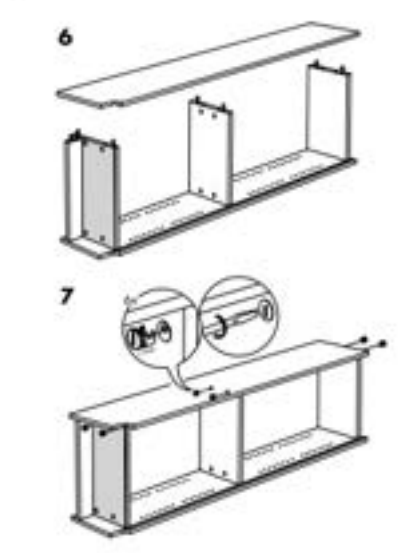
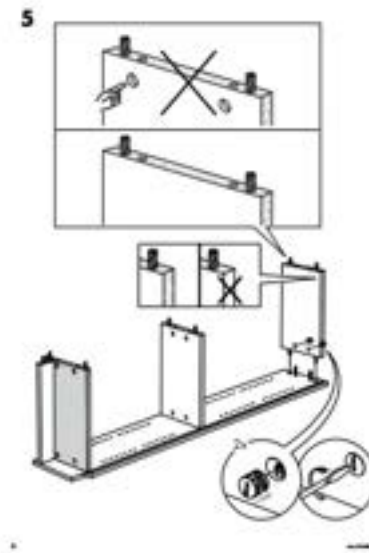
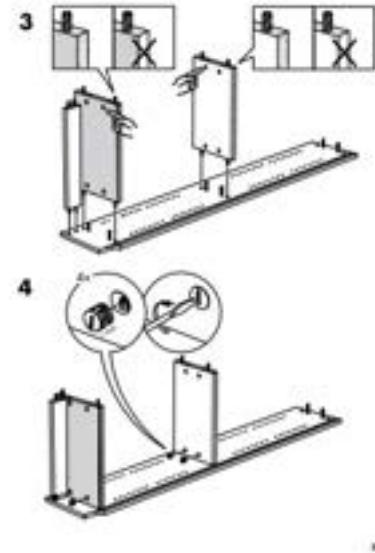
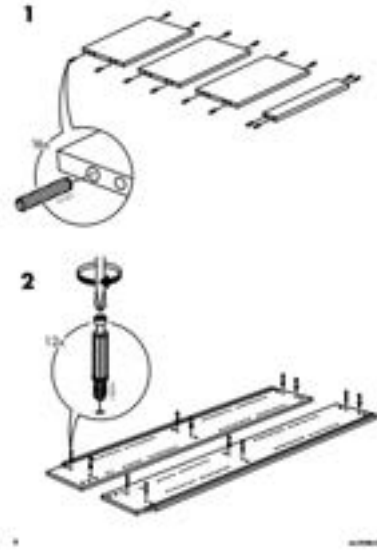


Documenter et décrire ses  
données de recherche



# Standard ?

## BILLY



# Exemple de métadonnées

Nom de fichier :

Type :

Auteurs : [Tania Zuber-](#)

Enregistre miniature

Propriétés	
Taille	0 octets
Pages	1
Mots	0
Temps total d'édition	0 minute(s)
Titre	Ajouter un titre
Balises	Ajouter un mot-clé
Commentaires	Ajouter des comme...
Modèle	Normal
État	Ajouter du texte
Catégories	Ajouter une catégorie
Objet	Spécifiez l'objet
Répertoire web	Ajouter du texte
Société	Indiquer la société

Dates associées	
Dernière modification	
Création	Aujourd'hui, 17:55
Dernière impression	

Personnes associées	
Responsable	Indiquer le responsa...
Auteur	Zuber Tania
	Ajouter un auteur
Dernière modification par	Pas encore enregistré

Documents associés	
	Ouvrir l'emplacement du fichier

[Afficher moins de propriétés](#)

Propriétés	
Taille	44.9 Ko
Pages	2
Mots	602
Temps total d'édition	45 minute(s)
Titre	Metadonnees des d...
Balises	Formation GC; Orga...
Commentaires	Ajouter des comme...
Modèle	Normal
État	Projet
Catégories	Formation
Objet	Spécifiez l'objet
Répertoire web	Ajouter du texte
Société	UNIL - UNIRIS

Dates associées	
Dernière modification	Aujourd'hui, 14:53
Création	Aujourd'hui, 13:39
Dernière impression	

Personnes associées	
Responsable	Indiquer le responsa...
Auteur	Zuber Tania
	Ajouter un auteur
Dernière modification par	Zuber Tania


Documents associés	
	Ouvrir l'emplacement du fichier

[Afficher moins de propriétés](#)



# Exemple de métadonnées

No HES-SO: 22791      Acronyme REM      Statut: En préparation      VS - HES-SO Valais-Wallis


**Définition**    Suivi    Clôture    Documents    

Données de base | Compléments | **Partenaires** | Budget | Valorisation prévue

No écran: 323 ?

**Retour**    **Héber**

**Partenaire**     Public     Privé

Raison sociale : VS - Institut Economie et tourisme 

**Nom partenaire :** VS - Institut Economie et tourisme

Service:

Type de financement :

Adresse:

NPA / Localité:

Pays: Suisse

Remarque:

**Types de partenariat :**

- Partenaire de recherche / formation
- Partenaire financier
- Partenaire de terrain
- Prestataire/Ss Traitance

**Enregistrer**

# Métadonnées DataCite

ID	DataCite-Property	Occ	Definition	Allowed values, examples, other constraints
1	Identifier	1	The Identifier is a unique string that identifies a resource. For software, determine whether the identifier is for a specific version of a piece of software, (per the Force11 Software Citation Principles <sup>11</sup> ), or for all versions.	DOI (Digital Object Identifier) registered by a DataCite member. Format should be "10.1234/foo"
1.a	identifierType	1	The type of Identifier.	<i>Controlled List Value:</i> DOI
2	Creator	1-n	The main researchers involved in producing the data, or the authors of the publication, in priority order. To supply multiple creators, repeat this property.	May be a corporate/institutional or personal name. Note: DataCite infrastructure supports up to 8000-10000 names. For name lists above that size, consider attribution via linking to the related metadata.
2.1	creatorName	1	The full name of the creator.	Examples: Charpy, Antoine; Jemison, Mae; Foo Data Center Note: The personal name, format should be: family, given. Non-

# Métadonnées disciplinaires

Spécifications par discipline !

Répertoires de schémas de métadonnées

- [Research Data Alliance](#)
- [Digital Curation Center](#)
- [FAIRsharing.org](#)

[Plus d'informations sur ces répertoires](#)

# Fichier descriptif README

Fichier readme.txt

- Pour le projet entier, à la racine du projet
- Pour chaque dossier / fichier si nécessaire

Exemples

- [Guide to writing “readme” style metadata](#)
- [AUTHOR\\_DATASET\\_ReadmeTemplate.txt](#)

# Fichier descriptif README

## Informations générales

Titre

Auteur·e, institution, information(s) de contact

Date de la collecte/production

Bailleur(s) de fonds

Informations sur des projets connexes

## Accès, liens/relations

Licence(s) des sets de données (conditions de partage/réutilisation possible)

Citation suggérée

Liens vers des publications qui citent ou utilisent ces données

Liens vers des données connexes ou similaires

## Organisation des données

Convention de nommage, structure des fichiers

Relations entre les fichiers

Liste d'autres fichiers avec documentation

## Set/Jeu de données

Liste de tous les fichiers avec une courte description

Méthodologie de collecte/production

Logiciel utilisé/nécessaire

Standards utilisés

## Livre de code

Liste et explication des abréviations/symboles utilisés

Format utilisé pour les dates (AAAAMMJJ ou AAAA-MM-JJ)

Traitement des données manquantes

## Traitement des données

Méthodologie appliquée

Logiciel utilisé/nécessaire (y.c. version)

Formats de fichiers

Procédure de contrôle de qualité appliquée

Historique des versions et raisons des mises à jour

Un fichier **README**, ou **LISEZMOI**, est un fichier au format texte (.txt) contenant des informations sur un ensemble de données. Ces informations aident à correctement interpréter ou comprendre les données, pendant ou après la fin du projet, par l'équipe qui les a produites, ou par d'autres personnes lors d'un partage ou d'une publication. Elles seront également nécessaires en cas de remplacement au sein de l'équipe.

Il est nécessaire de créer un fichier README pour un ensemble « logique » de données : fichiers apparentés, formatés de manière similaire, ou des fichiers regroupés logiquement pour être utilisés ensemble. Parfois, il peut être judicieux de créer un README pour un seul fichier de données. Si on en rédige plusieurs, il convient d'utiliser un format identique, en présentant les informations dans le même ordre et avec la même terminologie.

Il est préférable de créer un README dès le début de la collecte des données, voire pendant la phase de planification.

Il existe plusieurs modèles de fichiers README, par exemple :

- README en anglais de l'université de Cornell (USA) : <https://cornell.app.box.com/v/ReadmeTemplate>
- README en français de DoRANum (France) : [https://doranum.fr/wp-content/uploads/gabarit\\_readme.txt](https://doranum.fr/wp-content/uploads/gabarit_readme.txt)
- README en français avec l'affiliation HEIG-VD : [https://res.cloudinary.com/dxyvngy9l/bibliotheque/documents/heigvd\\_template\\_ReadmeDataSet](https://res.cloudinary.com/dxyvngy9l/bibliotheque/documents/heigvd_template_ReadmeDataSet)
- README en anglais avec l'affiliation HEIG-VD : [https://res.cloudinary.com/dxyvngy9l/bibliotheque/documents/heigvd\\_template\\_ReadmeDataSet\\_ENG](https://res.cloudinary.com/dxyvngy9l/bibliotheque/documents/heigvd_template_ReadmeDataSet_ENG)

### Contenu recommandé

Le contenu **minimum** recommandé est en gras.

Cette liste est une suggestion d'éléments que vous pourriez inclure dans un fichier README : certains n'ont de sens qu'avec des types de données précis. L'objectif n'est pas de rendre votre README trop long : il ne faut inclure que les éléments qui sont utiles et/ou nécessaires pour interpréter, évaluer et réutiliser correctement l'ensemble de données.

Les dates sont affichées selon des formats normalisés, par ex. selon la norme ISO 8601 : AAAA-MM-JJ.

### Informations générales

**Nom de l'ensemble de données**

**Nom, institution et informations de contact (adresse physique, e-mail) :**

- du/de la requérant e principal e,
- de la personne responsable de la collecte des données (si différente),
- des co-requérant es
- de toute autre personne de contact utile

Le fichier README.txt a été généré le AAAA-MM-JJ par NOM  
<le texte d'aide est inclus entre les chevrons, et peut être supprimé avant la sauvegarde>

#### INFORMATIONS GENERALES

1. Titre de l'ensemble de données:
2. Informations sur l'auteur:
  - A. Informations de contact du requérant principal
    - Nom:
    - Institution:Haute école d'ingénierie et de gestion du canton de Vaud, HES-SO Haute Ecole Spécialisée de Su
    - Adresse:
    - E-mail:
  - B. Informations de contact du co-requérant
    - Nom:
    - Institution:Haute école d'ingénierie et de gestion du canton de Vaud, HES-SO Haute Ecole Spécialisée de Su
    - Adresse:
    - E-mail:
  - C. Informations de contact alternatif
    - Nom:
    - Institution:Haute école d'ingénierie et de gestion du canton de Vaud, HES-SO Haute Ecole Spécialisée de Su
    - Adresse:
    - E-mail:
3. Date de collecte des données (date unique, intervalle de temps, date approximative) <format suggéré AAAA-MM-JJ> :
4. Emplacement géographique de collecte des données <latitude, longitude, or ville/région, canton, pays, le cas échéant>:

This readme file was generated on [YYYY-MM-DD] by [NAME]  
<help text in angle brackets should be deleted before finalizing your document>  
<[text in square brackets should be changed for your specific dataset]>

#### GENERAL INFORMATION

Title of Dataset:

<provide at least two contacts>

Author/Principal Investigator Information

Name:

ORCID:

Institution:School of Engineering and Management Vaud, HES-SO University of Applied Sciences and Arts Western Switzerland

Address:

Email:

Author/Associate or Co-Investigator Information

Name:

ORCID:

Institution:School of Engineering and Management Vaud, HES-SO University of Applied Sciences and Arts Western Switzerland

Address:

Email:

Author/Alternate Contact Information

Name:

ORCID:

Institution:School of Engineering and Management Vaud, HES-SO University of Applied Sciences and Arts Western Switzerland

Address:

Email:

Date of data collection: provide exact date, range or approximate date, suggested format YYYY-MM-DD

# La gestion du code

## Logiciel de gestion de version

- DAUZON Samuel, 2023. *Git : maîtrisez la gestion de vos versions (concepts, utilisation et cas pratiques)* [en ligne]. St-Herblain : Editions ENI. ISBN 9782409039621 [Consulté le 19 juin 2023]. Disponible à l'adresse : <https://bibliotheques.hes-so.ch/eni/> [accès par abonnement]

# Tout n'est pas données !

Logiciels de gestion bibliographique

- [Zotero](#) : [Guide de l'Infothèque de la HEG Genève](#)
- [EndNote](#) : [Guide de la Bibliothèque universitaire de médecine](#)

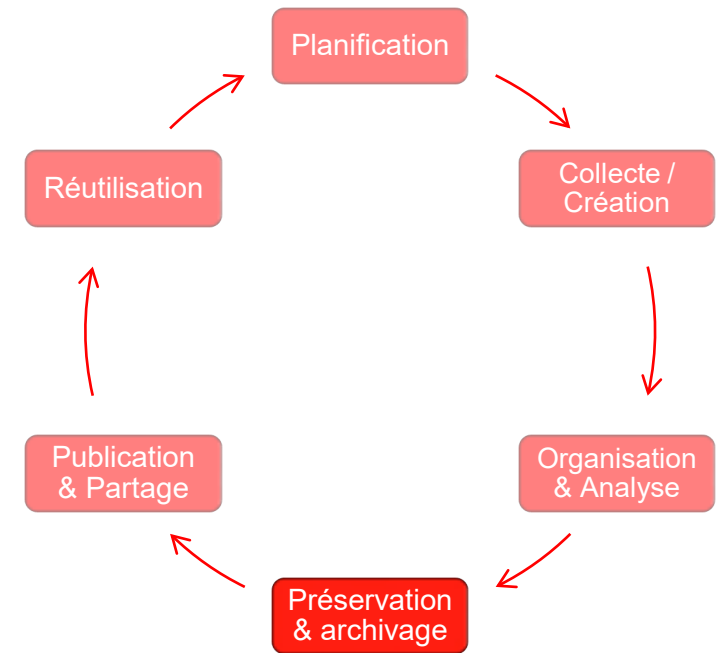


# Préservation et archivage

Comment mes données sont sauvegardées ?

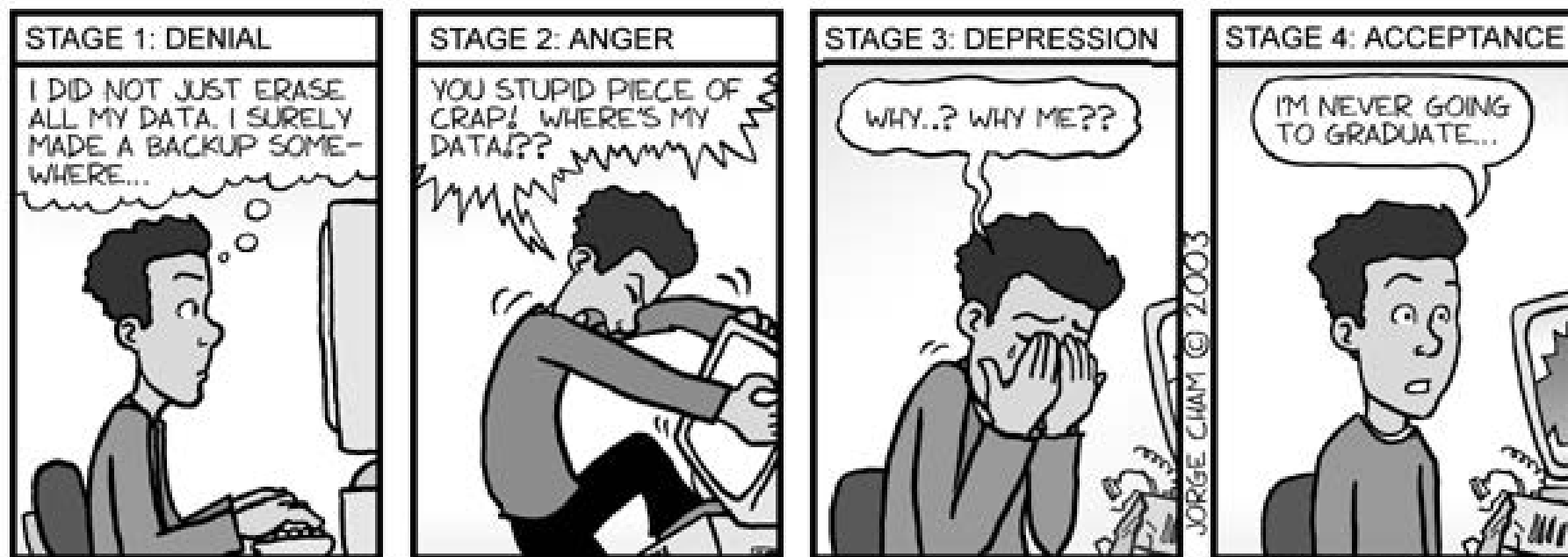
Suis-je certain·e de pouvoir les récupérer en cas de problème ?

Combien de temps dois-je les garder ?



# Stockage & sécurité

## THE FOUR STAGES OF DATA LOSS DEALING WITH ACCIDENTAL DELETION OF MONTHS OF HARD-EARNED DATA



[www.phdcomics.com](http://www.phdcomics.com)



TOUTE L'ACTUALITÉ / CLOUD / SAUVEGARDE

# Un bug dans Google Drive supprime certains fichiers

Célia Seramour, publié le 27 Novembre 2023



Plusieurs utilisateurs de Google Drive affirment avoir perdu des fichiers sans que cela ne soit lié à une fausse manipulation. Certains se retrouvent même avec un Drive dans l'état où il se trouvait plusieurs mois en arrière. Pour

## SUIVRE TOUTE L'ACTUALITÉ

Newsletter

Recevez notre newsletter comme plus

# Stockage & sécurité



## Stockage [Storage]

- Stockage actif ou courant sur les serveurs institutionnels
- Travailler ou collaborer sur les données

01

02

## Sauvegarde [Backup]

- Stockage actif additionnel
  - Copie des données originales pendant le projet
- Sur un autre support et/ou endroit



## Archive

- Stockage « froid »
- À des fins de conservation, généralement après la fin du projet : données conservées et accessibles à long terme
- Afin de réutiliser des données plus tard
- Garder le lien entre données et publication

03

## Stockage & sécurité

Identifier les données personnelles/sensibles nécessitant des mesures techniques spécifiques

Sauvegardes régulières et multi-supports :

règle 3-2-1 → règle 3-2-1-1-0

# Règle 3 - 2 - 1 - 1 - 0



● **COPIES**  
À JOUR  
DE VOS  
DONNÉES



● **TYPES DE**  
SUPPORTS  
DIFFÉRENTS



● **COPIE**  
HORS  
SITE



● **COPIE**  
HORS  
LIGNE



● **ERREUR**  
LORS DU TEST  
DE RESTAURATION

# Stockage & sécurité

Privilégier les [infrastructures institutionnelles](#)

Éviter les solutions commerciales Cloud

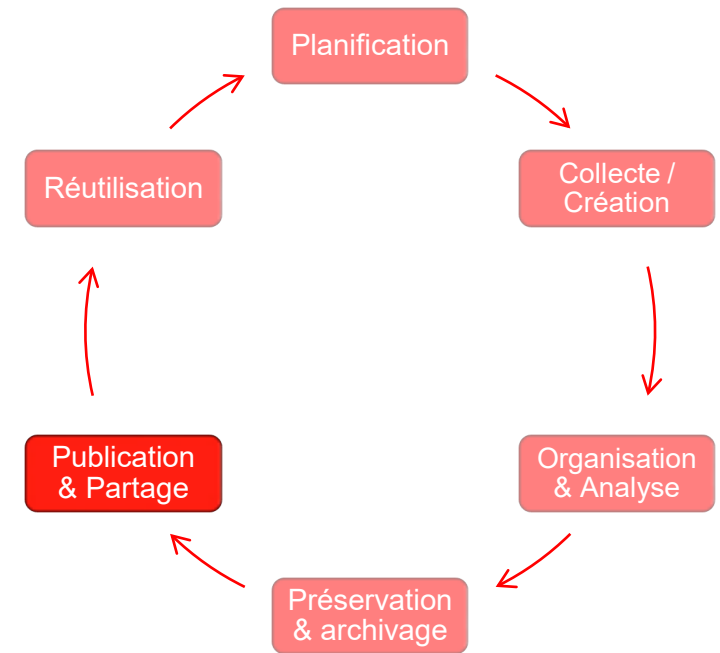
Microsoft / One Drive

CH : Interdiction de sauvegarder des données sensibles et des documents confidentiels dans le nuage de Microsoft

# Partage & dépôt

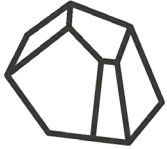
Sélectionner les données à déposer

Vérifier ses données, les nettoyer si nécessaire,  
les documenter





# Données brutes vs données traitées



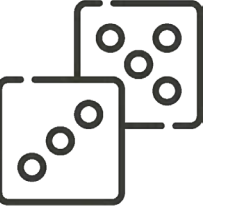
Données brutes



Données traitées



Données analysées : telles qu'apparaissant dans les publications



## Quelle plateforme permet la diffusion de données de recherche ?

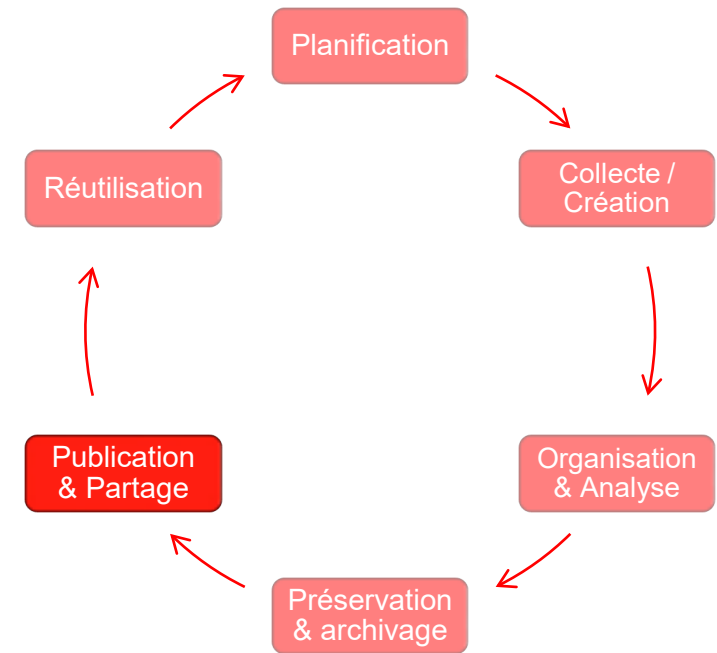
- A. Un dépôt comme Zenodo
- B. Un dossier Dropbox partagé
- C. Votre compte SWITCHdrive
- D. Toutes ces réponses

# Dépôt de données de recherche

Espace de publication des données de recherche

Fournit souvent un identifiant unique et permanent ([DOI](#))

Pas pour le stockage et le partage entre l'équipe au cours de la recherche





Published June 5, 2019 | Version v5

Dataset

Open

# Industry-scale Application and Evaluation of Deep Learning for Drug Target Prediction

Noé Sturm<sup>1</sup>; Andreas Mayr<sup>2</sup>; Thanh Le Van<sup>3</sup>; Vladimir Chupakhin<sup>4</sup> ; Hugo Ceulemans<sup>3</sup>;  
Joerg Wegner<sup>5</sup>; Jose-Felipe Golib-Dzib<sup>5</sup>; Nina Jeliaskova<sup>6</sup> ; Yves Vandriessche<sup>7</sup>; Stanislav Bohm<sup>8</sup>;  
Vojtech Cima<sup>8</sup>; Jan Martinovic<sup>8</sup>; Nigel Greene<sup>1</sup>; Tom Vander Aa; Thomas J. Ashby;  
Sepp Hochreiter<sup>2</sup>; Ola Engkvist<sup>9</sup>; Günter Klambauer<sup>2</sup>; Hongming Chen<sup>9</sup>

[Show affiliations](#)

Artificial intelligence (AI) is undergoing a revolution thanks to the breakthroughs of machine learning algorithms in computer vision, speech recognition, natural language processing and generative modelling. Recent works on publicly available pharmaceutical data showed that AI methods are highly promising for Drug Target prediction. However, the quality of public data might be different than that of industry data due to different labs reporting measurements, different measurement techniques, fewer samples and less diverse and specialized assays. As part of a European funded project (ExCAPE), that brought together expertise from pharmaceutical industry, machine learning, and high-performance computing, we investigated how well machine learning models obtained from public data can be transferred to internal pharmaceutical industry data. Our results show that machine learning models trained on public data can indeed maintain their predictive power to a large degree when applied to industry data. Moreover, we observed that deep learning derived machine learning models outperformed comparable models, which were trained by other machine learning algorithms, when applied to internal pharmaceutical company datasets. To our knowledge, this is the first large-scale study evaluating the potential of machine learning and especially deep learning directly at the level of industry-scale settings and moreover investigating the

972  
VIEWES526  
DOWNLOADS[Show more details](#)

## Versions

Version v5	Jun 5, 2019
10.5281/zenodo.3559987	
Version v5	Jun 5, 2019
10.5281/zenodo.3239499	

[View all 2 versions](#)

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.3239499](https://doi.org/10.5281/zenodo.3239499). This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

## Notes

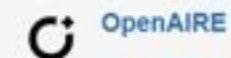
Dataset Format to reproduce manuscript results

## Files

Files (1.2 GB)		
Name	Size	Download all
<a href="#">activities.txt.gz</a> md5: 8dc6e1b61c1ef73c577d7434d9cc0132	190.9 MB	Download
<a href="#">activities_levels.txt.gz</a> md5: 19410b01e9c5da2fbb77cd6440f2f5d5	189.9 MB	Download
<a href="#">clustering.txt.gz</a> md5: 4e428221f720322083c24048b66a4c34	11.9 MB	Download
<a href="#">ecfp6_counts.txt.gz</a> md5: e26b290fbbd31d4ab27485ced96855e5	405.7 MB	Download
<a href="#">ecfp6_counts_var005.txt.gz</a> md5: b1f33641eda1a79aed6824c1fed0f6d3	166.5 MB	Download
<a href="#">ecfp6_folded.txt.gz</a> md5: 0848288b6f097c7f8d87dc54c1b2bad4	92.7 MB	Download

## External resources

Indexed in



## Keywords and subjects

QSAR machine learning chemogenomics  
benchmark dataset artificial intelligence

## Details

### DOI

DOI [10.5281/zenodo.3559987](https://doi.org/10.5281/zenodo.3559987)

### Resource type

Dataset

### Publisher

Zenodo

### Published in

Journal of Cheminformatics, 12(1), 26, 2019.

### Languages

English

## Additional details

### Related works

#### Is supplement to

Journal article: [10.1186/s13321-020-00428-5](https://doi.org/10.1186/s13321-020-00428-5) (DOI)

### Funding

#### ExCAPE – Exascale Compound Activity Prediction Engine

671555



European Commission

## Rights



Creative Commons Attribution 4.0  
International

## Citation

Noé Sturm, Andreas Mayr, Thanh Le Van, Vladimir Chupakhin, Hugo Ceulemans, Joerg Wegner, Jose-Felipe Golib-Dzib, Nina Jeliaskova, Yves Vandriessche, Stanislav Bohm, Vojtech Cima, Jan Martinovic, Nigel Greene, Tom Vander Aa, Thomas J. Ashby, Sepp Hochreiter, Ola Engkvist, Günter Klambauer, & Hongming Chen. (2019). Industry-scale Application and Evaluation of Deep Learning for Drug Target Prediction [Data set]. In *Journal of Cheminformatics* (Version v5, Vol. 12, Numéro 1, p. 26). Zenodo. <https://doi.org/10.5281/zenodo.3559987>

Style

APA



# GitHub

The Zenodo logo, consisting of the word "zenodo" in a white, lowercase, sans-serif font, centered on a solid blue rectangular background.

zenodo

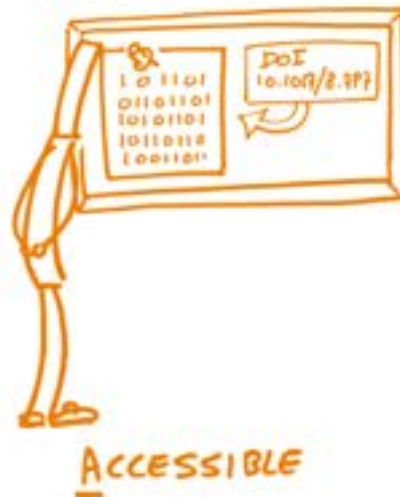
= code citable

The logo for ACOUA features the word "ACOUA" in a bold, white, sans-serif font. The text is centered within a rectangular area that has a dark blue background with a pattern of light blue binary code (0s and 1s) and a subtle, glowing effect.The logo for YARETA consists of a green, stylized flower-like icon on the left, made of overlapping petal shapes. To its right, the word "YARETA" is written in a bold, green, sans-serif font.



# Choisir un dépôt selon les principes FAIR

## FAIR DATA PRINCIPLES



# Principes FAIR : augmenter le potentiel des données

Attribuer un identifiant pérenne

Décrire les données avec  
des métadonnées

Toujours associées  
à leur source

Licence d'utilisation  
appropriée

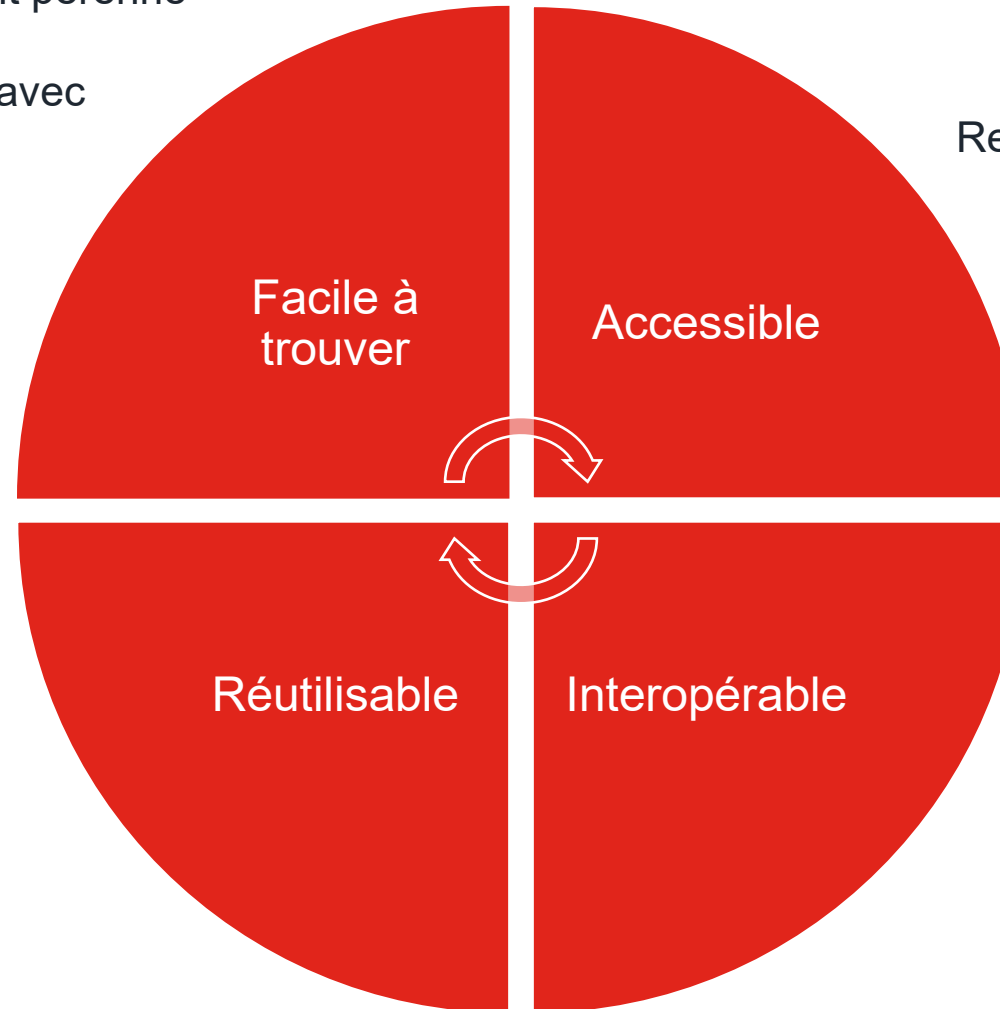
Gérer de façon pérenne l'accès  
aux données

Restriction si nécessaire

Utiliser des formats  
ouverts

Utiliser un vocabulaire  
partagé et/ou contrôlé

Standards communs



# Comment trouver un dépôt adapté ?

re3data.org  
REGISTRY OF RESEARCH DATA REPOSITORIES

🔍 Search

# Zenodo

**General****Institutions****Terms****Standards**

Name of repository

**Zenodo**

Additional name(s)

Research. Shared

Repository URL

<https://zenodo.org/>

Subject(s)

Humanities and Social Sciences

Life Sciences

Natural Sciences

Engineering Sciences

Description

ZENODO builds and operates a simple and innovative service that enables researchers, scientists, EU projects and institutions to share and showcase multidisciplinary research results (data and publications) that are not part of the existing institutional or subject-based repositories of the research communities. ZENODO enables researchers, scientists, EU projects and institutions to: easily share the long tail of small research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science. display their research results and get credited by making the research results citable and integrate them into existing reporting lines to funding agencies like the European Commission. easily access and reuse shared research results.

Contact

info@zenodo.org

<https://zenodo.org/contact>

Content type(s)

Standard office documents

Networkbased data

Images

Structured graphics

Audiovisual data

Scientific and statistical data formats

Raw data

Plain text

Structured text

Archived data

other

Source code

Keyword(s)

multidisciplinary

FAIR

Persistent identifier(s) of the repository

RRID:SCR\_004129

RRID:nlx\_158614

# Comment trouver un dépôt adapté ?

[Check-list du FNS](#) afin de vérifier le degré « FAIR » du dépôt choisi

Listes de dépôts recommandés

- [HES-SO](#)
- [FNS](#)
- [Open Research Europe](#)

# Tester le « FAIR » de vos données

Total across FAIR

0%

Q12 Answered Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12

## Findable

The data has sufficiently rich metadata and a unique and persistent identifier to be easily discovered by others. This includes assigning a persistent identifier (like a DOI or Handle), having rich metadata to describe the data and making sure it is findable through disciplinary local or international discovery portals.

1 Does the dataset have any identifiers assigned? [What is this?](#)

Globally unique, citable, and persistent (e.g. DOI, PURL, ARK or Handle)  Web Address (URL)

Local Identifier  No Identifier

2 Is the dataset identifier included in all metadata records/files describing the data?

Yes  No

3 How is the data described with metadata?

Comprehensively using a formal machine-readable metadata schema  Comprehensively, but in a text-based, non-standard format

Brief title and description  The data is not described

4 What type of repository or registry is the metadata record in?

Data is in one place but discoverable through several registries  Generalist public repository

Domain-specific repository  Local institutional repository

# HEIG<sup>VD</sup> Opter pour une licence ouverte (Creative Commons)

Plus ouverte



Attribution  
CC BY



Attribution / Partage dans les mêmes conditions  
CC BY-SA



Attribution / Pas de modification  
CC BY-ND



Attribution / Pas d'utilisation commerciale  
CC BY-NC



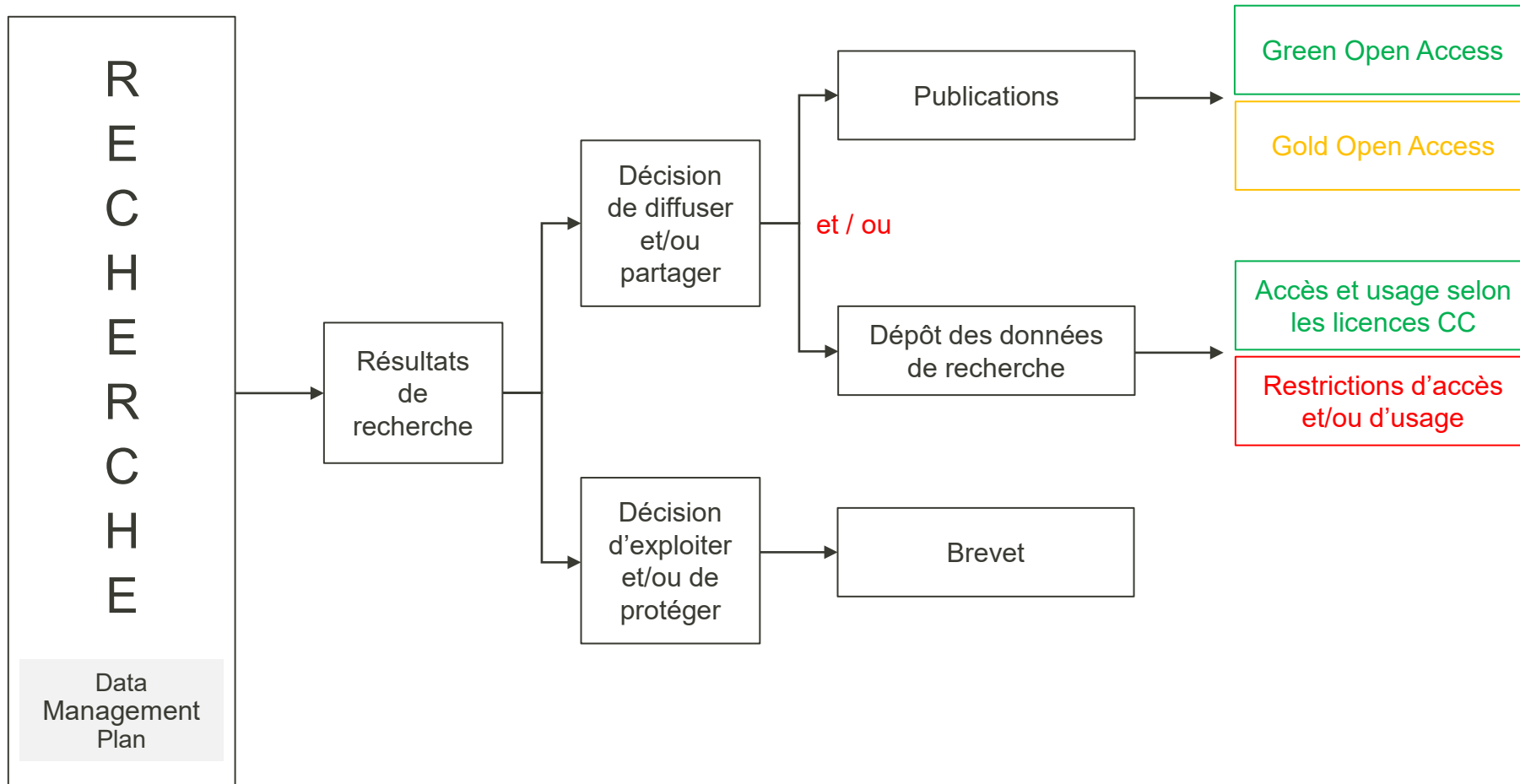
Attribution / Pas d'utilisation commerciale / Partage dans les mêmes conditions  
CC BY-NC-SA



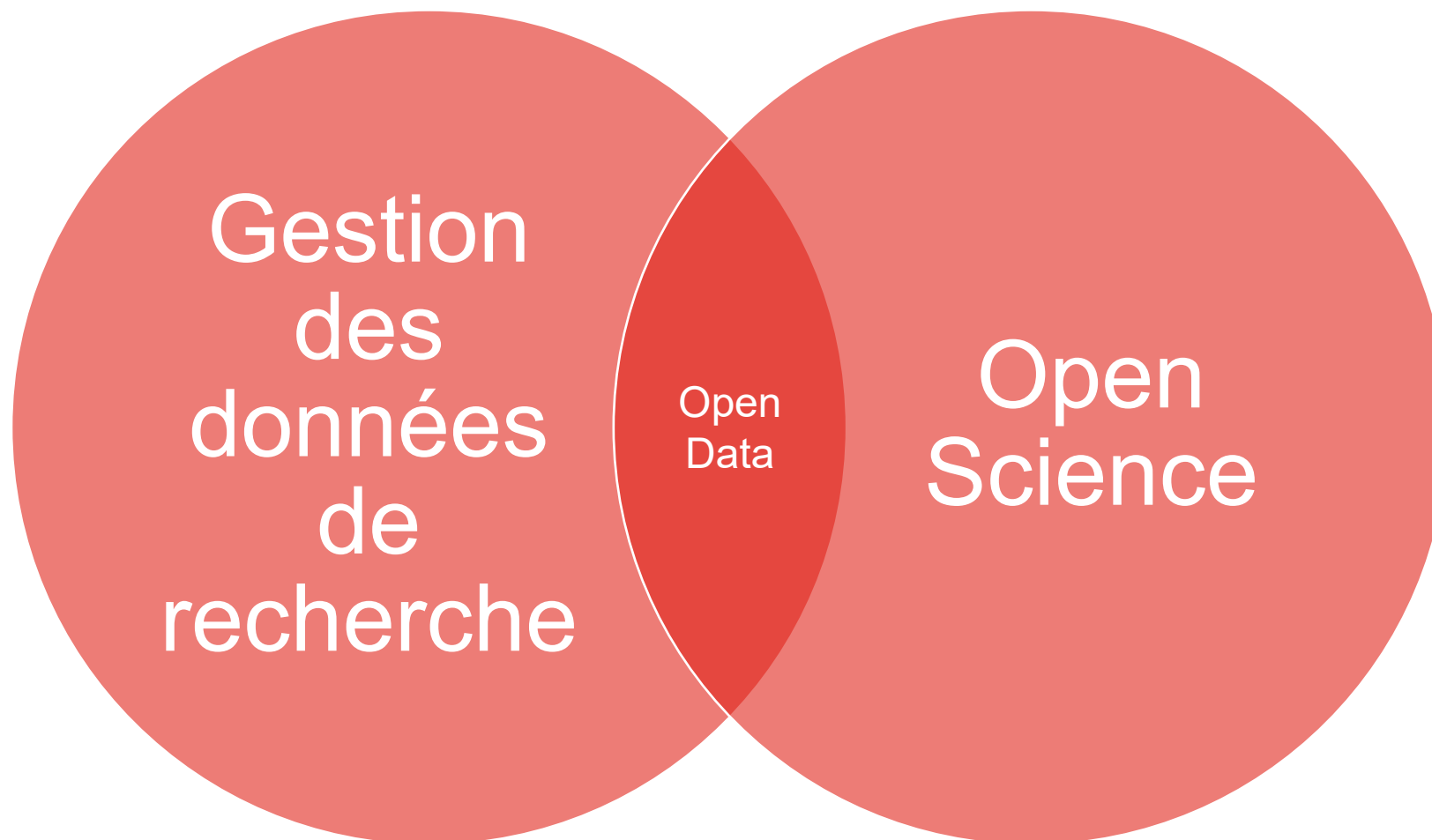
Attribution / Pas d'utilisation commerciale / Pas de modification  
CC BY-NC-ND

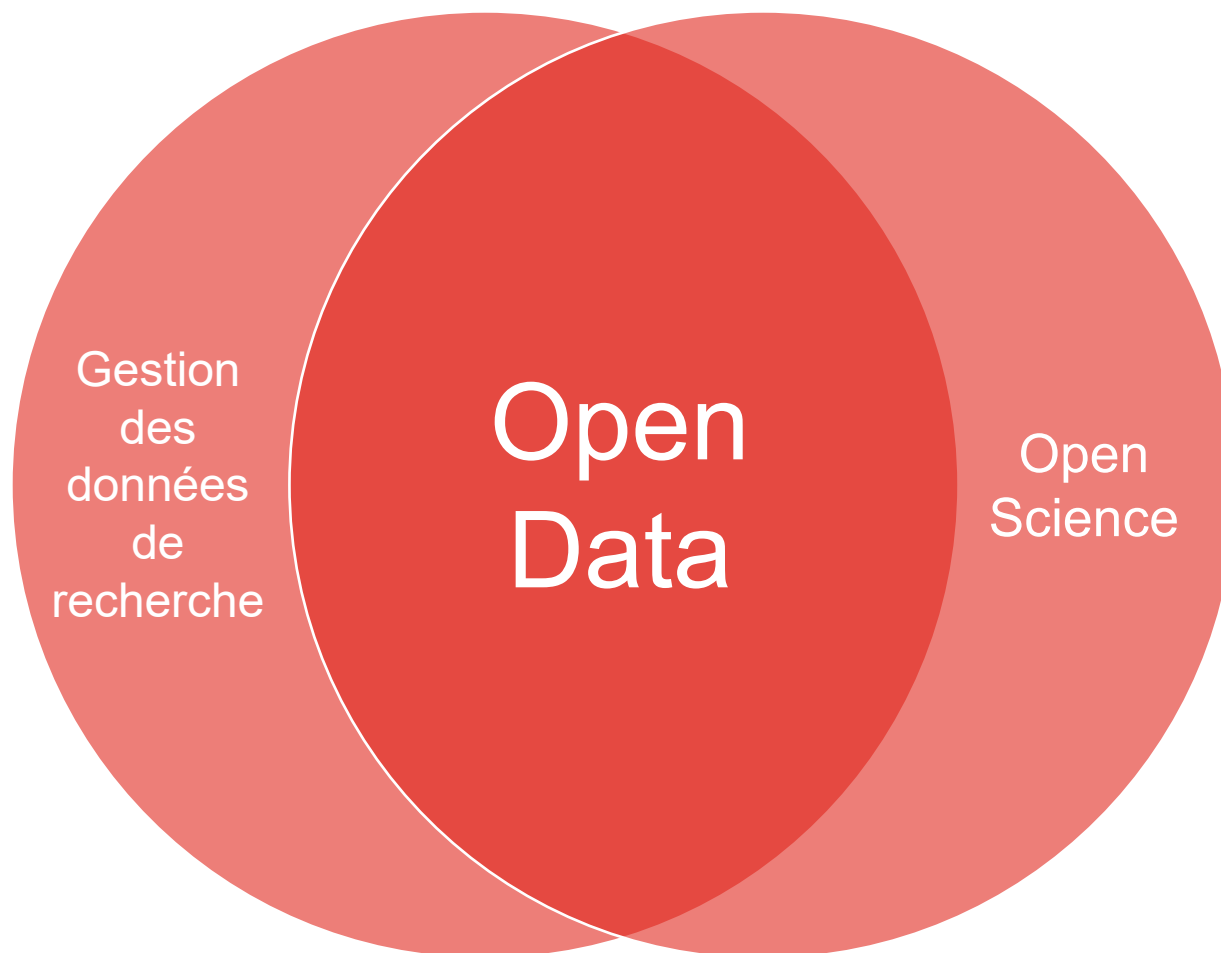
Moins ouverte

# Ouvrir ses données... oui, mais !

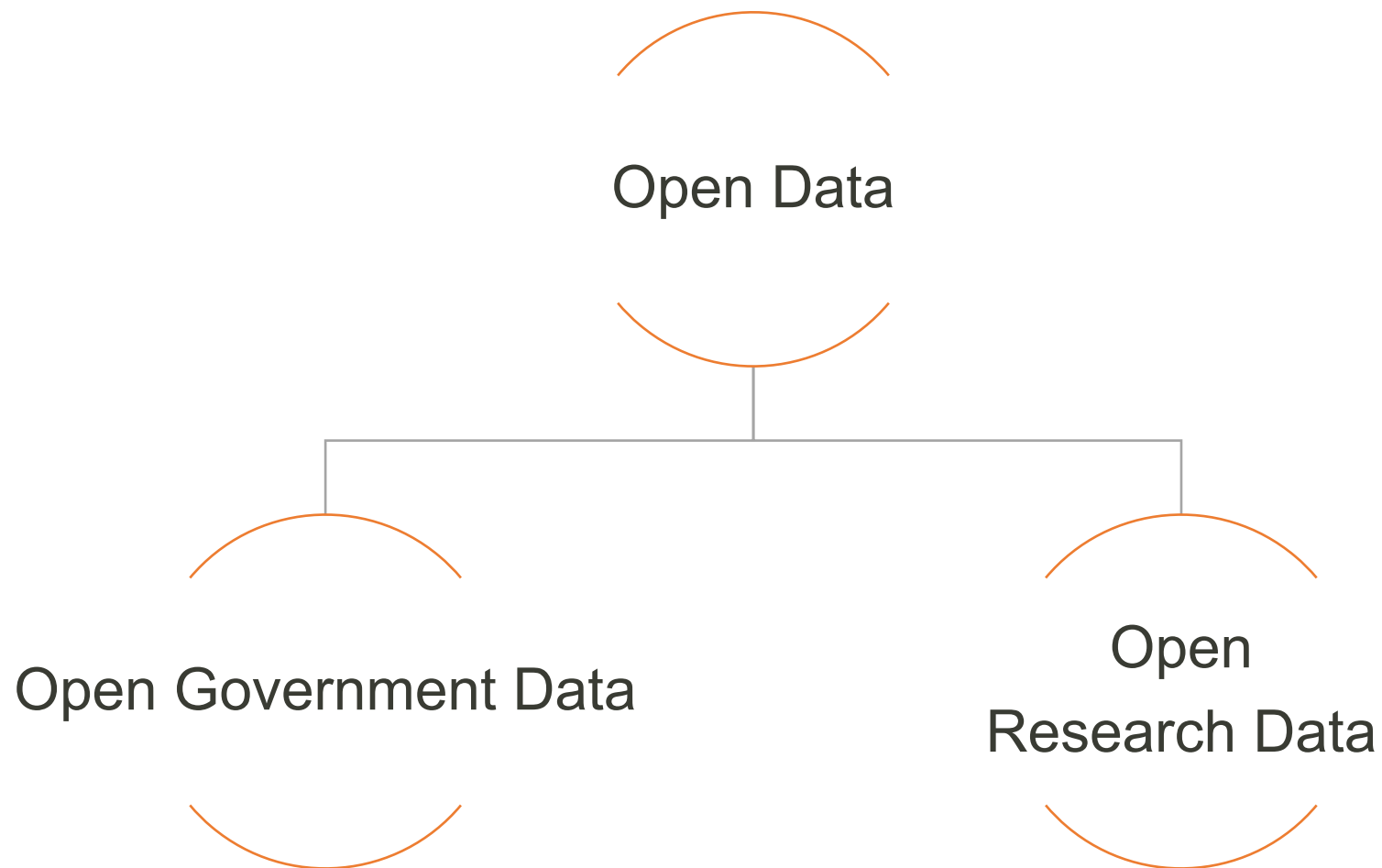








N'oubliez pas !



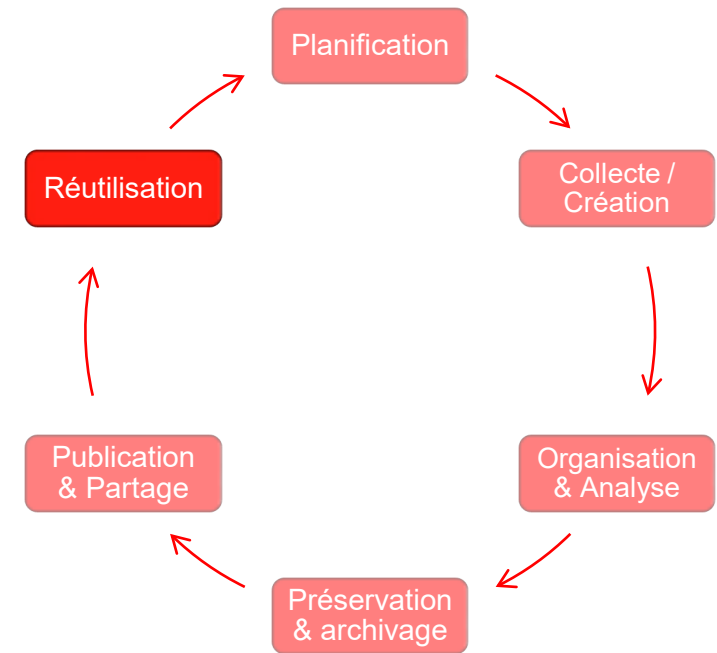
# Réutilisation et fin du projet

Qu'allez-vous faire des données ?

Poursuite du projet ?

Utilisation dans l'enseignement ?

Et plus encore ?



## Réutilisation et fin du projet

Est-ce qu'il y a encore des données qui pourraient/devraient être publiées ?

Qui est responsable des données après la fin du projet ?

Comment seront stockées (et organisées) les données à garder ?

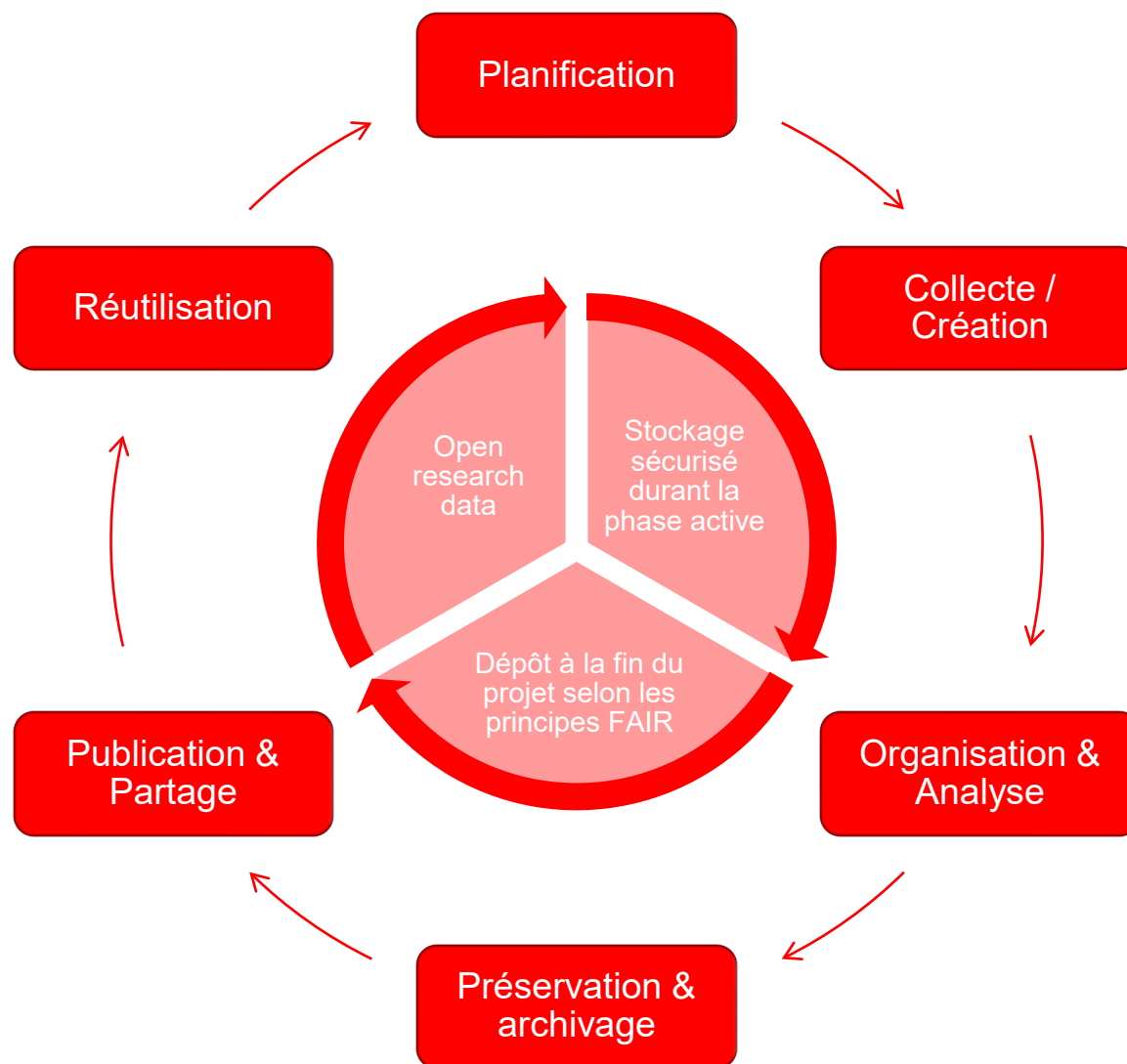
- Accès futur aux données nécessaire ?
- Si oui, comment ?

## Réutilisation et fin du projet

Est-ce qu'il y a des données à nettoyer ?

- Peuvent être supprimés : fichiers à double ou versions antérieures
- Données à supprimer à la fin du projet (selon contrat)
- Vérifier la documentation : est-ce que les fichiers README sont à jour et compréhensibles ?

# Gérer ses données de recherche



## **Vous souhaitez**

Une aide personnalisée pour votre publication / DMP

Une formation spécifique à un ou plusieurs outils

...

[biblio@heig-vd.ch](mailto:biblio@heig-vd.ch)  
[marylene.grzesiak@heig-vd.ch](mailto: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

Partager ses données en 5 étapes

Introduction à  
l'Open Access

Être FAIR avec ses données

Utiliser ORCID

Utiliser les licences Creative  
Common

Déposer un set de données sur  
Zenodo

Et plus encore...

# Ressources

Latour, M., Mahé, A., Copin, O., & Sauvage, B. (2023). On fait le point sur les données de la recherche avec Sorella ! (Version 1). [Consulté le 28 mai 2024]. Disponible à l'adresse :

<https://doi.org/10.5281/zenodo.10108736>

Ministère de l'enseignement supérieur, de la recherche et de l'innovation, centre national de la recherche scientifique et réseau des unités régionales de formation à l'information scientifique et technique, 2021. *DoRANum* - Données de la recherche : apprentissage Numérique [en ligne]. 2021. [Consulté le 30 septembre 2021]. Disponible à l'adresse : <https://doranum.fr>

PASQUIER, Guillaume, 2021. LibGuides: Research Data Management. In : *The Graduate Institute Geneva* [en ligne]. 27 septembre 2021. [Consulté le 30 septembre 2021]. Disponible à l'adresse :

<https://libguides.graduateinstitute.ch/rdm>

UK DATA SERVICE, 2021. Research data management. In : *UK Data Service* [en ligne]. 2021. [Consulté le 30 septembre 2021]. Disponible à l'adresse : [https://www.ukdataservice.ac.uk/manage-](https://www.ukdataservice.ac.uk/manage-data.aspx)

[data.aspx](https://www.ukdataservice.ac.uk/manage-data.aspx)

Icônes by [Freepik – Flaticon](#)

# Ressources

## Bibliothèque de la HEIG-VD

- [Convention de nommage HEIG-VD](#) : recommandations sur la structure et le nommage des fichiers
- [Proposition d'arborescence de vos fichiers](#)
- Arborescence de dossiers directement [téléchargeable au format ZIP](#)
- [Utilisation d'un fichier README](#) : que devrait-il contenir au minimum?
- Exemple de [fichier README en français](#) et [en anglais](#), avec l'affiliation HEIG-VD

## Data Management Plan

- DCC, ©2004-2021. [Checklist for a data management plan](#) ; [Develop a DMP](#)
- FNS, 2017. [Data Management Plan – Contenu du formulaire mySNF](#)
- EPFL Library, 2023. [Cost Calculator for Data Management](#)

# Ressources

## Organisation des données

- UK Data Archive, ©2012-2021. [Organising data](#) ; [Quality assurance](#) ; [Versioning](#)

## Documentation & métadonnées

- UK Data Archive, ©2012-2021. [Document your data](#)
- DCC, ©2004-2021. [Disciplinary Metadata](#)
- Cornell University, 2022. [Guide to writing "readme" style metadata](#)

## Formats de fichiers recommandés

- AFS, 2020. [Formats de fichiers adaptés à l'archivage](#)
- CINES, 2016. [FACILE - Service de validation de formats](#)
- ETH-Bibliothek, 2024. [File formats for archiving](#)
- UK Data Archive, ©2012-2021. [Recommended formats](#)

# Ressources

## Accès & sécurité

- UK Data Archive, ©2012-2021. [Data security](#) ; [Data encryption](#) ; [File sharing](#) ; [Regulating access to data](#)
- University of Bath, 2021. [Archiving and sharing data](#)

## Éthique

- [Commission cantonale d'éthique de la recherche sur l'être humain \(CER-VD\)](#)

# Ressources

## Droit d'auteur & propriété intellectuelle

- [Loi fédérale sur le droit d'auteur et les droits voisins](#) (Loi sur le droit d'auteur, LDA)
- Loi 419.01 sur les hautes écoles vaudoises de type HES (LHEV) [Art. 62 « Propriété intellectuelle »](#)
- Ball, A., DCC, 2014. [How to License Research Data](#)
- [Creative Commons](#)

# Ressources

## Dépôt & partage des données

- Dépôts (repositories) : [re3data.org](https://re3data.org) ; [OLOs](https://olos.org) ; [SWISSUbase](https://www.swissubase.ch) ; [Zenodo](https://zenodo.org) ; [Dryad](https://www.dryad.org)
- University of Bath, 2021. [Sharing data](#)
- Whyte, A., DCC, 2015. [Where to keep research data](#)

## Archivage & préservation à long terme

- Whyte, A., DCC, 2014. [Five steps to decide what data to keep](#)
- University of Bath, 2021. [Archiving data](#)
- Formats → cf. slide « [Formats de fichiers recommandés](#) »

# Licence de la présentation

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

**Bibliothèque de la HEIG-VD, 17 juin 2024**

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



Cette formation est librement adaptée de :  
BIBLIOTHEQUES UdeM, 2021. *Qu'est-ce que la gestion des données de recherche?* [en ligne]. Consulté le 28 septembre 2021. Disponible à l'adresse :  
<https://youtu.be/WKBqSnFoD2k>