



Plan de gestion des données

PGD / DMP

Data Managment Plan

Laurence Rageot et Anne Garcia-Fernandez

Définition d'un DMP ou PGD

Le DMP (Data management plan) ou PGD (plan de gestion de données) est un **outil de gestion dédié aux données de la recherche**.

Il explique comment sont gérées les données, depuis leur création ou collecte, jusqu'à leur partage et leur archivage. Il aide à organiser et anticiper toutes les étapes du cycle de vie de la donnée.

Par ailleurs, la gestion des données de la recherche est inscrite dans la politique nationale avec le Plan National pour la Science Ouverte en respectant le principe « **aussi ouvert que possible, aussi fermé que nécessaire** ».

Qu'est-ce qu'une donnée ?

«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 des résultats de recherche.»

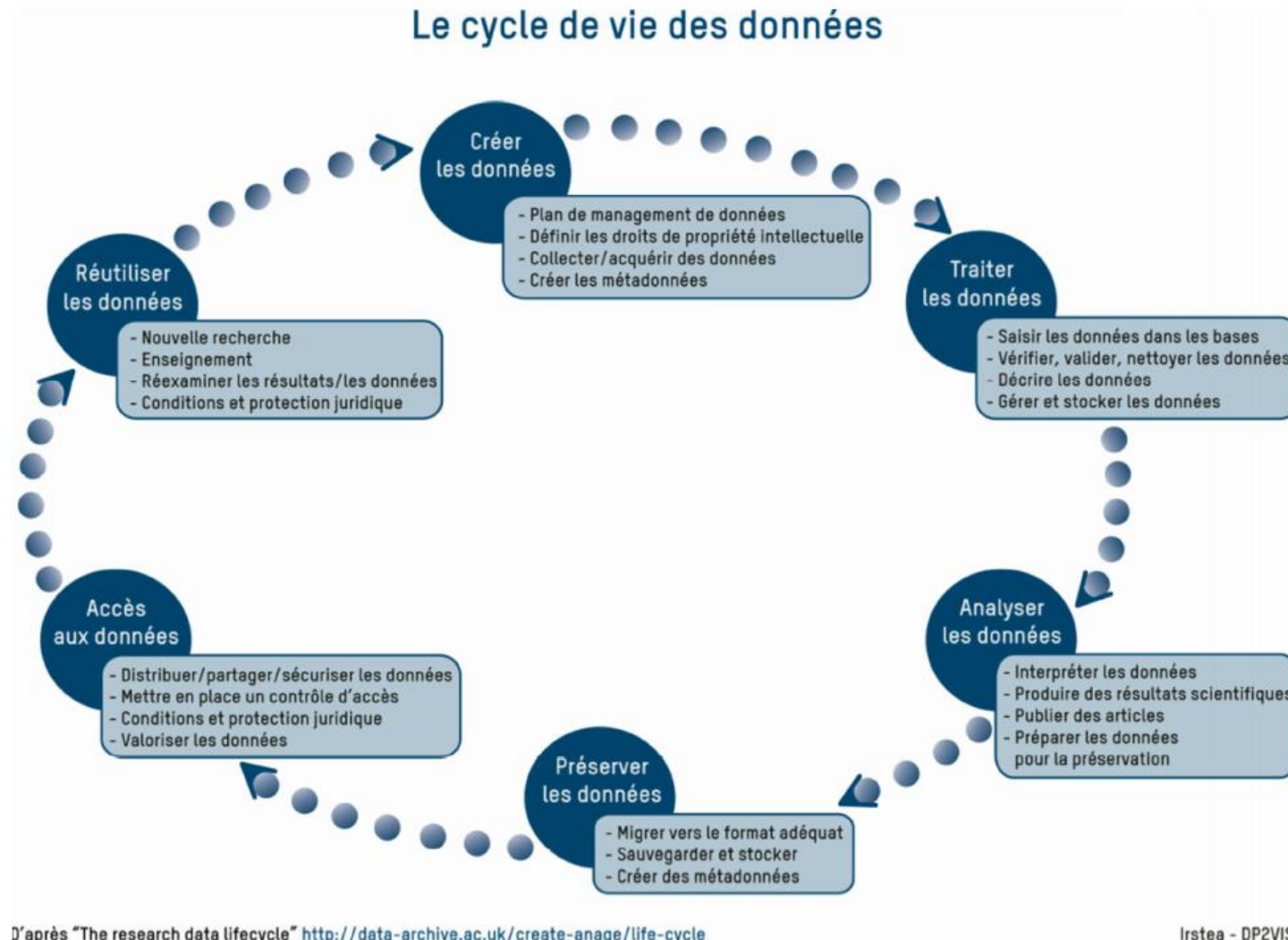
OCDE, Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics, OCDE

- Par exemple dans votre corpus de travail :
 - images
 - vidéos
 - enregistrements sonores

Données

- Dans la définition des données :
 - distinguer les données créées dans le cadre du projet des données collectées
 - si utilisation de données collectées, préciser comment elles ont été obtenues : utilisation de données produites dans un de vos projets antérieurs, utilisation des données d'un projet d'un de vos partenaires, récupération de données libres de droit...
- Pour chaque jeu de données
 - préciser les métadonnées choisies et les standards utilisés
 - préciser les formats

Cycle de vie de la donnée



Pourquoi ?

- **Demande** de plus en plus systématique **des financeurs** (Europe, ANR...) et des institutions
- **Outils de réflexion :**
 - organisation des données et de leur gestion au cours d'un projet,
 - permet d'éviter les pertes de données,
 - de décider qui fait quoi,
 - d'être attentif à certains aspects que l'on aurait oublié
 - ...

Et concrètement ?

Modèle de l'ERC fourni par la commission européenne

FAIR data and resources

1. Making data findable

Recommendations:

This should consider the dataset description

2. Making data openly accessible

Recommendations:

- Which data will be made openly available?
- Where the data and associated metadata are stored
- How the data can be accessed? And by whom?

3. Making data interoperable

Recommendations:

Which standard or field-specific data are used?

4. Increase data reuse

Recommendations:

- What data will remain re-usable after the project?
- Is embargo foreseen?
- How the data is licensed?
- Any data quality assurance procedures?

5. Allocation of resources and data security

Recommendations:

- Estimated costs for making the project FAIR
- Procedures for data backup and recovery
- Transfer of sensitive data and security

Modèle de l'ANR : exhaustif (lourd ?) et généraliste

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

Quelles métadonnées et quelle documentation accompagneront les données ?

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

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

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 ?

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 ?

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

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

Etc, etc.

Consortium MASA : guide pas à pas, "pensé SHS"

▼ MASA Consortium: MASA - Modèle de PGD (français) - Contexte

Identification du projet

Descriptif du projet et mots clés

Identification des opérations archéologiques associées (optionnel)

Historique du document

▼ MASA Consortium: MASA - Modèle de PGD (français) - Gestion des données

Responsabilités

Ressources

Contrôle qualité

▼ MASA Consortium: MASA - Modèle de PGD (français) - Généralités sur les données

Constitution du corpus

Organisation

Volume

Cadre juridique

▼ MASA Consortium: MASA - Modèle de PGD (français) - Jeux de données

Jeux de données

Description des données

Métadonnées associées

Stockage des données

Sécurité des données

Accès aux données

Protection des données sensibles au cours du projet

▼ MASA Consortium: MASA - Modèle de PGD (français) - À l'issue du projet

Politique d'archivage des données produites

Dissémination des données

Pas de panique !

- un PGD est

Évolutif

Pour exemple, l'ANR indique :

Un PGD devra être transmis à l'ANR à **T0+6** avec mise à jour à **mi-parcours** (si projet >30 mois) et à **la fin du projet**

Attention : les règles évoluent vite, se renseigner et lire les dernières infos !

Pointeurs

Découvrir et comprendre



vidéos et [tutoriels](#)



présentation [2019](#)

Modèles de PGD

ANR, ERC :

https://dmp.opidor.fr/public_templates
(ou sur les sites officiels)



<https://masa.hypotheses.org/3492>

Et dans votre institution ? À vérifier !

Outils



créer, collaborer, partager son PDG : <https://dmp.opidor.fr>
+ tutoriel en 4 vidéos : <https://oaamu.hypotheses.org/2208>)

Contacts



SOS-PGD, répertoire des Services Opérationnels de Soutien à la rédaction des Plans de Gestion des Données :

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