



UNIVERSITÉ DE
MONTPELLIER

Pourquoi rédiger un Plan de Gestion de Données ?

Webinaire du 15 mars 2022

Equipe d'Appui à la Science Ouverte de l'UM



Objectifs et plan schématique de la formation

Objectif :

Comprendre ce qu'est un PGD et l'intérêt d'en remplir un dans le cadre de son projet de recherche

Axes du webinaire :

1. Intro de contextualisation
2. Le PGD : définition, objectifs, modèles et contenu
3. Obligations contractuelles
4. Quels outils pour le rédiger ?
5. Où trouver des PGD ? Que faire du mien ?
6. D'autres types de PGD

I. La science ouverte : définition et contexte

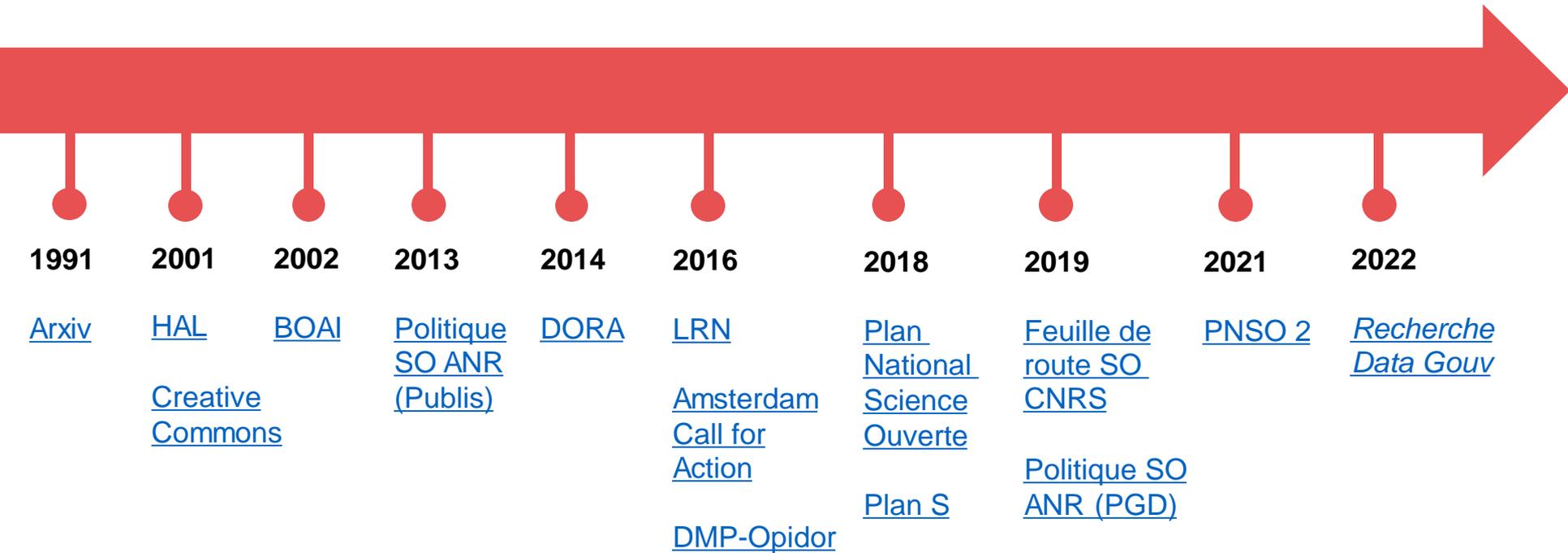
La science ouverte est la diffusion sans entrave des publications et des données de la recherche. Elle s'appuie sur l'opportunité que représente la mutation numérique.

https://www.ouvrirlascience.fr/category/science_ouverte/

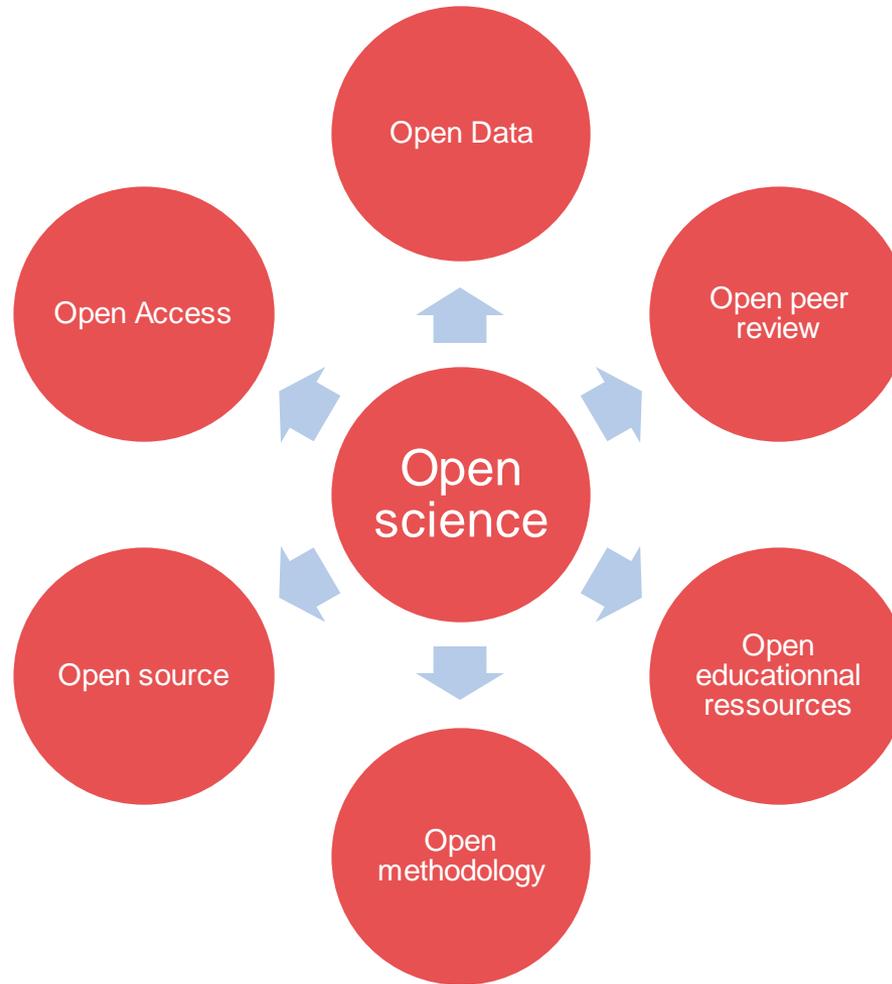
La science ouverte est la pratique de la science qui rend possible la collaboration et la contribution avec d'autres, où les données de recherche, les notes de laboratoire et autres processus de recherche sont librement disponibles, dans des conditions permettant la réutilisation, la redistribution et la reproduction de la recherche, ainsi que de ses données et méthodes sous-jacentes.

<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>

I. La science ouverte : développement



I. La science ouverte : périmètre



I. L'Open Data : définition

L'Open Data est une démarche de communication des documents ou données publics, afin qu'ils soient diffusés de manière structurée selon une méthode garantissant leur libre accès et leur réutilisation par tous, sans restriction technique, juridique ou financière injustifiée.

Source : [Ouvrir la science](#) « Ouverture des données de recherche : Guide d'analyse du cadre juridique en France », 2021, MESRI

I. L'Open Data : à noter

Les données de la recherche financée sur fonds publics sont des données publiques dont le principe est l'ouverture.

Néanmoins, des **restrictions juridiques et éthiques** s'imposent notamment pour les données personnelles, les données sensibles, les données protégées par le droit d'auteur.

En revanche, la communication des données relatives aux secrets professionnels, au secret de la défense nationale, à la sûreté de l'Etat et à la sécurité publique est interdite par principe.

I. Données de la recherche : définition

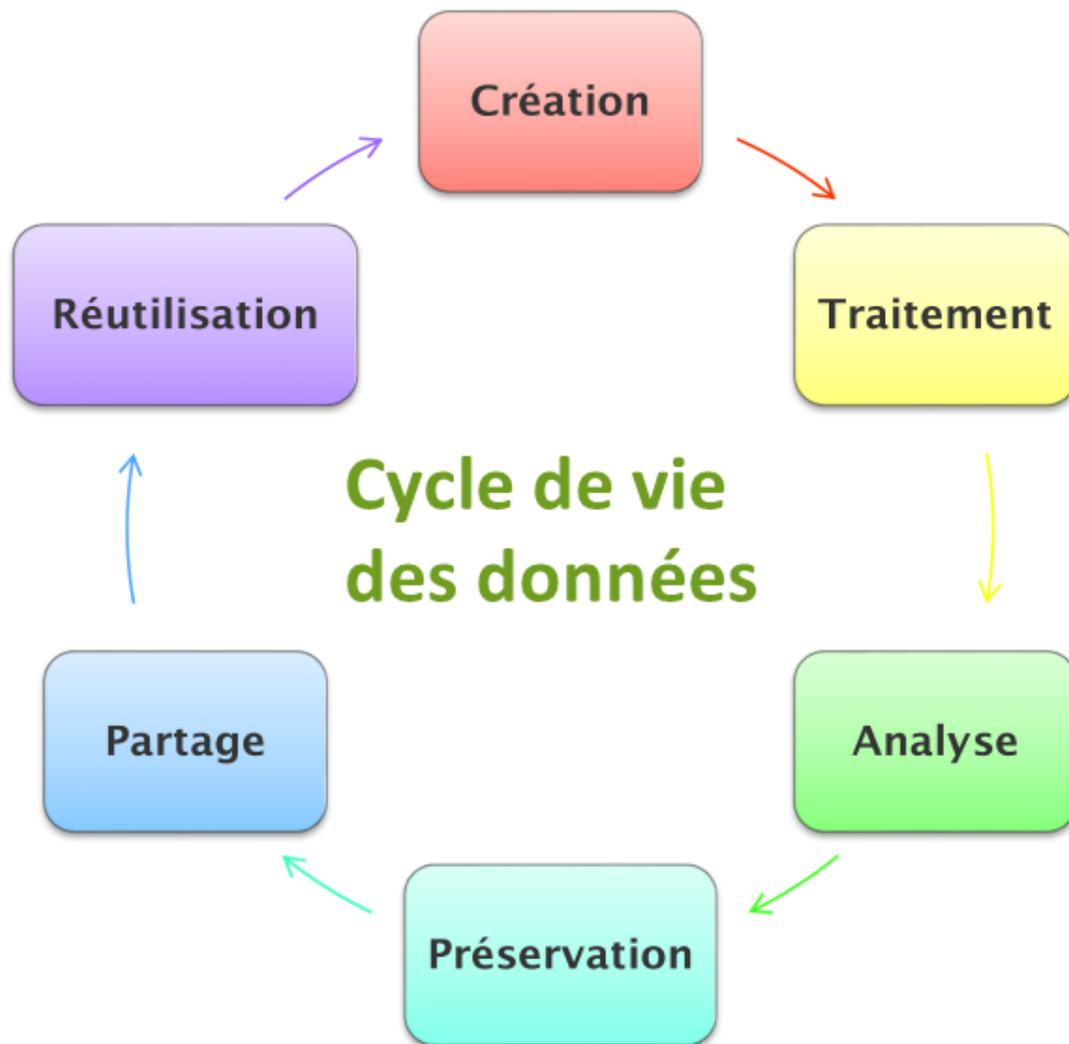
« Les données de la recherche sont définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la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

Définition de l'OCDE, rapport 2017 - *Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics*

Elles peuvent avoir différents :

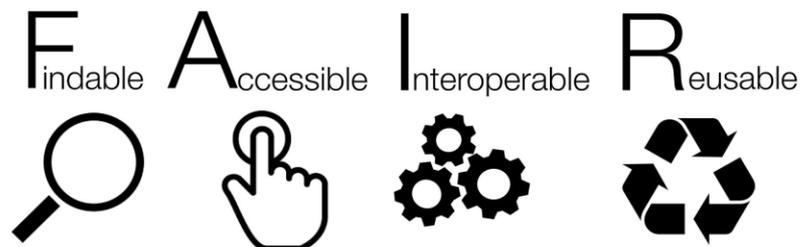
- **supports** : carnets de laboratoire, documents électroniques, logiciels, programmes informatiques...
- **types** : archives, audio, vidéo, bases de données, codes sources, données géospatiales, images, photographies, numérisations, scans, qualitatives, quantitatives, statistiques...
- **natures** : brutes, expérimentales, produites ou collectées, issues d'observation ou de simulation...

I. Données de la recherche : cycle de vie



Source : <https://data.ird.fr/utiliser-des-donnees/>

I. Données de la recherche et principes FAIR



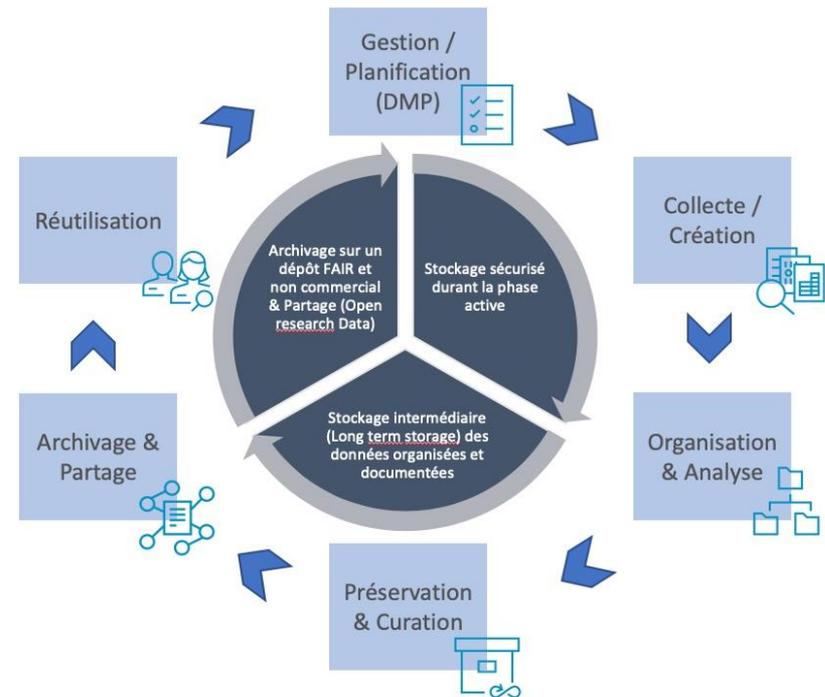
- **Faciles à trouver** : dotées d'identifiants uniques et pérennes (ex. DOI), décrites par des métadonnées précises, déposées dans un entrepôt de données
- **Accessibles** : récupérables par des protocoles de communication standardisés, libres et ouverts, avec authentification en cas de données sensibles
- **Interopérables** : décrites avec un vocabulaire contrôlé permettant l'interopérabilité et les liens entre les données
- **Réutilisables** : publiées avec une licence d'utilisation claire et accessible, avec indication de leur provenance

— II. Le PGD / DMP : définition

Le PGD ou Data Management Plan (DMP) est un document formalisé rédigé de préférence au démarrage d'un projet de recherche qui décrit la façon dont les données seront produites ou obtenues, (ré)utilisées, traitées, organisées, stockées, sécurisées, préservées, documentées, partagées, au cours et à l'issue d'un projet, explicitant notamment leur mise à disposition.

— II. Le PGD / DMP : objectifs

- Le PGD aide à mettre en place de **bonnes pratiques de gestion des données**. Il interroge le chercheur sur ses pratiques à chaque étape du cycle de vie des données.
- Il vise à garantir **des données fiables** tout au long du projet, **compréhensibles, disponibles et préservées** sur le long terme pour une réutilisation future.



Source : [Université de Lausanne](#)

— II. Le PGD / DMP : modèles

Le PGD est un **document formel** qui décrit dans le détail les méthodes et processus de création, de fourniture, de maintenance, de conservation et de protection de données.

Ainsi, de nombreux modèles ont été réalisés par :

- les **financeurs** : Science Europe, ANR...
- les **institutions** : l'université de Montpellier, l'INRAE...

Ces différents modèles existent en français et/ou en anglais.

— II. Le PGD / DMP : contenu

Le fond reste sensiblement identique entre tous ces différents modèles (tous les aspects du cycle de vie des données sont interrogés).

On y retrouve les mêmes éléments :

- Informations administratives et **responsabilités**
- Description des **données** et des **métadonnées**
- **Sauvegarde et partage** des données **durant le projet**
- **Stockage et partage** des données **après le projet**
- **Archivage** pérenne
- **Aspects juridiques et éthiques**

Les différences vont porter sur la **forme** (il y aura plus ou moins de rubriques et de questions) et sur les **recommandations** (qui sont ajoutées en complément du modèle pour aider à la rédaction).

— III. Le PGD / DMP : obligations contractuelles

C'est un **livrable obligatoire** de plus en plus demandé par les financeurs (**Horizon Europe, ANR depuis 2019**) pendant les projets de recherche, voire en amont, pendant le montage du projet.

Il existe plusieurs versions pour un PGD :

- La **V1** six mois après le début du projet
- La **V2** intervient à mi-parcours (pour les projets de plus de 30 mois)
- La **V3** à la fin du projet

— III. Le PGD / DMP : obligations contractuelles

Les différentes versions du PGD sont à déposer sur la plateforme [Coheris](#) : la plateforme usuelle de l'ANR.

A noter : pour les projets ANR 2021 et suivants, le **dernier versement financier** est conditionné par le dépôt de la V3.

Perspectives à court terme : vers une **évaluation des PGD** pour les projets européens et ARN dans les années à venir ?

— III. Le PGD / DMP : conclusion

En plus d'être un document administratif, on peut également voir un PGD comme un **atout** dans la construction et le développement **de son projet de recherche** :

- Il offre la possibilité de mettre en place de nouvelles pratiques en adéquation avec les **principes FAIR** pour gérer ses données
- C'est un document évolutif et collaboratif qui vise à devenir un **automatisme** pendant les projets de recherche
- C'est un document qui peut être **publié** suite au projet de recherche (dans un entrepôt, une revue et bientôt sur HAL)

— IV. Le PGD / DMP : quels outils pour les rédiger ?

Il existe des outils utiles pour simplifier la réalisation d'un PGD : les **formulaire en ligne**. Ils permettent de minimiser et de faciliter le travail de rédaction tout en assurant une certaine qualité des informations (notamment l'assurance que les données exigées par les financeurs sont bien présentes).

Voici les **trois principaux outils gratuits et en ligne** :

- DMP Opidor
- Argos
- DSW



— IV. Le PGD / DMP : quels outils pour les rédiger ?

DMP-Opidor (**DMP** pour une **O**ptimisation du **P**artage et de l'**I**nteropérabilité des **D**onnées de la **R**echerche)

Il permet de créer en ligne (et de partager) un PGD en fonction des exigences d'un financeur ou d'un organisme.

Ce service, mis à disposition par l'Inist-CNRS, est basé sur l'outil **DMPOnline**. Il est recommandé par l'ANR et est actuellement le plus utilisé par la communauté scientifique française.

Enfin, c'est un outil en perpétuelle évolution pour s'adapter aux besoins des chercheurs (V3 en novembre 2021).



— IV. Le PGD / DMP : quels outils pour les rédiger ?

Argos est un service intégré dans la plate-forme OpenAIRE.

Il s'appuie sur des processus automatisés pour la création, la gestion, le partage et le lien entre les plans de gestion de données et les données de la recherche auxquels ils correspondent.

Les PGD rédigés sont exploitables par les machines.

The logo for Argos, featuring the word "argos" in a lowercase, teal-colored, rounded sans-serif font.

— IV. Le PGD / DMP : quels outils pour les rédiger ?

DSW (Data Stewardship Wizard)

C'est un outil en Open Source développé par une équipe européenne.

Il possède quelques atouts intéressants :

- Historique des démarches
- Filtre des questions
- Alerte pour les éléments non finalisés ou à valider
- Possibilité d'avoir une trame individuelle prédéfinie



— V. Le PGD / DMP : Où trouver des PGD ? Que faire du mien ?

Il existe plusieurs ressources pour rendre disponible ou pour trouver un PGD :

- Les **entrepôts généralistes** (comme [Zenodo](#))
- Les **entrepôts spécialisés** ([DMP-Opidor](#), [DMP Tool](#), [DMP Online](#), [l'université de Vienne...](#))
- [HAL](#) (fonctionnalité disponible courant 2022)
- Le [Research Ideas and Outcomes](#) (RIO) Journal

— VI. Le PGD / DMP : les autres types

En dehors des PGD réalisés dans le cadre des projets européens et ANR, on note l'utilisation de PGD pour d'autres usages, voire d'autres échelles :

- Les **PGD projets** pour les doctorants
- Les **PGD structure**
- Les **PGD jeux de données**
- Les **Plans de Gestion de Logiciels** ([PGL](#))

Webinaires à venir

Créer son IdHAL et son CV dans HAL en lien avec ORCID

5 avril 2022 de 11h à 12h

Peer Community In : Une alternative gratuite pour évaluer, valider (et publier ?) les preprints

19 avril 2022 de 13h à 14h

Pour se tenir informé des webinaires thématiques bimensuels, c'est [ici](#) !

Merci pour votre écoute !

Gardons contact pour la suite :

- Pour des questions autour des publications ou de l'utilisation de HAL : hal-assistance@umontpellier.fr
- Pour des questions autour de la gestion des données de la recherche ou pour toute relecture de PGD : donnees-recherche@umontpellier.fr
- Pour toute autre question : laure.lefrancois@umontpellier.fr

Chargée de mission Appui à la recherche pour l'université

Webographie récapitulative

- [Dorandum](#) (le site incontournable d'auto-formation pour les données de recherche)

Pour en savoir plus sur :

- les [pratiques FAIR](#)
- les [PGD](#) (plan de gestion de données)
- les [PGL](#) (plan de gestion de logiciels)
- le [cycle de vie des données](#)
- [Deuxième Plan national pour la science ouverte](#) (MESRI)
- [La feuille de route 2021-2024 du MESRI sur la politique des données, des algorithmes et des codes sources](#)
- [Plan d'action 2022 de l'ANR](#)
- [Parcours interactif sur la gestion des données de la recherche](#) (Dorandum)
- [Conseils pour l'organisation des données](#) (Dorandum)
- [Trouver des exemples de PGD](#) (INRAE)