

# Introduction aux données de la recherche et atelier DMP

24/03/2023

URFIST Occitanie

**Amélie Barrio - URFIST**

**[amelie.barrio@univ-toulouse.fr](mailto:amelie.barrio@univ-toulouse.fr)**

**URFIST**  
occitanie

Université  
Fédérale  
Toulouse  
Midi-Pyrénées



**Données de la recherche : définition(s) ?**

**Enjeux**

**Comment gérer les données ?**

**Partage et réutilisation**

# De quelles données parle-t-on ?

L'article 15 de la Déclaration des droits de l'homme et du citoyen de 1789 dispose que "la société a le droit de demander compte à tout agent public de son administration".

Les données produites et détenues par l'État constituent une ressource essentielle pour dynamiser la vie démocratique (transparence et concertation), stimuler l'innovation économique et sociale, voire pour améliorer le fonctionnement même de l'État (décloisonnement, meilleur pilotage des politiques publiques, etc).

La transformation numérique de l'État ne peut se passer d'une politique stratégique de la donnée. Un État intelligent doit s'appuyer sur ces données pour être plus efficace. Un accès plus large et de meilleure qualité aux données publiques permettra d'encourager l'ensemble de la société à créer de la valeur et de nouveaux services numériques utiles au public.

Source :

<https://www.gouvernement.fr/action/l-ouverture-des-donnees-publiques>

## L'OPEN DATA, C'EST QUOI?

L'open data désigne l'effort que font les institutions, notamment gouvernementales, qui partagent les données dont elles disposent. Ce partage doit être gratuit, dans des formats ouverts, et permettre la réutilisation des données. La politique du Gouvernement s'appuie notamment sur le droit d'accès aux documents administratifs (loi CADA 1978), qui considère que les données produites ou détenues par les administrations, dans le cadre de leurs missions de service public, doivent être mises à disposition du public. Cela ne concerne ni les informations personnelles, ni celles touchant à la sécurité nationale, ni celles couvertes par les différents secrets légaux.

Source :  
<https://www.gouvernement.fr/action/l-ouverture-des-donnees-publiques>

# Les données publiques

Des données en provenance de tous les secteurs



Source :  
<https://www.gouvernement.fr/action/l-ouverture-des-donnees-publiques>

**Introduction aux données de la recherche et atelier DMP**

24/03/2023

# Données de la recherche : tentative de définition

- Définition de l'OCDE :

« 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. »

(Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics - 2006)

- Définition de la Royal Society de Londres :

« Des informations qualitatives ou quantitatives [...] qui sont factuelles. Ces données peuvent être brutes ou primaires (directement issue d'une mesure), ou dérivée de données primaires, mais ne sont pas encore le produit d'analyse ou d'interprétation autres que le calculs. »

- Définition de l'Association des archivistes français :

« Les données de la recherche sont des informations, spécimens et matériaux produits, recueillis et documentés. Elles sont collectées ou exploitées à des fins de recherche et de preuve par les chercheurs et leurs équipes. A ce titre, elles constituent une partie des archives de la recherche. »

(Section Aurore AAF – 2014)

# Données de la recherche : tentative de définition

- Loi pour une République Numérique :

Article 38 : « Les copies ou reproductions numériques réalisées à partir d'une source licite, en vue de l'exploration de textes et de données incluses ou associées aux écrits scientifiques pour les besoins de la recherche publique, à l'exclusion de toute finalité commerciale. Un décret fixe les conditions dans lesquelles l'exploration des textes et des données est mise en œuvre, ainsi que les modalités de conservation et de communication des fichiers produits au terme des activités de recherche pour lesquelles elles ont été produites ; ces fichiers constituent des données de la recherche. »

(LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique)

- Définition de la Commission Européenne pour Horizon 2020 :

« Les données de la recherche désignent les informations, et en particulier les faits ou chiffres collectés pour être analysés et traités pour alimenter des réflexions, discussions ou calculs. Dans un contexte de recherche, ces données sont, par exemple, des statistiques, des résultats d'expériences, des mesures, des observations sur le terrain, des résultats d'enquêtes, des enregistrements d'entretiens ou des images. Il s'agit plus spécifiquement de données disponibles sous forme numérique . »

(Lignes directrices sur les règles pour le libre accès aux publications scientifiques et pour le libre accès aux données de la recherche dans Horizon 2020)

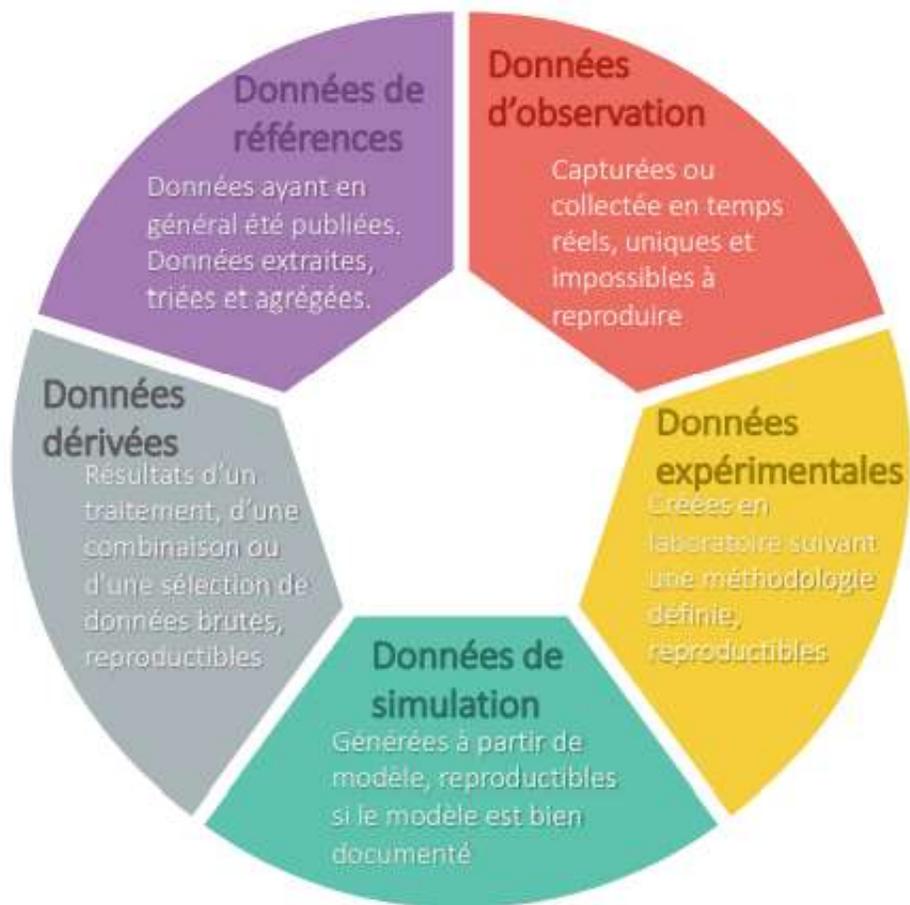
- Définition Plan pour la Science Ouverte:

« Enregistrements factuels (chiffres, textes, images, son, vidéo) utilisés comme sources primaires pour la recherche et qui sont habituellement acceptés par la communauté scientifique comme étant nécessaire pour valider les résultats de la recherche. »

(Plan pour la Science Ouverte, 04 juillet 2018)

Diapo Chloée Fabre - 2019

# Données de la recherche : quel périmètre ?



# Enjeux des données de la recherche

**Juridique** : <https://espacechercheurs.enpc.fr/fr/donnees-recherche-contexte-juridique>

**Scientifique (reproductibilité)**

**Ethique et intégrité**

**Sécurité**

PIXELS · ARMÉE

## Des soldats américains révèlent malencontreusement des secrets nucléaires

Contraints de mémoriser un grand nombre d'informations et d'acronymes, des militaires chargés de l'arsenal nucléaire ont utilisé des applications en ligne comme outils de révision, dévoilant des informations sensibles sans s'en rendre compte.

Le Monde avec AFP

Publié le 28 mai 2021 à 22h37 · Mis à jour le 29 mai 2021 à 05h42 · Lecture 2 min

[https://www.lemonde.fr/pixels/article/2021/05/28/des-soldats-americains-revelent-malencontreusement-des-secrets-nucleaires\\_6081945\\_4408996.html](https://www.lemonde.fr/pixels/article/2021/05/28/des-soldats-americains-revelent-malencontreusement-des-secrets-nucleaires_6081945_4408996.html)

**Politique**

**Economique**

PIXELS · CYBERCRIMINALITÉ

## Affaire Sébastien Raoult : qui sont les ShinyHunters, ces cybercriminels spécialisés dans le vol et la vente de données ?

En enquêtant sur ce gang spécialisé dans le hameçonnage ciblant particulièrement les développeurs présents sur GitHub, le FBI met un pied dans l'underground français du piratage informatique. Outre Sébastien Raoult interpellé au Maroc, deux autres Français sont visés par la justice américaine.

Par Gabriel Thierry

Publié le 04 août 2022 à 11h07 · Mis à jour le 06 août 2022 à 00h52 · Lecture 3 min · [Read in English](#)

[https://www.lemonde.fr/pixels/article/2022/08/04/affaire-sebastien-raoult-qui-sont-les-shinyhunters\\_6137126\\_4408996.html](https://www.lemonde.fr/pixels/article/2022/08/04/affaire-sebastien-raoult-qui-sont-les-shinyhunters_6137126_4408996.html)

**Introduction aux données de la recherche et atelier DMP**

24/03/2023

Stockage

Sauvegarde

Archivage

## Incendie chez OVH : protection des données, serveurs informatiques... un expert de l'université de Strasbourg décrypte

Après l'incendie du 10 mars qui a détruit des serveurs d'OVH à Strasbourg (Bas-Rhin), des sites Internet sont inaccessibles. Faute de protection suffisante de leurs données hébergées sur ces serveurs. Romaric David, responsable du centre de données de l'université de Strasbourg (Unistra), décrypte.

Publié le 11/03/2021 à 19h11 • Mis à jour le 12/03/2021 à 07h15

Source : <https://france3-regions.francetvinfo.fr/grand-est/bas-rhin/strasbourg-0/incendie-chez-ovh-cloud-sites-en-panne-entreprises-paralysees-un-expert-de-l-universite-de-strasbourg-explique-1994359.html>

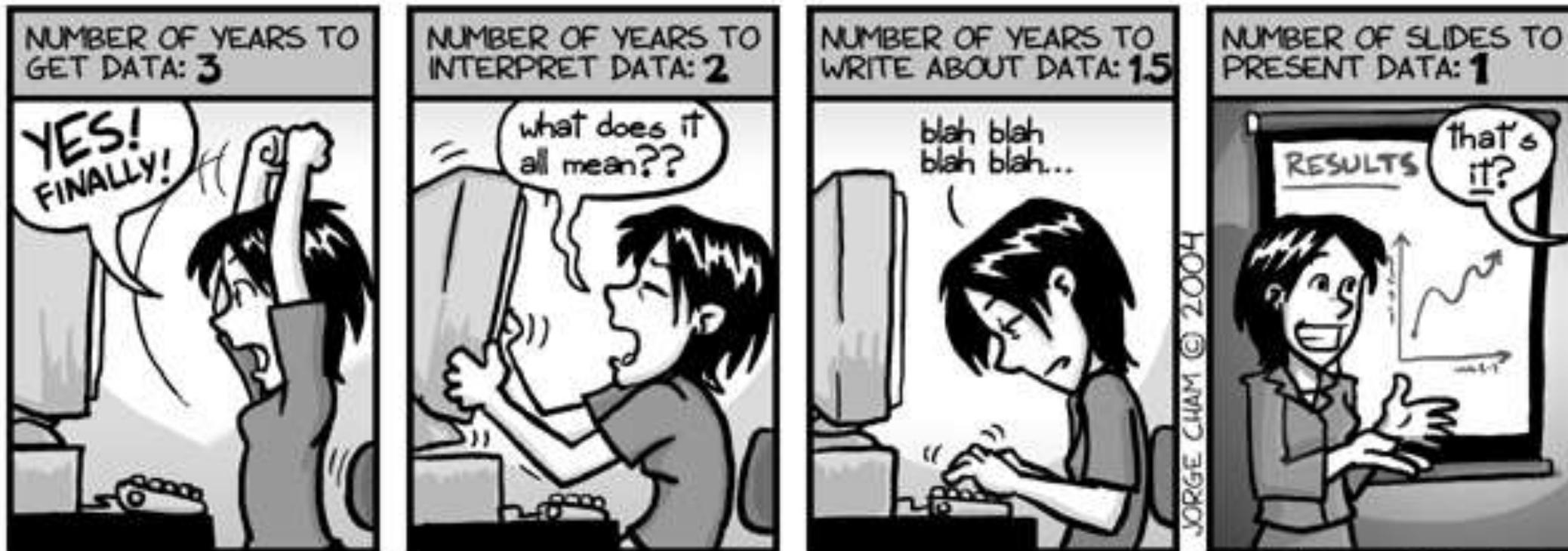
# **(Bien) gérer les données de recherche**

**Stockage** (charte de nommage, arborescence, standards de métadonnées, formats de fichiers, stockage sécurisé)

**Sauvegarde** (3 copies identiques, sur 2 supports différents dont 1 copie à un endroit différent des autres)

**Archivage** (charte d'archivage, tri et sélection, conservation pérenne)

## DATA: BY THE NUMBERS



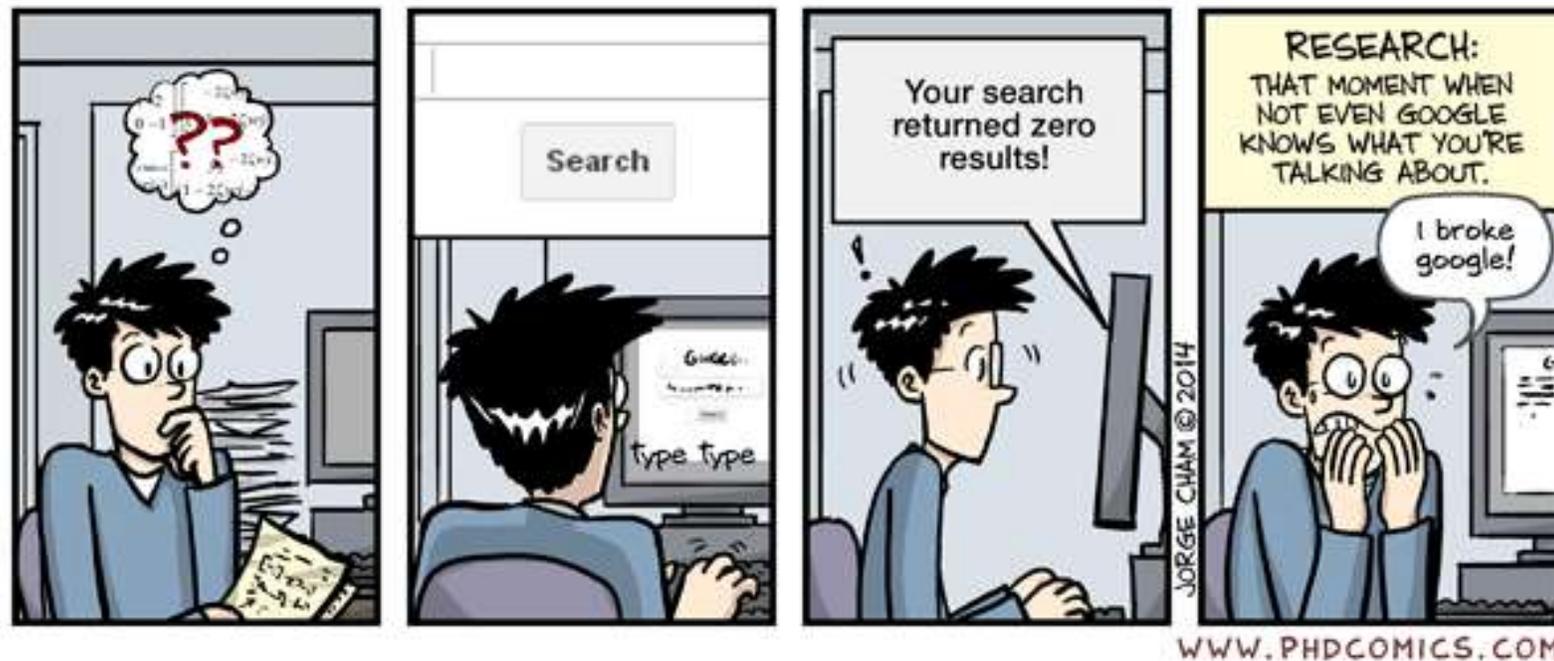
www.phdcomics.com

# La FAIRisation des données : pré-requis au partage



Source Doranum

# Faciles à trouver ?



- Des moteurs de recherche généralistes
  - Datacite Search
  - Dataset Search
  - OpenAire
  - Mendeley Data (Elsevier)
  - DataSearch (Elsevier)
  - Dimension et Lens
- Des entrepôts multidisciplinaires

Source : DataCC

# Comment FAIRiser des données ?

Déposer ses jeux de données dans des entrepôts de données

« certifiés » (ex : Recherche Data Gouv)

Attribuer un identifiant pérenne aux jeux de données

Enrichir les métadonnées des différents jeux

Coupler les métadonnées à l'identifiant pérenne

Rédiger et publier un Data Paper

Bien citer les jeux de données

# Comment FAIRiser des données ?

Adopter et suivre un plan de gestion de données

Quels mots-clés ?

Traduction, médiation, vulgarisation

Accessibilité totale ou partielle

Anonymisation des données sensibles

Attributions de licences d'usage



[ACCUEIL](#)

[RESSOURCES](#) ▾

[ACTUALITES](#)

[A PROPOS](#)

[CONTACT](#)



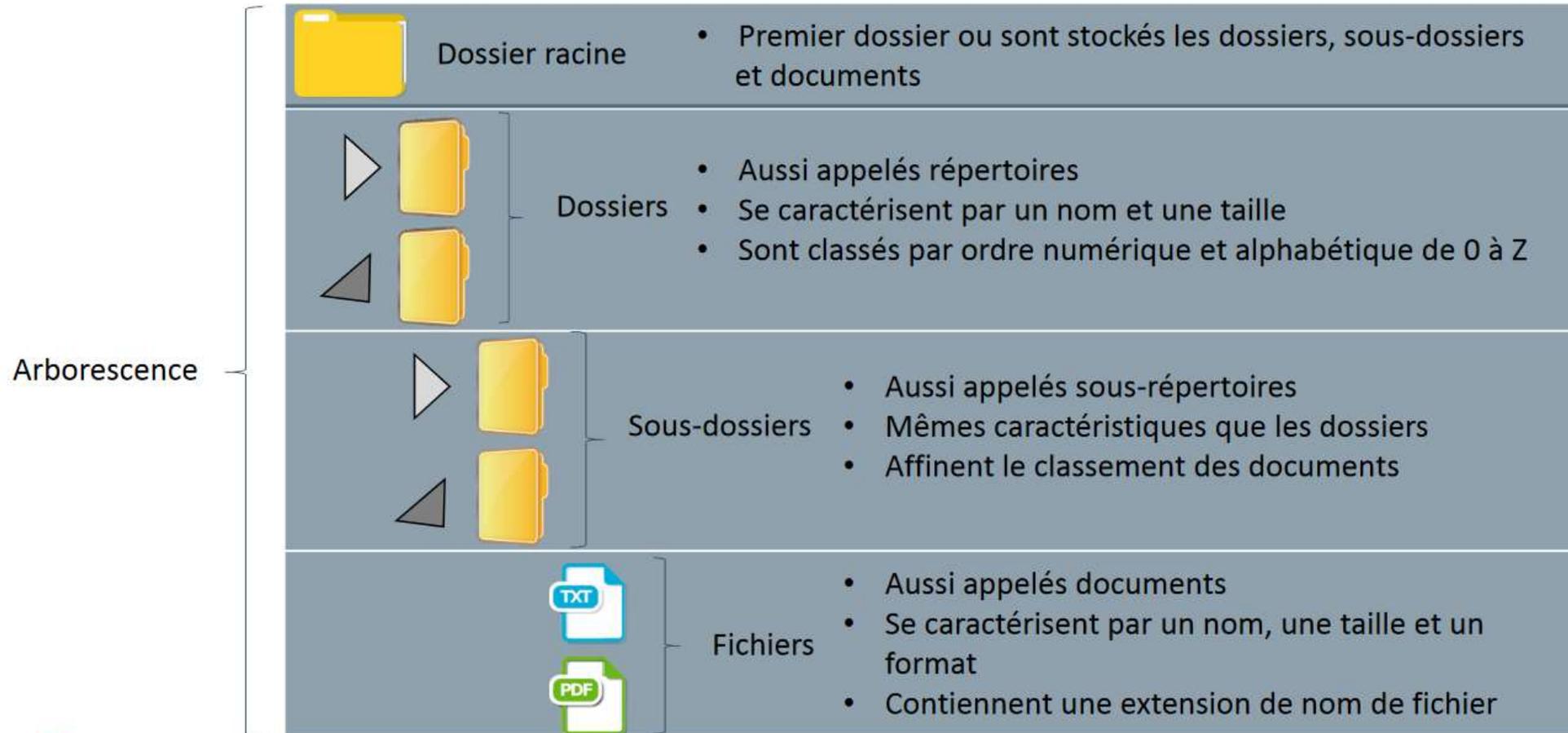
## DONNÉES DE LA RECHERCHE APPRENTISSAGE NUMÉRIQUE

DES RESSOURCES POUR ACCOMPAGNER LA COMMUNAUTÉ SCIENTIFIQUE DANS LA GESTION ET  
LE PARTAGE DE LEURS DONNÉES

TROUVER UNE RESSOURCE : VIDEO, FICHE, QUIZ...



[VOIR TOUTES LES RESSOURCES](#)



# Ressources DoRANum

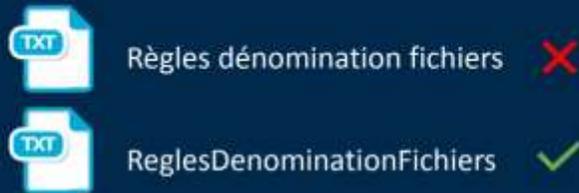
## 1 Nom bref et explicite

- Choisir un nom court reflétant son contenu
- La dénomination peut contenir le sujet, la date, la version
- Ne pas mettre de mots vides
- Utiliser des abréviations communément compréhensibles
- Ne pas répéter les termes entre le nom du fichier et le nom du dossier



## 2 Pas d'espaces ni de caractères spéciaux

- Ne pas mettre d'espaces mais des majuscules ou des underscores « \_ »
- Ne pas mettre de caractères spéciaux, seulement des caractères alphanumériques



## 3 Dates au format AAAAMMJJ

- Indiquer les dates au format AAAAMMJJ (Année, Mois, Jour)
- Commencer par l'année permet de maintenir les documents dans un ordre chronologique (le plus récent se retrouvera toujours en fin de liste)

## 4 L'élément le plus important en premier

- Mettre l'élément le plus pertinent en premier pour faciliter la recherche du document (la date par exemple)

## 5 Version des documents

- Mettre la version du document si nécessaire
- Mettre la lettre V suivie du numéro de la version en deux unités ou préciser les versions provisoires (VP) et finales (VF)

[https://doranum.fr/stockage-archivage/comment-nommer-fichiers\\_10\\_13143\\_wgqw-aa59/](https://doranum.fr/stockage-archivage/comment-nommer-fichiers_10_13143_wgqw-aa59/)

**Ouvert  
ou  
Fermé ?**

*Parmi les formats de fichiers qui vous sont proposés, lesquels sont "ouverts" ou "fermés" ? Testez vos connaissances avec ce mini jeu !*

Bonne chance 🍀

**JOUER**

CC BY NC SA

20

<https://doranum.fr/stockage-archivage/quiz-format-ouvert-ou-ferme-10-13143-mc-wq-qs64/>

# Pour aller plus loin



## FACILE - Service de validation de formats

Vérifier l'éligibilité de vos documents à un archivage sur la plateforme PAC du CINES.

Validation

Correction PDF

Tutoriels

Web Service

+ Choisissez des fichiers - Taille max 2,5 Go

↗ Valider

⊗ Annuler

[Cliquez ici pour demander l'aide d'un expert du CINES](#)

Liste des formats validables

⚠ Attention : le validateur de formats permet de valider certains formats qui ne sont pas pris en charge par la plateforme d'archivage du CINES.

Format	Nom	PRONOM PIUD	Type MIME	Commentaire	Archivable dans PAC
<a href="#">AAC AAC</a>	Advanced Audio Codings	[fmt/199]		Format Mpeg-4 contenant uniquement un flux audio au format AAC.	✓
<a href="#">AIFF PCM</a>	Audio Interchange File Format	[fmt/414]	[audio/x-aif, audio/x-aiff]	Format audio contenant uniquement un flux PCM.	✓
<a href="#">APNG</a>	Animated Portable Network Graphics	[fmt/935]	[image/vnd.mozilla.apng, image/apng]	L'APNG est une extension du format PNG permettant de réaliser des animations graphiques.	✗

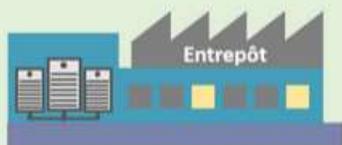
<https://facile.cines.fr/>

Introduction aux données de la recherche et atelier DMP

24/03/2023

## Stockage, partage et archivage : quelles différences ?

### PARTAGE



Entrepôt généraliste  
ou disciplinaire



Le partage consiste à déposer vos données dans un entrepôt de données afin de les rendre accessibles facilement et de permettre leur réutilisation (principes FAIR) sur le court et le moyen terme (5 à 10 ans), pour des chercheurs du même domaine, ou d'un autre domaine.

#### Comment faire ?

- Sélectionner les données à partager.
- Estimer la volumétrie des données et éventuellement le budget.
- Vérifier la compatibilité et l'interopérabilité des formats de données en se posant par exemple ces questions :
  - Le contenu est-il valorisé ou toujours valorisable ?
  - Le fichier et son contenu sont-ils compréhensibles en l'état ?
  - Le format de fichier utilisé nécessite-t-il un logiciel pour être lisible ?
- Migrer si nécessaire vers un format adapté, le plus ouvert possible.
- Préparer si nécessaire les codes sources (ex : scripts) qui permettront de lire et traiter les données.
- Compléter et enrichir les métadonnées (en fonction de l'entrepôt choisi).
  - Si ce n'est pas déjà fait, choisir un standard de métadonnées.
  - S'il n'en existe pas d'adapté, créer un schéma de métadonnées.
  - Compléter les champs pour chaque jeu de données, suivant le standard adopté.

<https://doranum.fr/stockage-archivage/stockage-partage-archivage-quelles-differences-10-13143-5d-ax-qp58/>



Because good research needs good data

About	▼
News	▼
Events	▼
Services	▼
Guidance	▲
Briefing Papers	
How-to Guides	
Case Studies	

**Home » Guidance » Metadata**

## Disciplinary Metadata

While data curators, and increasingly researchers, know that good metadata is key for research data access and re-use, figuring out precisely what metadata to capture and how to capture it is a complex task. Fortunately, many academic disciplines have supported initiatives to formalise the metadata specifications the community deems to be required for data re-use. This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them.

# Trouver de l'aide



recherche.data.gouv.fr

Déposer / Publier un jeu de données

FR

À propos Membres de l'écosystème Données Aide en ligne Actualités

## Un écosystème au service du partage et de l'ouverture des données de recherche

FÉDÉRER, ACCOMPAGNER, PARTAGER, OUVRIR, RÉUTILISER

Rechercher un jeu de données

Trouver des données

Déposer / Publier un jeu de données

<https://recherche.data.gouv.fr/fr>

Introduction aux données de la recherche et atelier DMP

24/03/2023

# Pour aller plus loin : codes sources et algorithmes



The screenshot shows the CodeGouv website interface. At the top left is the French Republic logo with the text 'RÉPUBLIQUE FRANÇAISE' and 'Liberté Égalité Fraternité'. Next to it is the 'CodeGouv' logo and the tagline 'Les codes sources du secteur public'. On the right, there are social media icons for GitHub (@codegouvfr), Twitter (@codegouvfr), an RSS feed icon, and a settings icon labeled 'Paramètres d'affichage'. Below this is a navigation menu with links: Accueil, Dépôts, Organisations, Bibliothèques, Dépendances, Versions taguées, SILL, Services, Chiffres, and À propos. The main content area features four cards: 1. 'Dépôts' with a keyboard image, text 'Parcourez la liste des dépôts de code source publiés par des organismes publics pour les réutiliser ou y contribuer.', and a right arrow. 2. 'Organisations' with a modern building image, text 'Quelles sont les organisations qui publient des codes sources ?', and a right arrow. 3. 'Bibliothèques' with a bookshelf image, text 'Quelles sont les bibliothèques de code publiées à', and a right arrow. 4. 'Dépendances' with a rock stack image, text 'Quelles sont les bibliothèques de code les plus', and a right arrow.

<https://code.gouv.fr/>

**Introduction aux données de la recherche et atelier DMP**

24/03/2023

# Pour aller plus loin : codes sources et algorithmes



## Les Codes Sources

🔍 Chaque ligne de code compte !

```
session_start();
ob_start();
include 'sys/login.php';
include 'init.php';

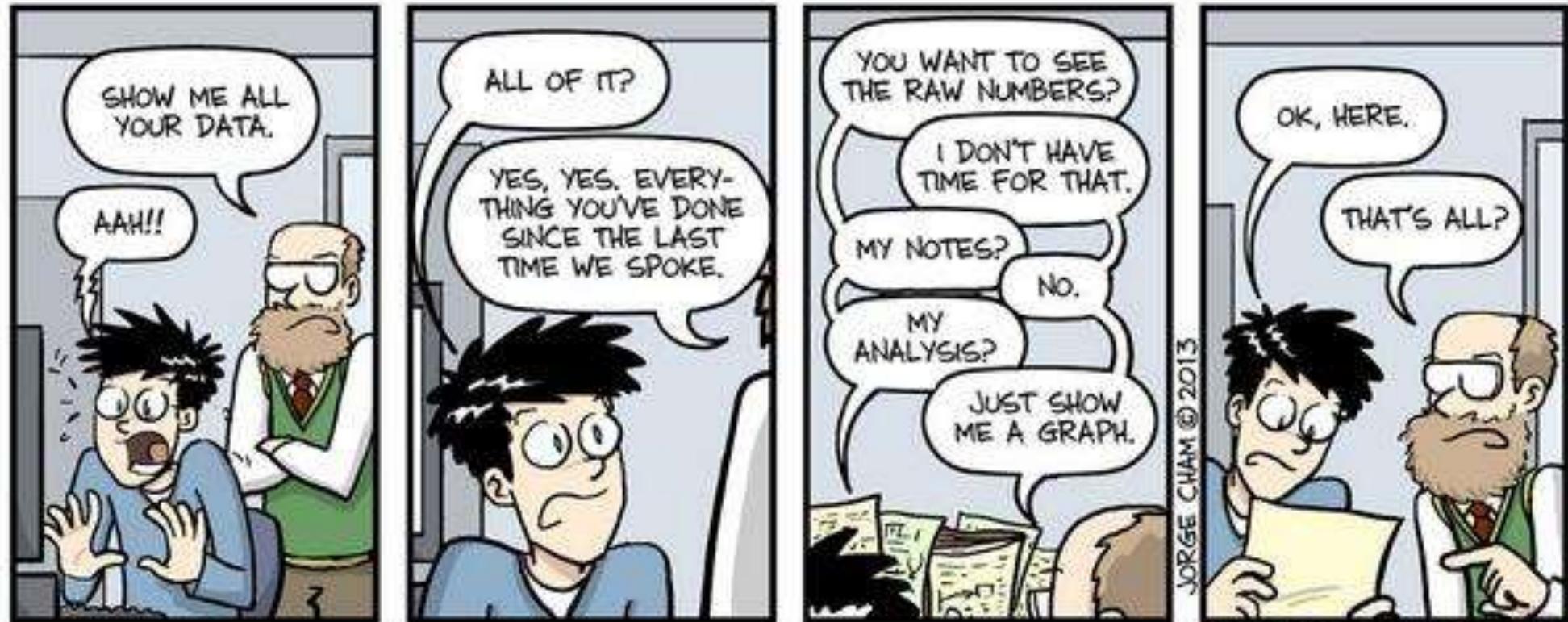
<DOCTYPE html>
<html>
<head>
<meta content="text/html"/>
<meta charset="utf-8" session="cnsconfig" ["charset"]>
<meta http-equiv="Content-Language" content="fr" session="cnsconfig" ["content-language"]>
<meta http-equiv="Content-Type" content="text/html" session="cnsconfig" ["content-type"]>
<!-- Basic Meta -->
<title>+$_SESSION["cnsconfig"]["title"]+</title>
<meta name="description" content="+$_SESSION["cnsconfig"]["description"]+</meta>
<meta name="keywords" content="+$_SESSION["cnsconfig"]["keywords"]+</meta>
<meta name="author" content="+$_SESSION["cnsconfig"]["author"]+</meta>
<meta name="owner" content="+$_SESSION["cnsconfig"]["owner"]+</meta>
<meta name="robots" content="+$_SESSION["cnsconfig"]["robots"]+</meta>
<meta name="distribution" content="Global"/>
<link rel="shortcut icon" href="media/images/favicon.png"/>
<?php
isset($_REQUEST['admin']) || isset($_REQUEST['administration'])
? $scheme=$_SESSION["cnsconfig"]["theme"];
: $scheme=$_SESSION["cnsconfig"]["mobile"];
$path="media/themes/".$scheme."/mobile";
if($isMobile && !is_dir($path))
print "<meta name='viewport' content='width=device-width, initial-scale=1'>";
<?-- OpenGraph Meta -->
<meta property="og:title" content="+$_SESSION["cnsconfig"]["title"]+>
<meta property="og:description" content="+$_SESSION["cnsconfig"]["description"]+>
<meta property="og:type" content="+$_SESSION["cnsconfig"]["og:type"]+>
<meta property="og:image" content="+$_SESSION["cnsconfig"]["og:image"]+>
<meta property="og:url" content="+$_SESSION["cnsconfig"]["og:url"]+>
<?-- CSS Files -->
<script src="sys/jquery-1.11.1.min.js" type="text/javascript"></script>
<script src="sys/prop.js" type="text/javascript"></script>
<script src="lib/jquery/jquery.js" type="text/javascript"></script>
<script src="lib/jquery/jquery.js" type="text/javascript"></script>
<script src="lib/jquery/jquery.js" type="text/javascript"></script>
<link rel="stylesheet" type="text/css" href="sys/style.css">
```

Commencer



<https://doranum.fr/stockage-archivage/les-codes-sources-definition-enjeux-et-preservation> 10 13143 7 tj2-gw58/





WWW.PHDCOMICS.COM

**LIBER Data Management Plan Catalogue**  
Un catalogue de DMPs évalués dans différentes disciplines.  
Donnez votre avis à l'aide de ce questionnaire

<https://dmp.opidor.fr>

## Bienvenue !

DMP OPIDoR vous accompagne à travers l'élaboration et la mise en pratique de plans de gestion de données et de logiciels.



Accessible à la communauté scientifique de l'ESR et à ses partenaires français ou étrangers



Personnalisable par tout organisme de recherche pour la mise en place de sa politique de données



Enrichi par des exemples et des recommandations adaptés à l'environnement de recherche



Collaboratif : il facilite les échanges entre les partenaires d'un même projet et les services d'accompagnement

DMP OPIDoR évolue grâce à vos retours. Les développements s'inscrivent dans le cadre d'une collaboration internationale autour du logiciel open source DMPRoadmap

Connexion

S'inscrire

Accès institutionnel

Connexion

Accès individuel

\* Courriel

\* Mot de passe

[Mot de passe oublié ?](#)

Remember me

Connexion

# Logiciels d'aide à la rédaction de DMP

- DMProadmap <https://github.com/DMPRoadmap>
- DSW : <https://ds-wizard.org>
- Argos : <https://argos.openaire.eu/splash/>
- OpenDMP : <https://github.com/rhinoman/odmp>
- DMPtool <https://dmptool.org>
- DMPonline : <https://dmponline.dcc.ac.uk>
- EasyDMP <https://easydmp.no>
- DMP canvas generator <https://dmp.vital-it.ch/#/login>

Comparatif d'outils <https://zenodo.org/record/6513467#.Y-NZDa2ZM2x>

**Plan de gestion des données**

**PGD et principes FAIR**

**Présentation DMP Opidor**

**Trouver des exemples de PGD**

# L'adoption de bonnes pratiques : les DMP/PGD

## Le Plan de gestion des données (PGD) ou *Data Management Plan (DMP)*

Document qui décrit la façon dont les données seront :

- obtenues,
- traitées,
- organisées,
- stockées,
- sécurisées,
- préservées,
- partagées,
- etc.

au cours et à l'issue d'un projet.

- Comment la gestion et le partage des données sont-ils financés, en particulier à long terme ?

Ressources

- En quoi consiste le projet ?
- Qui sont les partenaires ?
- Quelle est la politique de gestion des données ?
- Qui est responsable de la gestion des données ?

Responsabilités dans le projet

- Quelles données seront produites/utilisées au cours du projet ? (type, format, volume et accroissement...).
- Comment seront-elles produites ou transformées ?

Collecte des données

- Comment les données seront-elles identifiées, décrites ?
- Quels standards de métadonnées utilisera-t-on ?
- Comment seront générées les métadonnées ?

Documentation des données

- Qui pourra accéder aux données ?
- Les données seront-elles publiées ?
- Comment ?
- Dans quel délai ?
- Sous quelle licence ?

Accès et partage des données

What's the plan?

D'après S. Cocaud et D. L'Hostis, « Pourquoi et comment rédiger un plan de gestion de données ? », 2019 [\(en ligne\)](#).

- Comment, où, par qui, seront stockées, sauvegardées et sécurisées les données ?

Sauvegarde des données

- Qui sera propriétaire des données produites ?
- Des données externes seront-elles utilisées ?

Propriété intellectuelle

- Des données sensibles seront-elles produites ou utilisées ?
- Comment sera assurée leur anonymisation ?

Ethique

- Quel est le plan d'archivage et de préservation à long terme ?

Archivage et préservation des données

## 10 sujets à aborder dans un bon PGD

1. Préparation du projet
2. Coûts liés au projet
3. Collecte et réutilisation des données
4. Documentation des données
5. Stockage et organisation
6. Accès aux données
7. Partage et publication
8. Archivage
9. Aspects éthiques et juridiques
10. Informations générales sur le projet et ses données



[10 Topics in a good DMP template,](#)  
Edugroepen, 2018

Répondre aux exigences de financeurs

Organiser et anticiper les étapes du cycle de vie de la donnée

Document évolutif (PGD projet comme PGD structure)

Définition des responsabilités, respect des principes FAIR

Evaluation des coûts <https://www.openaire.eu/how-to-create-a-data-management-plan>

**Description des données**

**Documentation et qualité des données**

**Sauvegarde et stockage des données au cours du projet**

**Exigences légales et éthiques**

**Stratégie de partage et d'ouverture des données**

**Préservation à long terme des données**

## Description des données

### 1. Description des données et collecte ou réutilisation de données existantes

#### 1.1 Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

1.1.a Les méthodologies et/ou logiciels utilisés pour produire ou recueillir de nouvelles données sont-ils explicités ?  Oui  En partie  Non  N/A

1.1.b L'origine et les conditions d'utilisation des données préexistantes réutilisées pour le projet sont-elles explicitées ?  Oui  En partie  Non  N/A

1.1.c Les raisons pour lesquelles il a été décidé de réutiliser ou de ne pas réutiliser de sources de données existantes sont-elles explicitées ?  Oui  En partie  Non  N/A

1.1.d La façon dont sera documentée la provenance de toutes les données est-elle explicitée ?  Oui  En partie  Non  N/A

#### 1.2 Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

1.2.a Trouve-t-on des détails sur le(s) type(s) de données utilisées (ex. : données numériques, textuelles, iconographiques...) ?  Oui  En partie  Non  N/A

1.2.b Trouve-t-on des détails sur le(s) format(s) des données utilisées ?  Oui  En partie  Non  N/A

1.2.c L'utilisation de ce(s) format(s) est-elle justifiée ? Le(s) format(s) de données choisi(s) sont-ils prioritairement des format(s) standards et/ou ouverts ?  Oui  En partie  Non  N/A

1.2.d Trouve-t-on des détails sur les volumes de données manipulés (espace de stockage requis, quantité de fichiers...) ?  Oui  En partie  Non  N/A

## Documentation et qualité des données

**2.1 Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?**

2.1.a Sait-on quelles informations (ex. : intitulé, date de production, nom du producteur...) seront créées ou récupérées pour décrire les données du projet et aider à leur recherche et à leur identification ?

2.1.b Comment ces métadonnées seront-elles récupérées ou créées ? Avec quels outils et contrôle qualité ?  Oui  En partie  Non  N/A

2.1.c Sait-on quels standards de métadonnées seront utilisés pour décrire ces informations ? Le choix de ces standards est-il expliqué ? Leur utilisation est-elle appropriée ?

Oui  En partie  Non  N/A

2.1.d Sait-on comment les données seront organisées au cours du projet (nommage, versionnage, classement...) ?

Oui  En partie  Non  N/A

2.1.e Trouve-t-on des détails sur la documentation qui accompagnera les données ? Facilitera-t-elle leur réutilisation ?

Oui  En partie  Non  N/A

**2.2 Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?**

2.2.a Trouve-t-on des détails sur la façon dont la qualité et la conformité de la collecte des données seront contrôlées et documentées ?

Oui  En partie  Non  N/A

## Sauvegarde et stockage des données au cours du projet

**3.1 Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?**

3.1.a Sait-on où les données seront stockées au cours du processus de recherche ?

Oui  En partie  Non  N/A

3.1.b Sait-on où, comment et à quel rythme les données seront sauvegardées au cours du processus de recherche ?

Oui  En partie  Non  N/A

3.1.c Le système de stockage et sauvegarde utilisé est-il approprié ?

Oui  En partie  Non  N/A

**3.2 Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?**

3.2.a Sait-on comment les données seront récupérées en cas d'incident ?

Oui  En partie  Non  N/A

3.2.b Sait-on qui aura accès aux données au cours du processus de recherche, et comment l'accès aux données sera contrôlé ?

Oui  En partie  Non  N/A

3.2.c Si certaines données sont sensibles, la question de leur protection spéciale est-elle abordée ?

Oui  En partie  Non  N/A

3.2.d Trouve-t-on des détails sur la politique institutionnelle de protection des données ?

Oui  En partie  Non  N/A

## Exigences légales et éthiques

**4.1 Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?**

4.1.a Sait-on si/comment sera obtenu le consentement éclairé des personnes pour la préservation et/ou le partage de leurs données personnelles ?

Oui  En partie  Non  N/A

4.1.b Trouve-t-on des détails sur une anonymisation ou une pseudonymisation des données personnelles ?

Oui  En partie  Non  N/A

4.1.c Sait-on si/comment les données personnelles seront chiffrées ?

Oui  En partie  Non  N/A

4.1.d Trouve-t-on des détails sur une procédure d'accès spécifique à ces données ?

Oui  En partie  Non  N/A

**4.2 Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?**

4.2.a Sait-on qui sera(-ont) le(s) propriétaire(s) des données ?

Oui  En partie  Non  N/A

4.2.b Trouve-t-on des détails sur la façon dont les données seront rendues accessibles ?

Oui  En partie  Non  N/A

4.2.c Des licences sont-elles prévues pour l'accès et la réutilisation des données ?

Oui  En partie  Non  N/A

4.2.d Des droits particuliers de propriété intellectuelle s'appliquent-ils aux données (ex. : droit des bases de données) ?

Oui  En partie  Non  N/A

4.2.e Si certaines données sont fournies par des tiers, leurs conditions de (ré-)utilisation sont-elles précisées ?

Oui  En partie  Non  N/A

4.2.f Dans le cas d'une propriété des données partagée entre plusieurs partenaires, les droits de chacun sont-ils précisés ?

Oui  En partie  Non  N/A

## Exigences légales et éthiques

**4.3 Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?**

4.3.a Sait-on si le projet de recherche implique des questions d'éthique spécifiques ?

Oui  En partie  Non  N/A

4.3.b Trouve-t-on des détails sur la façon dont les questions éthiques seront prises en compte dans la façon dont les données seront stockées et transférées ?

Oui  En partie  Non  N/A

4.3.c Les questions d'éthique impliquées dans le projet sont-elles abordées dans le respect des codes de conduite nationaux, internationaux et institutionnels ? Le contrôle de ces questions par une instance tierce (ex. : comité d'éthique) est-il nécessaire et effectif ?

Oui  En partie  Non  N/A

## Stratégie de partage et d'ouverture des données

**5.1 Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?**

5.1.a Sait-on comment les données pourront être retrouvées et mises à disposition ?

Oui  En partie  Non  N/A

5.1.b Trouve-t-on des détails sur la façon dont les données seront préservées, notamment sur leur durée d'archivage pérenne ?

Oui  En partie  Non  N/A

5.1.c Sait-on si et quand les données seront rendues disponibles ?

Oui  En partie  Non  N/A

5.1.d Trouve-t-on des détails sur d'éventuelles conditions exclusives de réutilisation des données ? Ces conditions sont-elles justifiées (ex. : préservation du secret industriel en attendant un dépôt de demande de brevet) ?

Oui  En partie  Non  N/A

5.1.e Sait-on qui sera en mesure de réutiliser les données ?

Oui  En partie  Non  N/A

**5.2 Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un entrepôt de données ou une archive) ?**

5.2.a Sait-on quelles données ne devront pas être divulguées ou devront être détruites à la fin du projet ? Sait-on pourquoi ?

Oui  En partie  Non  N/A

5.2.b Sait-on quelles données seront sélectionnées pour être conservées sur le long terme, et comment s'opérera cette sélection ?

Oui  En partie  Non  N/A

5.2.c Trouve-t-on des exemples possibles de réutilisation des données dans un cadre de futures recherches ?

Oui  En partie  Non  N/A

5.2.d Trouve-t-on des détails sur le mode de préservation ou l'entrepôt choisi pour l'archivage pérenne des données ? Le mode de préservation ou l'entrepôt choisi est-il approprié ?

Oui  En partie  Non  N/A

## Stratégie de partage et d'ouverture des données

**5.3 Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?**

5.3.a Sait-on si les utilisateurs potentiels auront besoin d'outils ou de logiciels spécifiques pour accéder et/ou réutiliser les données ? Trouve-t-on des détails précis (version, système d'exploitation...) sur ces logiciels ?  Oui  En partie  Non  N/A

5.3.b Les outils et logiciels utilisés pour créer, traiter ou analyser les données sont-ils rendus disponibles en même temps que les données ?  Oui  En partie  Non  N/A

5.3.c Si les données sont partagées via un entrepôt, sait-on comment seront traitées les demandes d'accès aux données ?  Oui  En partie  Non  N/A

**5.4 Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?**

5.4.a Sait-on si les données se verront attribuer des identifiants pérennes ? De quel(s) type(s) ?  Oui  En partie  Non  N/A

5.4.b Trouve-t-on des détails sur la façon dont les données se verront attribuer des identifiants pérennes ?  Oui  En partie  Non  N/A

## Préservation à long terme des données

**6.1 Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?**

6.1.a Trouve-t-on des détails sur la façon dont seront répartis les rôles et responsabilités concernant les activités de gestion des données ? Des responsables sont-ils nommément identifiés ?

Oui  En partie  Non  N/A

6.1.b Trouve-t-on des détails sur la coordination des responsabilités de gestion des données entre partenaires ?

Oui  En partie  Non  N/A

6.1.c Sait-on qui est responsable de la mise à jour et de la mise en œuvre du PGD ?

Oui  En partie  Non  N/A

6.1.d Sait-on à quel rythme le PGD sera mis à jour ?

Oui  En partie  Non  N/A

**6.2 Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?**

6.2.a A-t-on des informations sur le temps humain qui sera dédié à la gestion des données (planification, collecte, traitement, stockage, analyse, mise à disposition etc.) ? Ces tâches de gestion donneront-elles lieu à un recrutement spécifique ?

Oui  En partie  Non  N/A

6.2.b A-t-on des informations sur les frais spécifiques à engager pour la gestion des données pendant le projet (achat d'accès à des données, commande de prestations de collecte de données, achat ou location de matériel de stockage, achat de licences logiciels pour le traitement...)?

Oui  En partie  Non  N/A

6.2.c A-t-on des informations sur le coût de préparation des données en vue de leur mise à disposition (anonymisation, reformatage, création de métadonnées...)?

Oui  En partie  Non  N/A

6.2.d A-t-on des informations sur le coût de l'archivage pérenne des données ?

Oui  En partie  Non  N/A

## Stratégie d'ouverture et de conservation des données

### PARTAGE

Entrepôt

Entrepôt généraliste ou disciplinaire

L'équipe de recherche

Autre équipe de recherche (même domaine)

Autre équipe de recherche (autre domaine)

Le partage consiste à déposer vos données dans un entrepôt de données afin de les rendre accessibles facilement et de permettre leur réutilisation (principes FAIR) sur le court et le moyen terme (5 à 10 ans), pour des chercheurs du même domaine, ou d'un autre domaine.

**Comment faire ?**

- Sélectionner les données à partager.
- Estimer la volumétrie des données et éventuellement le budget.
- Vérifier la compatibilité et l'interopérabilité des formats de données en se posant par exemple ces questions :
  - Le contenu est-il valorisé ou toujours valorisable ?
  - Le fichier et son contenu sont-ils compréhensibles en l'état ?
  - Le format de fichier utilisé nécessite-t-il un logiciel pour être lisible ?
- Migrer si nécessaire vers un format adapté, le plus ouvert possible.
- Préparer si nécessaire les codes sources (ex : scripts) qui permettront de lire et traiter les données.
- Compléter et enrichir les métadonnées (en fonction de l'entrepôt choisi).
  - Si ce n'est pas déjà fait, choisir un standard de métadonnées.
  - S'il n'en existe pas d'adapté, créer un schéma de métadonnées.
  - Compléter les champs pour chaque jeu de données, suivant le standard adopté.



**LIBER Data Management Plan Catalogue**  
Un catalogue de DMPs évalués dans différentes disciplines.  
Donnez votre avis à l'aide de ce questionnaire

<https://dmp.opidor.fr>

## Bienvenue !

DMP OPIDoR vous accompagne à travers l'élaboration et la mise en pratique de plans de gestion de données et de logiciels.



Accessible à la communauté scientifique de l'ESR et à ses partenaires français ou étrangers



Personnalisable par tout organisme de recherche pour la mise en place de sa politique de données



Enrichi par des exemples et des recommandations adaptés à l'environnement de recherche



Collaboratif : il facilite les échanges entre les partenaires d'un même projet et les services d'accompagnement

DMP OPIDoR évolue grâce à vos retours. Les développements s'inscrivent dans le cadre d'une collaboration internationale autour du logiciel open source DMPRoadmap

Connexion

S'inscrire

Accès institutionnel

Connexion

Accès individuel

\* Courriel

\* Mot de passe

[Mot de passe oublié ?](#)

Remember me

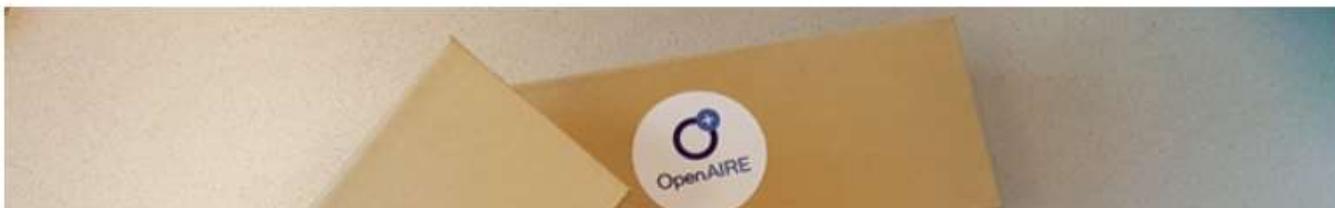
Connexion



🕒 3 MINUTES READING TIME (599 WORDS)

## Establishing a collection of 841 publicly available Horizon 2020 Data Management Plans

📁 scholarly communication | 👤 Gerda McNeill | 🕒 22 Jan 2021 | ⓘ 2227 Hits | 💬 0 Comments



Source :

<https://www.openaire.eu/blogs/establishing-a-collection-of-841-horizon-2020-data-management-plans>

# Exemples de PGD rédigés

Nom du DMP	Version	Sujet	Nombre de page du DMP	Lien
	Version finale		53	<a href="https://services.phaidra.univie.ac.at/api/object/o:1139874/diss/Content/get">https://services.phaidra.univie.ac.at/api/object/o:1139874/diss/Content/get</a>
<b>MELOA Data management plan</b>	Version 1	Appareil d'océanographie extra léger polyvalent/multi-capteur	13	<a href="https://services.phaidra.univie.ac.at/api/object/o:1139227/diss/Content/get">https://services.phaidra.univie.ac.at/api/object/o:1139227/diss/Content/get</a>
<b>OPERA Data management plan</b>	Première version	Une expérience d'exploitation en pleine mer pour réduire le coût de l'énergie houlomotrice	30	<a href="https://services.phaidra.univie.ac.at/api/object/o:1139216/diss/Content/get">https://services.phaidra.univie.ac.at/api/object/o:1139216/diss/Content/get</a>
	Version finale		29	<a href="https://phaidra.univie.ac.at/detail/o:1139215#?q=%22OPERA%20Data%20management%20plan%20(final%20version)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797">https://phaidra.univie.ac.at/detail/o:1139215#?q=%22OPERA%20Data%20management%20plan%20(final%20version)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797</a>
<b>Open4Citizens Data management Plan</b>	Version initiale (ébauche)	Donner aux citoyens les moyens d'utiliser utilement les données ouvertes	15	<a href="https://phaidra.univie.ac.at/detail/o:1139705#?q=%22Open4Citizens%20Data%20Management%20Plan%20(Draft)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797">https://phaidra.univie.ac.at/detail/o:1139705#?q=%22Open4Citizens%20Data%20Management%20Plan%20(Draft)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797</a>
	Version mi-parcours		20	<a href="https://phaidra.univie.ac.at/detail/o:1139704#?q=%22Open4Citizens%20Data%20Management%20Plan%20(Midterm)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797">https://phaidra.univie.ac.at/detail/o:1139704#?q=%22Open4Citizens%20Data%20Management%20Plan%20(Midterm)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797</a>
	Version finale		35	<a href="https://phaidra.univie.ac.at/detail/o:1139446#?q=%22Open4Citizens%20Data%20Management%20Plan%20(Final)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797">https://phaidra.univie.ac.at/detail/o:1139446#?q=%22Open4Citizens%20Data%20Management%20Plan%20(Final)%22&amp;page=1&amp;pagesize=10&amp;collection=o:1140797</a>

Source : Candice Fillaud,

Mémoire M1 EISO 2021

# Des DMP sur HAL

Déposez sur HAL

× Quitter

Choisir le type de document qui correspond à votre publication. L'icône '>' indique que des sous-types de documents sont disponibles et peuvent être sélectionnés ⓘ

×

 Article dans une revue >	 Brevet	 HDR	 Carte
 Communication dans un congrès	 Autre publication scientifique		
 Poster de conférence	 Pré-publication, Document de travail >		
 Ouvrages >	 Rapport >		
 Chapitre d'ouvrage	 Thèse		

Rapport de recherche

Rapport technique

Rapport contrat/projet

Rapport d'expertise collective

Plan de gestion de données/data management plan

## Bienvenue dans le patchwork de construction d'un PGD

**Un parcours d'accompagnement de bout-en-bout pour construire son Plan de Gestion des Données**

Bâti autour du cycle de vie de la donnée, vous y trouverez un **patchwork de savoir-faire, d'outils et de documentations** tissé pour construire votre PGD.

<https://opendatacanvas.org/sprint-pgd>

# Pour aller plus loin

Gestion et partage des

## Données de la recherche

Dans le cadre du mouvement de l'Open Science, la question de la gestion et du partage des données de la recherche devient centrale dans le montage de projets de recherche. De plus en plus d'organismes de financement public conditionnent le versement de leurs subventions à la diffusion de ces données. Afin d'y répondre, les porteurs de projets doivent réaliser des plans de gestion de données.



Définition et enjeux



Gestion des données de la recherche



Diffusion des données de la recherche



Ré-utilisation des données de la recherche

📅 Juillet 2022

🕒 2h de lecture

👤 Chloée Fabre

*Responsable Diffusion et valorisation de la production académique - Université*

*Toulouse Jean Jaurès*

📄 BY\_NC-SA 4.0



<https://callisto-formation.fr/course/view.php?id=126>

**Introduction aux données de la recherche et atelier DMP**

24/03/2023