

Pourquoi et comment rédiger un Plan de Gestion de Données? (PGD)

Cycladoc 2025

Nayeli DESBROSSES

Service appui à la recherche et science ouverte (SCD)



Objectifs pédagogiques



- Donner une définition de ce qu'est un PGD
- Interpréter un PGD
- S'orienter dans les obligations légales du traitement de données
- Comprendre comment la rédaction d'un PGD participe de la Science Ouverte
- Identifier les outils principaux et infrastructures disponibles pour gérer les données de recherche de son projet de thèse
- Utiliser l'outil DMP OPIDoR pour rédiger un plan de gestion de données
- Localiser des modèles de PGD

Cette formation est inspirée de :

Justine Ancelin, « Pourquoi et comment rédiger un plan de gestion des données (PGD ou DMP) », 28 novembre 2019, URFIST de Paris.

Justine Ancelin et Amélie Barrio, « Accompagner la rédaction d'un plan de gestion de données (PGD) ou Data Management Plan (DMP) [Formation de formateurs] », 20 et 21 janvier 2020, URFIST de Paris.

<https://urfistinfo.hypotheses.org/3520>

Et de la formation Cycladoc 2022-2023 « Pourquoi et comment rédiger un PGD ? » d'Amélie Collin (chargée des données de la recherche – SCD Paris 1 Panthéon-Sorbonne)

Programme

Rappel du contexte : le matériau de recherche

I. Le Plan de gestion des données : QQCOQP ?
(qui, quoi, comment, où, quand, pourquoi ?)

II. Le PGD avec l'outil DMP OPIDoR

III. Comment le PGD peut m'aider dans la gestion des données ?
Outils, ressources et infrastructures

Tour de table :

Vos questions/doutes, vos attentes. Pourquoi cette formation ?

1. Allez sur **wooclap.com**
2. Entrez le code d'événement dans le bandeau **ENEJAG**



Rappel du contexte : les données de la recherche

Annexes de la thèse ➡ Données de la recherche

Une **bonne gestion des données** de la thèse permet une meilleure réutilisation (ex. publications ultérieures)

La donnée de la recherche, une notion :

- Difficile à cerner dans le continuum de la recherche.
- Protéiforme.
- Essentiellement disciplinaire et multiforme mais qui fait l'objet de normes et de standards.



Les données personnelles peuvent être des données de recherche.

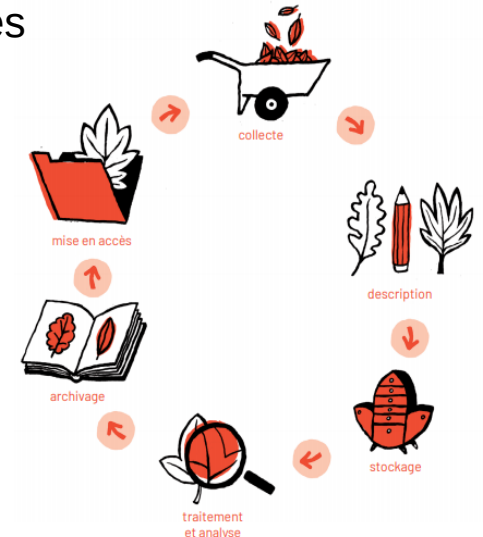
Les données sensibles forment une catégorie particulière des données personnelles.

Définition de l'OCDE

«Les données de la recherche sont définies comme des **enregistrements factuels** (chiffres, textes, images, sons, etc.), 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.»

Le respect du cycle des données

- ✓ Organisation
- ✓ Collecte
- ✓ Traitement
- ✓ Préservation
- ✓ Publication
- ✓ Réutilisation

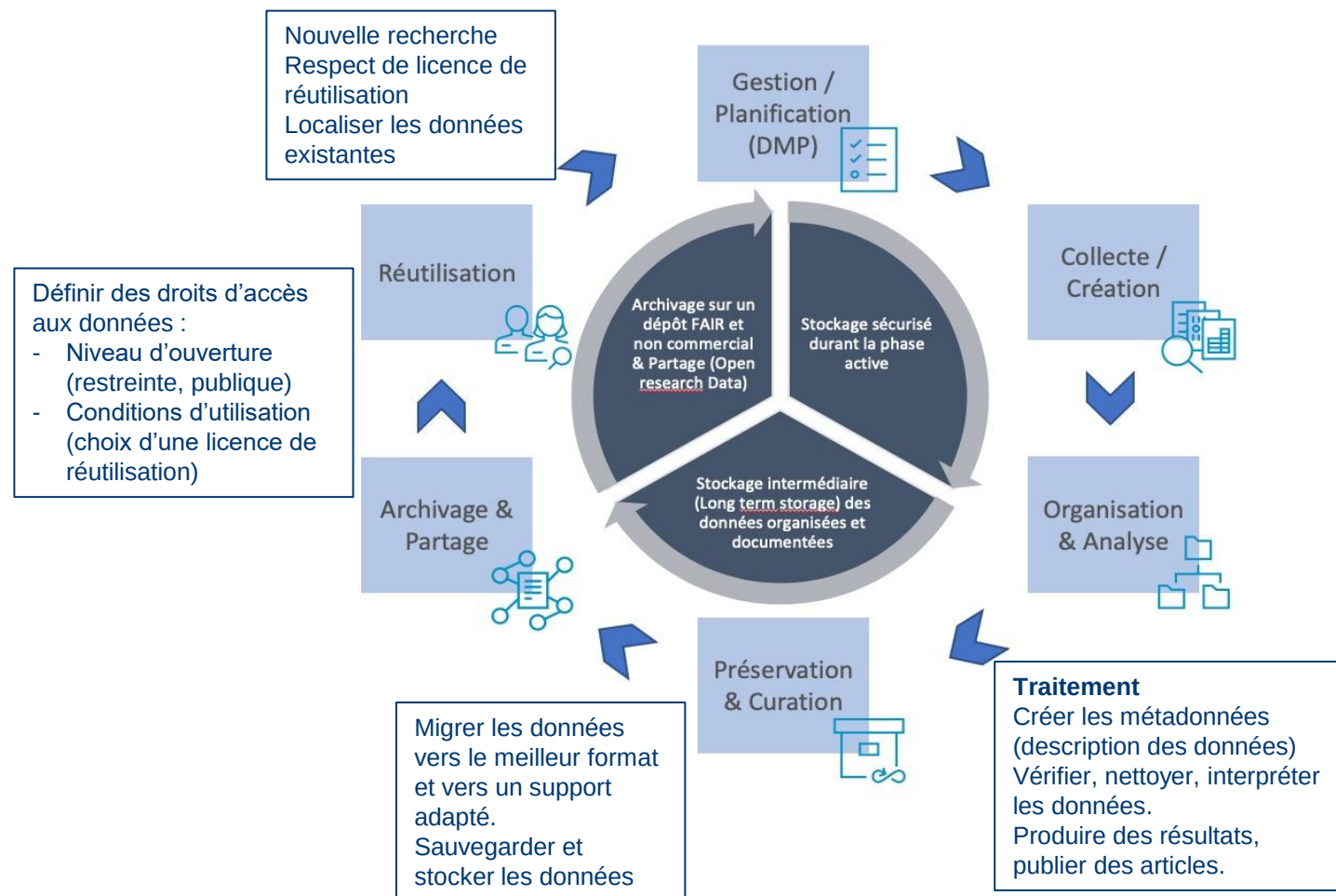


Rappel du contexte : les données de la recherche

Cycle de vie des données de recherche

DEUX questions à se poser pour identifier ses données de recherche :

- Quels sont les éléments auxquels je tiens vraiment et qui seraient irremplaçables ou très longs à remplacer en cas de perte, vol ou problème technique?
- Pour relire et évaluer mon travail, de quoi auraient besoin mes collègues pour vérifier mes résultats?



UNIL - UNIRIS 2019



Rappel du contexte : les données de la recherche



Focus sur : le droit d'auteur

La propriété intellectuelle est régie en France par le Code de la propriété intellectuelle (CPI, 1992) qui comprend deux branches :

- Le droit de la propriété littéraire et artistique (DPLA) : **droit d'auteur** (droits moraux et droits patrimoniaux) et **droits voisins**
- Le droit de la propriété intellectuelle (DPI) : regroupe les inventions, marques, dessins avec le **droit des brevets** et le **droit des marques**

Les bases de données possèdent une protection juridique propre en Europe depuis la directive européenne du 11 mars 1996.

Rappel du contexte : les données de la recherche



Focus : le droit des bases de données

- Rappel définition base de données :

Recueil de données, d'œuvres ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, individuellement accessibles par des moyens électroniques ou par tout autre moyen.

- Critère d'originalité :

Une base de données peut être considérée comme une œuvre de l'esprit uniquement si elle a une structure originale. Les données contenues dans une base de données ne sont pas des œuvres de l'esprit.

Si des choix originaux et justifiables ont été fait sur l'architecture, la structure d'une base de données pour classer, indexer les données, la créatrice/le créateur de la base détient sur celle-ci des droits d'auteur.

Si aucun choix original n'a été fait dans la construction de la base de données, le droit d'auteur ne peut s'appliquer. C'est le droit *sui generis* des bases de données, autre branche de la propriété intellectuelle qui entre alors en jeu. Dans ce cas, c'est le producteur/productrice de la base qui est titulaire du droit, c'est-à-dire la personne qui a fourni les investissements financiers, matériels, humains pour l'obtention et la vérification des données ou la constitution de la base.

I. Le plan de gestion de données

qui, quoi, comment, où, quand, pourquoi ?

I. Le Plan de gestion des données

QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

Le Plan de Gestion des Données (PGD)

- Aussi appelé DMP pour *Data Management Plan*
- Document plus ou moins formalisé qui décrit la façon dont les données liées à un projet de recherche vont être utilisées et manipulées **en amont, pendant et à l'issue du projet**.
- Pièce à usage externe, dont la rédaction est exigée par de plus en plus de financeurs de la recherche.
- Précise quelles données de recherche pourront être mises à la disposition des partenaires du projet ou du plus grand nombre, dans un souci de transparence scientifique et de partage des savoirs.
- Bien décrire la vie de ses données pour une recherche transparente, reproductible et intègre.



L'ANR (depuis 2018, AAPG 2019) et la Commission européenne (depuis 2017) exigent un plan de gestion détaillé des données de recherche mobilisées dans leurs projets.

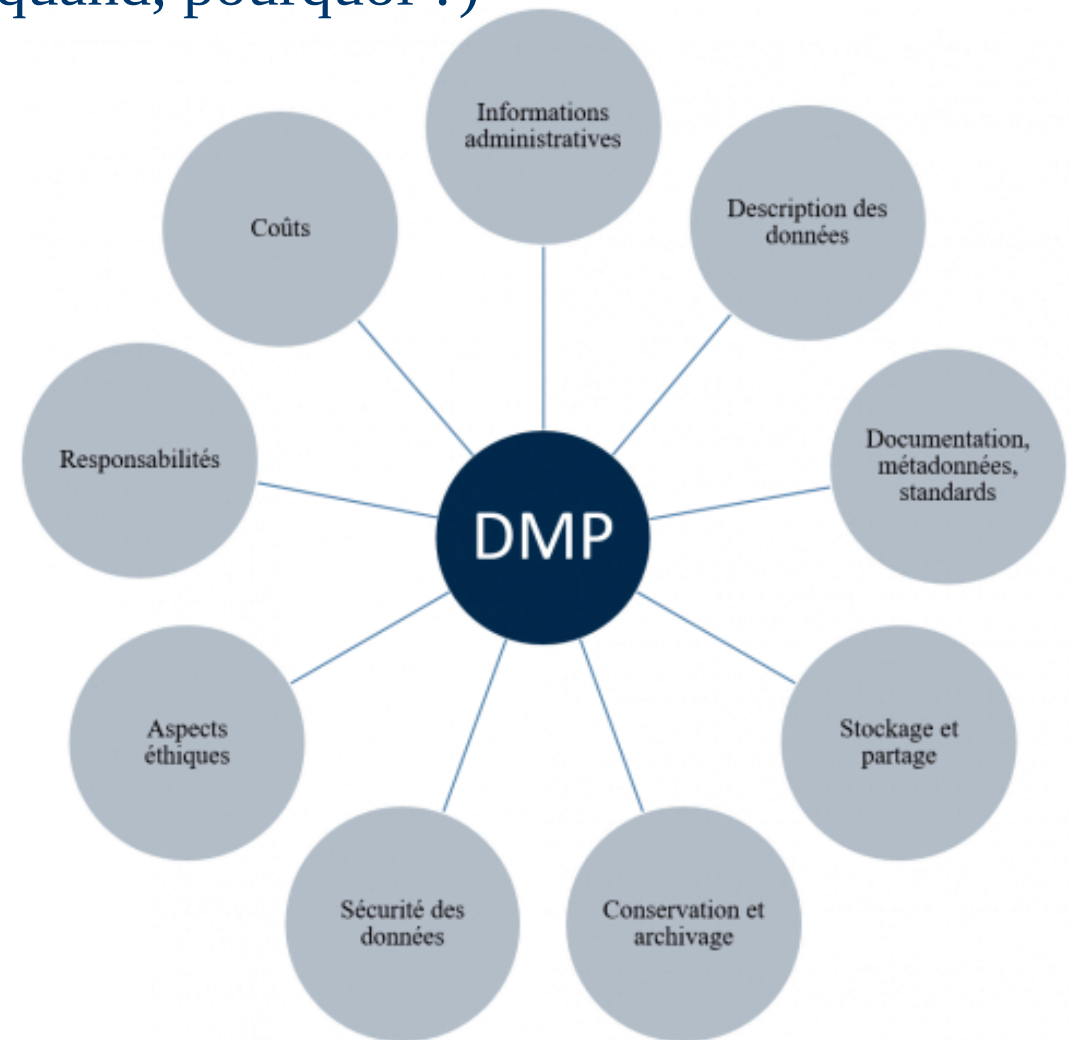
I. Le Plan de gestion des données

QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

Document outil qui décrit la façon dont les données sont :

- Obtenues
- Traitées
- Organisées
- Stockées
- Sécurisées
- Préservées
- Partagées
- etc.

... tout au long du projet et à l'issue de celui-ci.



I. Le Plan de gestion des données QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

A quelles questions répond un PGD ?



- 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?

Dans le cadre d'un projet de recherche ou non...

- 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

Source : Sylvie Cocaud, Dominique l'Hostis. Pourquoi et comment rédiger un plan de gestion de données ? Pourquoi et comment rédiger un plan de gestion de données ? (Stage de formation URFIST), 2019, pp.128 slides. <https://hal.inrae.fr/hal-02791507>

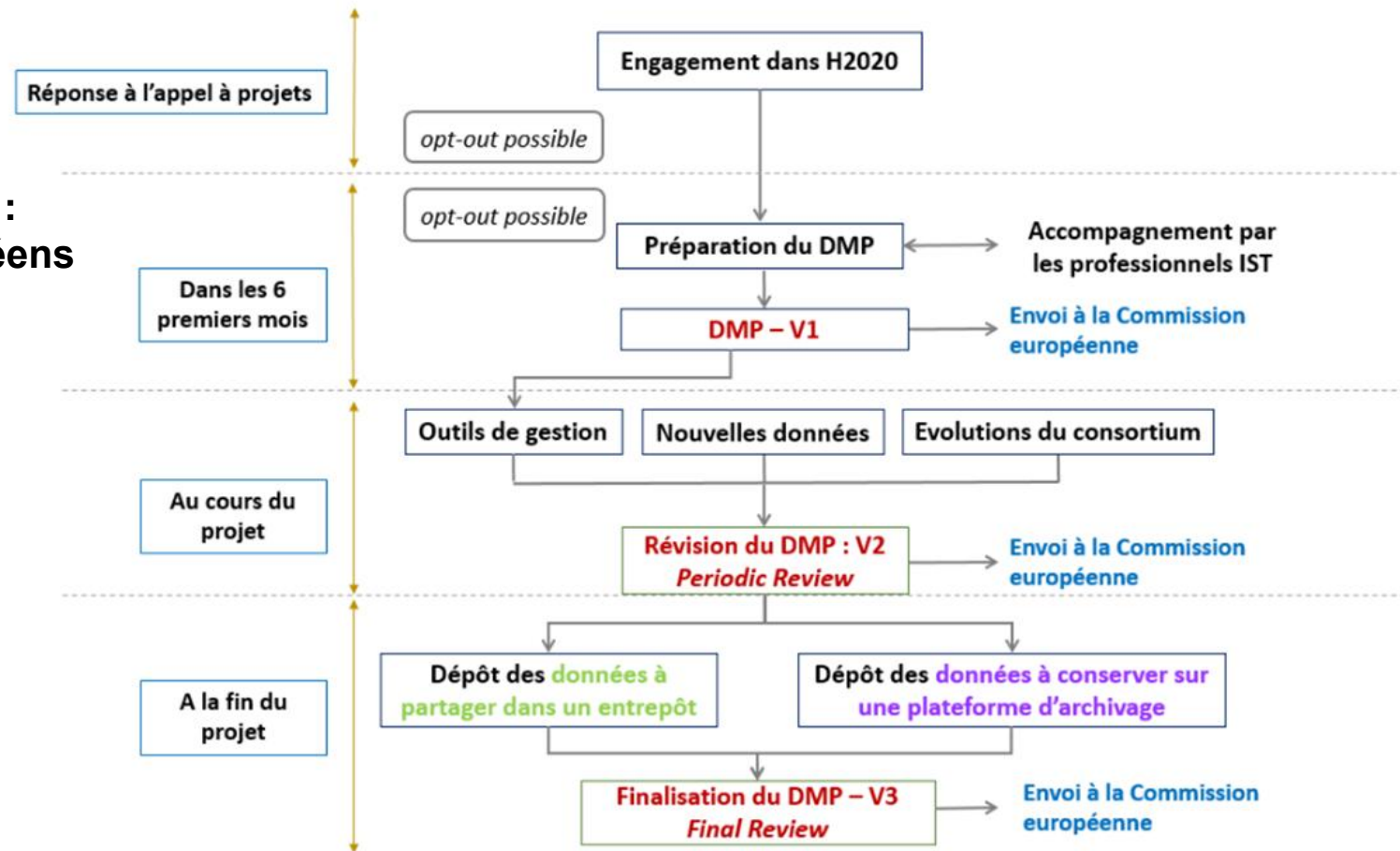
I. Le Plan de gestion des données QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

Les politiques nationales et européennes de science ouverte



I. Le Plan de gestion des données QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

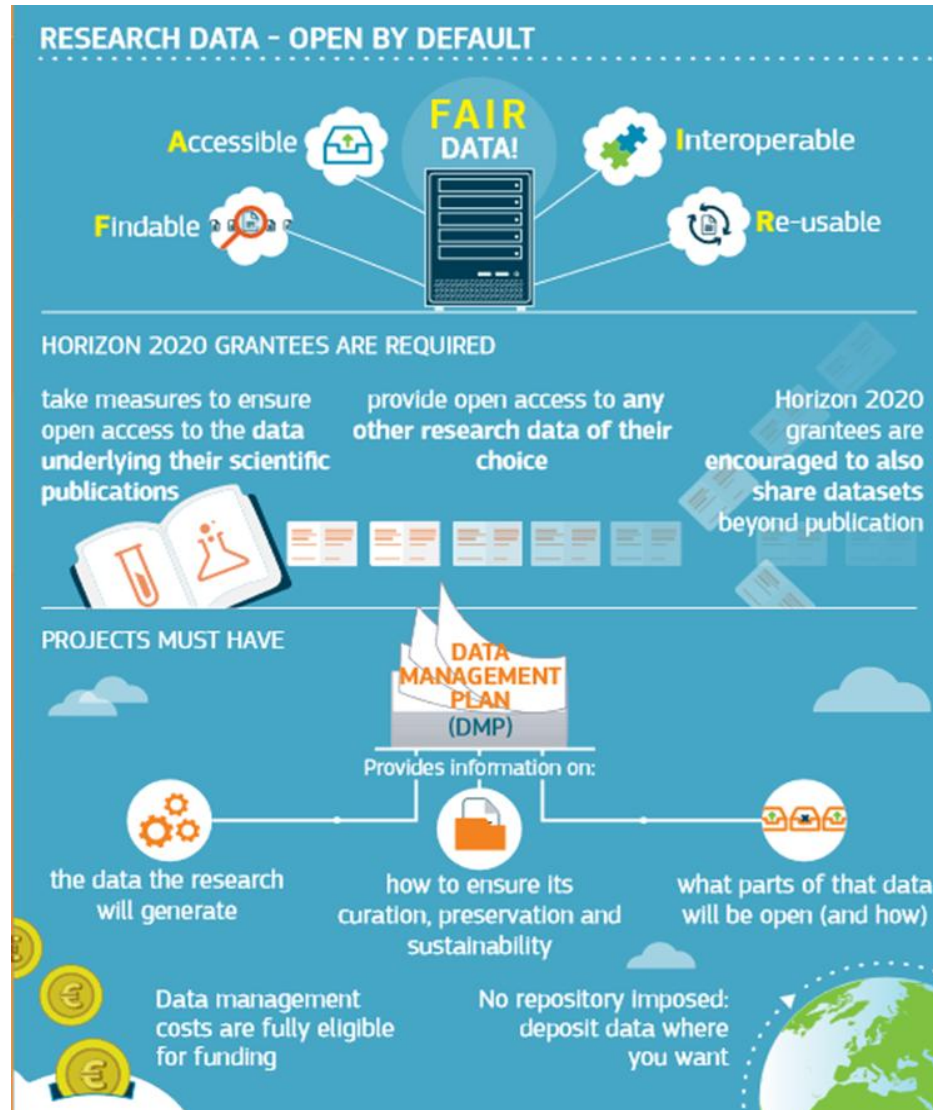
Le PGD est un document évolutif :
l'exemple dans les projets européens



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

I. Le Plan de gestion des données QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

■ Le PGD : intérêt des parties prenantes



- La Commission Européenne : politique d'accès ouvert aux publications et aux données
 - ✓ Accélérer l'innovation
 - ✓ Permettre de nouvelles formes de recherche sur les données
 - ✓ Inciter aux collaborations, éviter les doublons, améliorer l'efficacité
 - ✓ Améliorer la qualité et la transparence
 - ✓ Améliorer le retour sur investissement de la R&D
- Projets H2020, puis Horizon Europe
 - ✓ Libre accès à toutes les publications dans des revues avec peer review
 - ✓ Libre accès aux données de recherche (ORD Pilot étendu)
 - ✓ Données et métadonnées nécessaires à la validation des publications
 - ✓ Autres (méta)données "as open as possible, as closed as necessary"
 - ✓ PGD obligatoire
 - ✓ Désengagement possible (*opt out*)

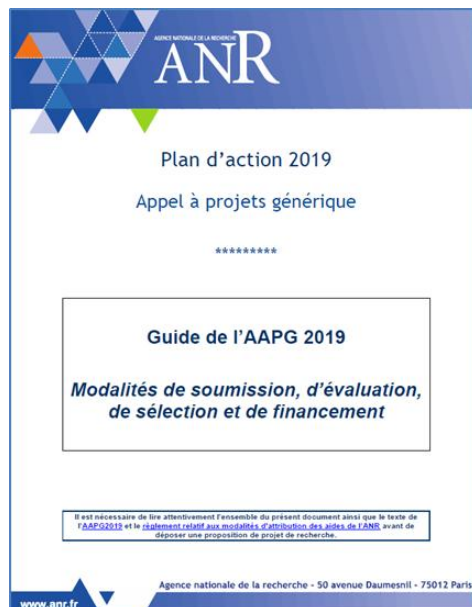
I. Le Plan de gestion des données QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

L'Agence Nationale pour la Recherche (ANR)
Impose un PGD pour les projets financés à partir de 2019

« Afin de favoriser la diffusion ouverte des données de recherche, l'ANR attire l'attention des déposants sur l'importance de considérer la question des données de recherche au moment du montage et tout au long du projet. »

<http://www.agence-nationale-recherche.fr/PA2019>

https://anr.fr/fileadmin/documents/2019/ANR_Modele_de_DMP_francais_DMPOPIDoR_2019_07_24_mis_en_page_2.pdf



Université Paris 1 Panthéon-Sorbonne

Plan national pour la science ouverte (2018), puis
PNSO 2 (2021-2024)

Les résultats de la recherche scientifique ouverts à tous, sans
entrave, sans délai, sans paiement.

3 axes :

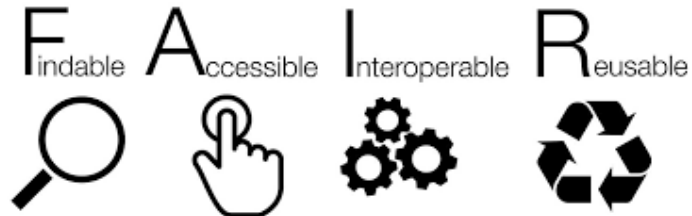
- généraliser l'accès ouvert aux publications
- structurer et ouvrir les données de la recherche
- s'inscrire dans une dynamique durable, européenne et internationale



I. Le Plan de gestion des données

QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

Le PGD et les principes FAIR



Facile à trouver

Accessibles

Interopérables

Réutilisables



I. Le Plan de gestion des données

QQCOQP ? (qui, quoi, comment, où, quand, pourquoi ?)

Le PGD et les principes FAIR

Findable

- F1 Les données et les métadonnées sont identifiées par un identifiant global unique et pérenne.
- F2 Les métadonnées décrivant les données sont riches.
- F3 Les métadonnées spécifient l'identifiant de la donnée.
- F4 Les données et les métadonnées sont enregistrées et indexées dans un dispositif permettant de les rechercher.

Accessible

- A1 Les données et les métadonnées sont accessibles par leur identifiant via un protocole de communication standardisé.
 - A1.1 Le protocole utilisé est ouvert, libre et peut être implémenté de manière universelle.
 - A1.2 Le protocole utilisé permet l'authentification et l'autorisation si besoin.
- A2 Les métadonnées sont accessibles même quand les données ne le sont plus.

Interoperable

- I1 Les données et les métadonnées utilisent un langage formel, accessible, partagé et largement applicable pour la représentation des connaissances.
- I2 Les données et les métadonnées utilisent des vocabulaires qui respectent les principes FAIR.
- I3 Les données et les métadonnées incluent des liens vers d'autres (méta)données.

Reusable

- R1 Les données et les métadonnées ont des attributs multiples et pertinents.
 - R1.1 Les données et les métadonnées sont mises à disposition selon une licence explicite et accessible.
 - R1.2 Les données et les métadonnées sont associées à leur provenance.
 - R1.3 Les données et les métadonnées correspondent aux standards des communautés indiquées.

II. Le PGD avec l'outil DMP OPIDoR

Relecture commentée de PGD

II. Le PGD avec l'outil DMP OPIDoR

Exemples de PGD et grille de lecture

PGD publics sur DMP OPIDoR

https://dmp.opidor.fr/public_plans#content

Grille de relecture de plans de gestion de données

https://doranum.fr/plan-gestion-donnees-dmp/grille-de-relecture-de-plans-de-gestion-de-donnees_10_13143_r7gm-6c38/



II. Le PGD avec l'outil DMP OPIDoR



DMP OPIDoR est l'acronyme de Data Management Plan pour l'Optimisation du Partage et de l'Interopérabilité des Données de Recherche.

Outil simple et intuitif pour l'élaboration des plans de gestion de données en ligne.

- ✓ Création de PGD à partir de modèles préexistants
- ✓ Modèles de structures ou financeurs pour faciliter la réutilisation automatique d'information
- ✓ Possibilité de partage de PDG pour une rédaction collective
- ✓ Mise à jour possible à tout moment
- ✓ Conforme aux principes FAIR : facilite saisie et réutilisation d'information (référentiels/thésaurus, identifiants de la recherche comme ORCID)
- ✓ Téléchargement sous différents formats

II. Le PGD avec l'outil DMP OPIDoR



Les modèles de PGD

Exemples de modèles d'établissements

- AgroParisTech - Institut des sciences et industries du vivant et de l'environnement
- Centre de coopération internationale en recherche agronomique pour le développement (CIRAD)
- "Memories of Archaeologists and Archaeological Sites" consortium (MASA)
- Université Paris 8

Exemples de modèles des financeurs

- Agence nationale de la recherche (ANR)
- Commission européenne (Horizon 2020 FAIR DMP)
- Conseil européen de la recherche (European Research Council, ERC)

https://dmp.opidor.fr/public_templates

II. Le PGD avec l'outil DMP OPIDoR



Démonstration et pratique de l'outil DMP OPIDoR

<https://dmp.opidor.fr/>

L'outil de rédaction DMP OPIDoR

https://doranum.fr/plan-gestion-donnees-dmp/le-pgd-et-loutil-de-redaction-dmp-opidor_10_13143_es2g-0f16/

III. Comment le PGD peut m'aider dans la gestion des données ?

Outils, ressources et infrastructures

III. Comment le PGD peut m'aider dans la gestion des données ?

Outils, ressources et infrastructures



- Le [datapaper](#) et les *data journals*
- Accompagnement par le dispositif [CoDataSorb](#) (dans le cadre du réseau des [Ateliers de la donnée](#) – services généralistes sur tout le territoire)

codatasorb@sorbonne-alliance.fr

III. Comment le PGD peut m'aider dans la gestion des données ?

Outils, ressources et infrastructures disponibles pour la gestion des données de recherche



III. Comment le PGD peut m'aider dans la gestion des données ?

Ressources en autoformation



III. Comment le PGD peut m'aider dans la gestion des données ?

Quelques outils et services complémentaires



III. Comment le PGD peut m'aider dans la gestion des données ?

Partager et diffuser ses données



Catégories d'entrepôts de données

- Institutionnels (en France)
- Pluridisciplinaires et internationaux
- Thématiques ou disciplinaires
- Liés à un éditeur



Les licences libres (Creative Commons, Etalab)
Les identifiants pérennes de la recherche (DOI, ORCID)



Liste des entrepôts de données thématiques de confiance

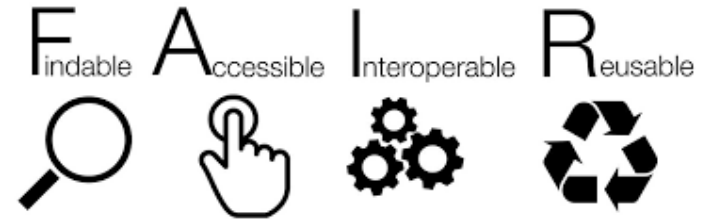


Vérifier le cadre légal de diffusion de ses données avec l'outil d'aide à la décision de Andro M., Morcrette, N., Gandon, N.

<http://www.bibliotheque-numerique.fr/DonneesDiffusables.php>

III. Comment le PGD peut m'aider dans la gestion des données ?

Mes données sont-elles FAIR ?



FAIR-Aware : un outil d'auto-évaluation de ses pratiques de gestion de données

<https://fairaware.dans.knaw.nl/>

Pour aller plus loin – Liens utiles sur Internet

Pour aller plus loin – Liens utiles sur Internet

1/2

Documentation généraliste sur la Science Ouverte

[Les bonnes pratiques de la Science Ouverte appliquées aux thèses de doctorat](#), support de formation doctorale (Université de Lille), 2020.

[Passeport pour la Science Ouverte](#), guide pratique à l'usage des doctorants et doctorantes.

[Déclinaison du passeport](#) relative à la gestion et à la diffusion des données de la recherche.

Documentation généraliste pour la gestion des données de recherche

[Logigrammes interactifs : questions juridiques et éthiques liées à mon projet de recherche](#)

Ressources complémentaires :

[Droit des bases de données et logiciels, DAJ, Mission appui au patrimoine immatériel de l'Etat, ministère de l'Économie, des finances et de la relance](#)

Répertoires de ressources et services

[Cat OPIDoR](#) Catalogue des services français dédiés aux données de la recherche - Inist-CNRS.

[DoRANum \(Données de la Recherche : Apprentissage NUMérique à la gestion et au partage\)](#) Inist-CNRS, Réseau des Urfist, BSN9, Dist-CNRS.

Le projet DoRANum a pour objectif de mettre en place un dispositif de formation à distance d'accès coordonné, intégrant différentes ressources d'auto-formation sur la thématique de la gestion et du partage des données de la recherche.

[SOS-PGD : répertoire des Services Opérationnels de Soutien à la rédaction des Plans de Gestion des Données](#) Groupe Données du GTSO Couperin

Ressources pour autoformation

[Gérer les données de la recherche](#) CoopIST Coopérer en information scientifique et technique - CIRAD

[IRD Data](#) Site support pour des données ouvertes et une science durable au Sud – IRD

[Données de la recherche : DMP Sciences Po](#) Libguides Données de la recherche - SciencesPo Bibliothèque

[Guide de bonnes pratiques sur la gestion des données de la Recherche](#), groupe de travail inter-réseaux « Atelier Données », CNRS, 2021

[Ressources et autoformations](#) sur la gestion et le partage des données de la recherche (DoRANum)

[Réaliser un plan de gestion de données 'FAIR' : modèle](#) [PDF, 33 pages] Reymonet N., Moysan M., Cartier A., Délémontez R. Université Paris Diderot et Université Paris Descartes. 2018. Ce document a pour vocation d'accompagner les chercheurs et chargés de projets dans la rédaction d'un PGD. 2018.

Pour aller plus loin – Liens utiles sur Internet

2/2

Ressources spécifiques

- **Logiciels de documentation**
Liste d'outils pour saisir les métadonnées et la documentation
[RDA Metadata Standards Catalog](#)
- **Standards de métadonnées/données**
Liste de standards de métadonnées ou données utilisés pour décrire les produits de recherche en application des principes FAIR
[RDA Metadata Standards Catalog](#)
- **Plateformes d'archivage de données**
Liste de plateformes accueillant des données de recherche au format numérique, dans le but de les conserver sur le long terme, tout en préservant la lisibilité et l'intelligibilité des fichiers dans le temps
- [Cat OPIDoR/Plateforme d'archivage](#)

Rédaction d'un DMP - Retours d'expériences

[Rédaction d'un plan de gestion de données dans le cadre du projet Hospitam](#)

Lauriane Locatelli, Post-doctorante ENS Lyon / HiSoMA (Le blog de la TGIR Huma-Num - Hypothèses - Mis à jour le 08/09/2020)

[La gestion de données d'archives orales en sciences du patrimoine](#)

Gwendoline Torterat, Post-doctorante, DIM Matériaux Anciens et Patrimoniaux Île-de-France (Le blog de la TGIR Huma-Num - Hypothèses - Mis à jour le 08/09/2020)

[Rédiger un Plan de gestion de données pour un projet ERC : l'expérience GRAPH-EAST](#)

Manon Durier, ingénieure de recherche en charge de l'infrastructure des données (ERC GRAPH-EAST - Hypothèses - Publié le 09/12/2021)

Data papers

URFIST [Rédiger et publier un data paper](#) (cours en ligne)

DATAACC [Publier un data paper : où et comment?](#)

Data journal (revue de données) en SHS : [Revue Data & Corpus](#)

Merci!

Nayeli Desbrosses

appui-recherche-scd@univ-paris1.fr
codatasorb@sorbonne-alliance.fr

Service Commun de la Documentation (SCD)