

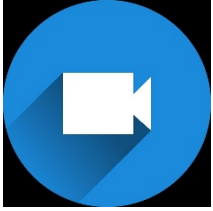
# Rechercher, gérer et valoriser les données pour sa thèse

---

Annabelle Filatre (AgroParisTech)

Hanka Hensens (IRD)

# Plan de Gestion de Données / Data Management Plan (PGD/DMP)



Une petite video pour commencer (Minute DMP Doranum) :

<https://www.youtube.com/embed/YFxyKRNv5QU>

En résumé :

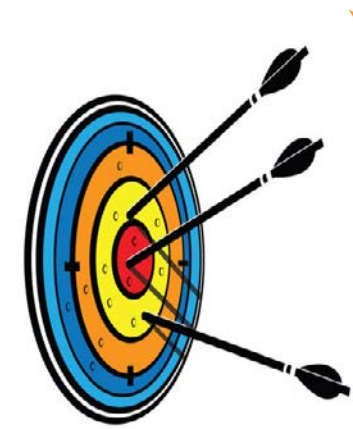
## Un PGD, qu'est-ce que c'est ?

- ✓ Un document formalisé (type formulaire)
- ✓ Vue globale sur l'ensemble des données produites ou utilisées dans le projet
- ✓ Liste toutes les questions à se poser
- ✓ A toutes les étapes du cycle de vie des données
- ✓ Donne des réponses spécifiques, précises, réalistes, budgétées et chiffrées
- ✓ A commencer le plus tôt possible (dans l'idéal au démarrage d'un projet)



# Objectifs

- ✓ Aide à la mise en place de bonnes pratiques de gestion des données
- ✓ Pour des données plus fiables, compréhensibles, disponibles durant le projet
- ✓ Préservées si elles le méritent pour une réutilisation future



 publicdomainvectors.org

- 1. Décrire les données → **Assurer la reproductibilité des expériences**
- 2. Clarifier le cadre juridique et éthique → **Respecter les droits et personnes**
- 3. Garantir la compréhension → **Permettre la réutilisation des données**
- 4. Assurer le stockage → **Eviter les pertes**
- 5. Définir les responsabilités → **Assurer la bonne gestion des données**
- 6. Spécifier les modalités de partage → **Clarifier les droits de réutilisation**

## Les PGD ont toujours les mêmes rubriques



D'après [Fiche synthétique Doranum](#)

1. **Description du projet de recherche** : but, partenaires, responsabilités, leadership,...
2. **Inventaire des données** : type, taille, méthode de collecte et traitement,...
3. **Listing des « métadonnées »** qui accompagneront les données: fichiers complémentaires, logiciel,...
4. **Accès et sauvegarde durant le projet** : matériel et méthodes, périodicité et nombre, versioning, nommage,...
5. **Droit et déontologie** : selon le type de données (personnelles, santé, sensibles,...), les lois,...
6. **Partage pendant et après le projet** : restrictions d'accès (quoi et pourquoi), licences, embargo, confidentialité,...
7. **Conservation à long terme** : quelles données, combien de temps, à quel endroit
8. **Personnes physiques ou morales chargées des données** aux différentes étapes de leur cycle de vie
9. **Ressources nécessaires à la gestion des données** : matérielles, financières, humaines, ...

# Contenu : questions à se poser

- Quels sont les besoins générés par la gestion des données ?
- Comment seront-ils financés, à court et 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, où, par qui, seront stockées, sauvegardées et sécurisées les données au cours du projet ?

## Sauvegarde des données

- Qui pourra accéder aux données ? Quelles données seront partagées ? publiées ?
- Avec qui ?
- Comment ?
- Dans quel délai ?
- Sous quelle licence ?

## Accès et partage des données



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

## Documentation des données

- Quel plan pour l'archivage et la préservation à long terme ?
- Pour quelles parties des données ?

## Archivage et préservation des données

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

## Ethique

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

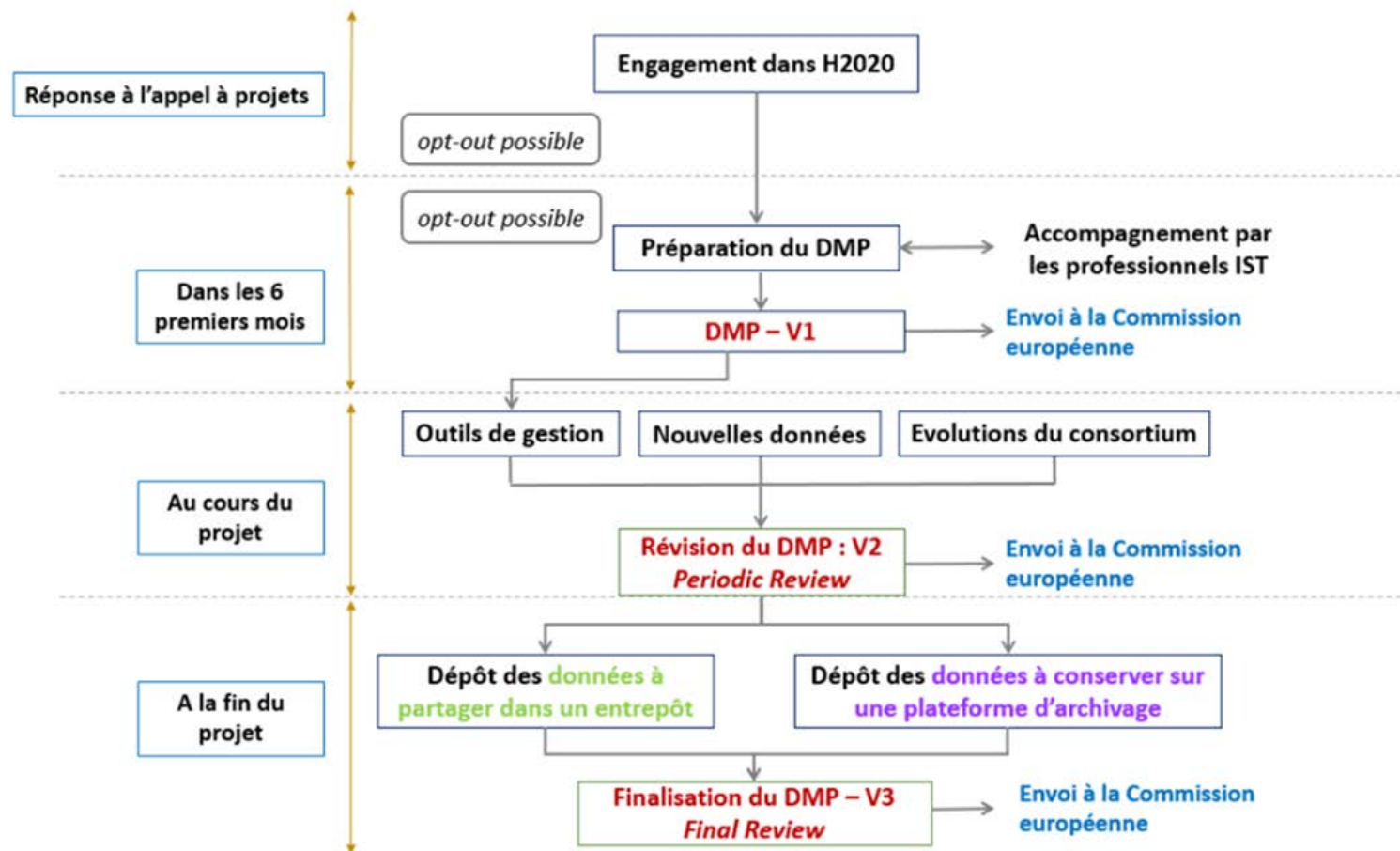
## Propriété intellectuelle

# Un document évolutif

Un PGD n'est pas figé :

- il doit être mis à jour tout au long du projet
- Il comporte des versions successives (certaines pouvant être des livrables)
  - *Exemple H2020 : 3 versions minimum (pas de maximum)*

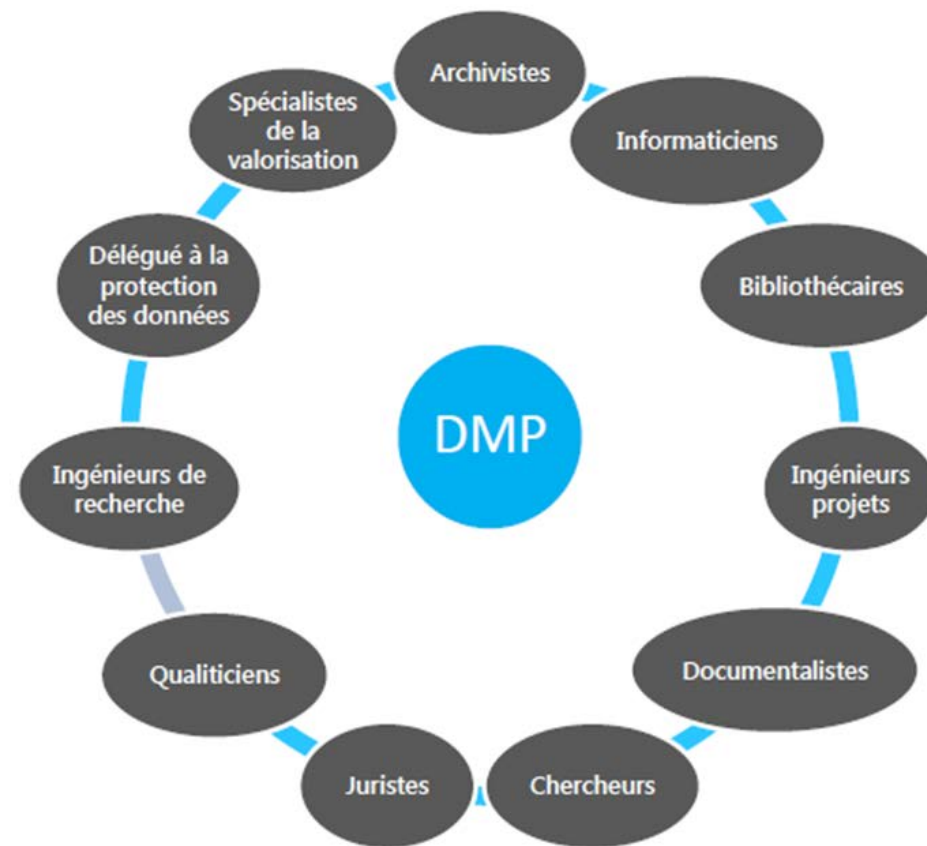
→ Pas d'obligation de répondre à toutes les questions dans la V1



A. Cartier, M. Moysan, N. Reymonet, d'après les *Guidelines on FAIR data management*

Participent à la gestion des données et donc à la rédaction du PGD :

- **Scientifiques** : description, validation
- **Ingénieurs-données** : sauvegarde, qualité
- **Informaticiens** : sécurisation, infrastructure
- **Spécialistes IST** : modèles de PGD, standards de métadonnées
- **Archivistes** : données à conserver, durées et solutions techniques
- **Juristes** : propriété intellectuelle, législation, données personnelles, sensibles,...
- **Chargés de valorisation** : rédaction des conventions à signer
- **TOUS LES PARTENAIRES** producteurs ou fournisseurs des données



Reymonet, N., Moysan, M., Cartier, A., & Délémontez, R. (2018). Réaliser un plan de gestion de données « FAIR »



# Comment ? Suivre les conseils de sites de référence

- Datapartage INRAE, en particulier : <https://datapartage.inrae.fr/Gerer/Rediger-un-plan-de-gestion>
- IRD Data, en particulier : <https://data.ird.fr/gerer/>
- Coopist CIRAD, en particulier : <https://coop-ist.cirad.fr/gerer-des-donnees/rediger-un-pgd/1-qu-est-ce-qu-un-pgd>
- Doranum, en particulier : <https://doranum.fr/plan-gestion-donnees-dmp/>
- Data Curation Center (DCC), toute la rubrique « resources », dont : <http://www.dcc.ac.uk/resources/data-management-plans>

# Comment ? Utiliser des ressources spécialisées

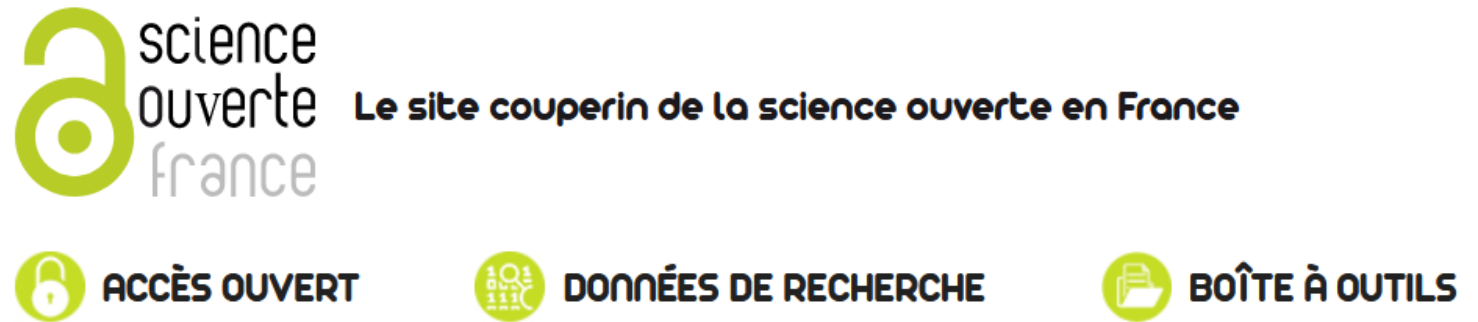
Fichiers	Métadonnées	Entrepôts	Licences
<ul style="list-style-type: none"><li>• <u>Types de fichier</u> (Datapartage INRAE)</li><li>• <u>Formats de fichiers ouverts/fermés</u> (Dorandum)</li><li>• <u>Nommage et organisation</u> (Datapartage INRAE)</li></ul>	<p><b>Choix de standard de métadonnées</b> par thématique sur :</p> <ul style="list-style-type: none"><li>• <u>DCC</u>,</li><li>• <u>RDA</u>,</li><li>• <u>Fairsharing</u></li></ul>	<ul style="list-style-type: none"><li>• <u>Quelles données conserver ?</u> (DCC)</li><li>• <b>Choix d'un entrepôt sur <u>re3data.org</u></b> (<u>Browse by subject</u>)</li></ul>	<ul style="list-style-type: none"><li>• <u>Mes données sont-elles partageables ?</u> Aides à la décision (IRD Data)</li><li>• <u>Choix d'une licence</u> : ODBL et Etalab (Datapartage INRA)</li><li>• <u>Choix d'une licence Creative Commons</u></li></ul>

- Pour les méthodologies, les formats, les logiciels, les entrepôts,...
- Recommandés par son organisme, son pays, ses partenaires, ses collègues de la même thématique scientifique...
- A défaut, Science Europe a établi des recommandations facilitant la rédaction et le contrôle des PGD grâce à l'établissement de protocoles de données par domaine (DDP) :

Science Europe. (2018). [Science Europe Guidance. Presenting a Framework for Discipline-specific Research Data Management.](#)



- <https://scienceouverte.couperin.org/sos-pgd/>



Vous êtes ici : [Science ouverte France](#) » [Données de recherche](#) » [SOS-PGD](#)

## SOS-PGD

Le répertoire des Services Opérationnels de Soutien à la rédaction des Plans de Gestion des Données (SOS-PGD) recense les services accompagnant la rédaction des plans de gestion des données au sein des établissements d'enseignement supérieur et de la recherche. Il vise à aider les chercheurs à identifier leurs interlocuteurs au sein de leur institution et à faciliter la mise en relation entre les services supports de différentes institutions pour les projets de recherche multi-partenariaux.

Si vous souhaitez que votre service soit ajouté à ce répertoire, veuillez remplir ce [formulaire](#). Pour modifier, rectifier ou supprimer des informations, veuillez écrire à [gtso\\_donnees@listes.couperin.org](mailto:gtso_donnees@listes.couperin.org)

Répertoire mis en ligne le 30 mars 2020. Dernière mise à jour le vendredi 26 juin 2020.

Institution	Service	Alias	Référent	Adresse mél
Aix Marseille Université	Service Commun de la Documentation	<a href="mailto:scd-dmpopidor-assistance[at]univ-amu.fr">scd-dmpopidor-assistance[at]univ-amu.fr</a>	Caroline Sophie Donati	<a href="mailto:caroline-sophie.donati[at]univ-amu.fr">caroline-sophie.donati[at]univ-amu.fr</a>

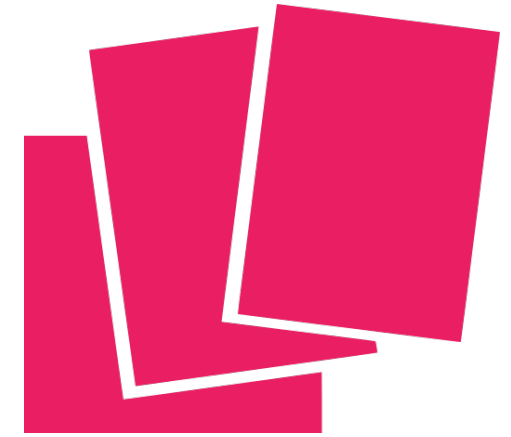
## de financeurs :

- [ANR](#)
- [H2020](#)
- [Welcome Trust](#)

## D'organismes de recherche :

- [INRAE](#)
- [CIRAD](#)
- [National Science Fondation](#) (USA)





- **DMP Opidor** : <https://dmp.opidor.fr/static/examples> et [https://dmp.opidor.fr/public\\_plans](https://dmp.opidor.fr/public_plans)
- **DCC** : <http://www.dcc.ac.uk/resources/data-management-plans/guidance-examples>
- **DMPTool** : [https://dmptool.org/public\\_plans](https://dmptool.org/public_plans)
- **Data-partage Inrae** : <https://datapartage.inrae.fr/Generer/Rediger-un-plan-de-gestion/Exemples-de-plans>
- **Journal *Research Ideas and Outcomes* (RIO), Section DMP** : [https://riojournal.com/browse\\_journal\\_articles.php?form\\_name=filter\\_articles&sortby=0&journal\\_id=17&search\\_in=0&section\\_type\[\]=231](https://riojournal.com/browse_journal_articles.php?form_name=filter_articles&sortby=0&journal_id=17&search_in=0&section_type[]=231)



Pour :

- ✓ Rédiger un PGD conforme aux demandes d'un financeur ou d'un organisme
  - de nombreux modèles au choix
  - aides et exemples spécifiques en ligne
- ✓ Partager un PGD
  - avec des collaborateurs identifiés aux droits définis (copropriétaire / éditeur / lecteur) ;
  - avec toutes les personnes de son organisme ;
  - avec tout le monde (publiquement).
- ✓ Télécharger votre PGD dans différents formats (pdf, doc,...), à plusieurs reprises (versions)

