

Ce webinaire a pour objectif de vous expliquer ce qu'est un Plan de Gestion de Données, à quoi il sert, ce qu'il contient et comment le rédiger.

Sommaire



01 Enjeux et objectifs d'un PGD

02 Contenu et rédaction

03 Présentation rapide de DMP OPIDoR

cnrs Inist

01

Enjeux et objectifs d’un PGD

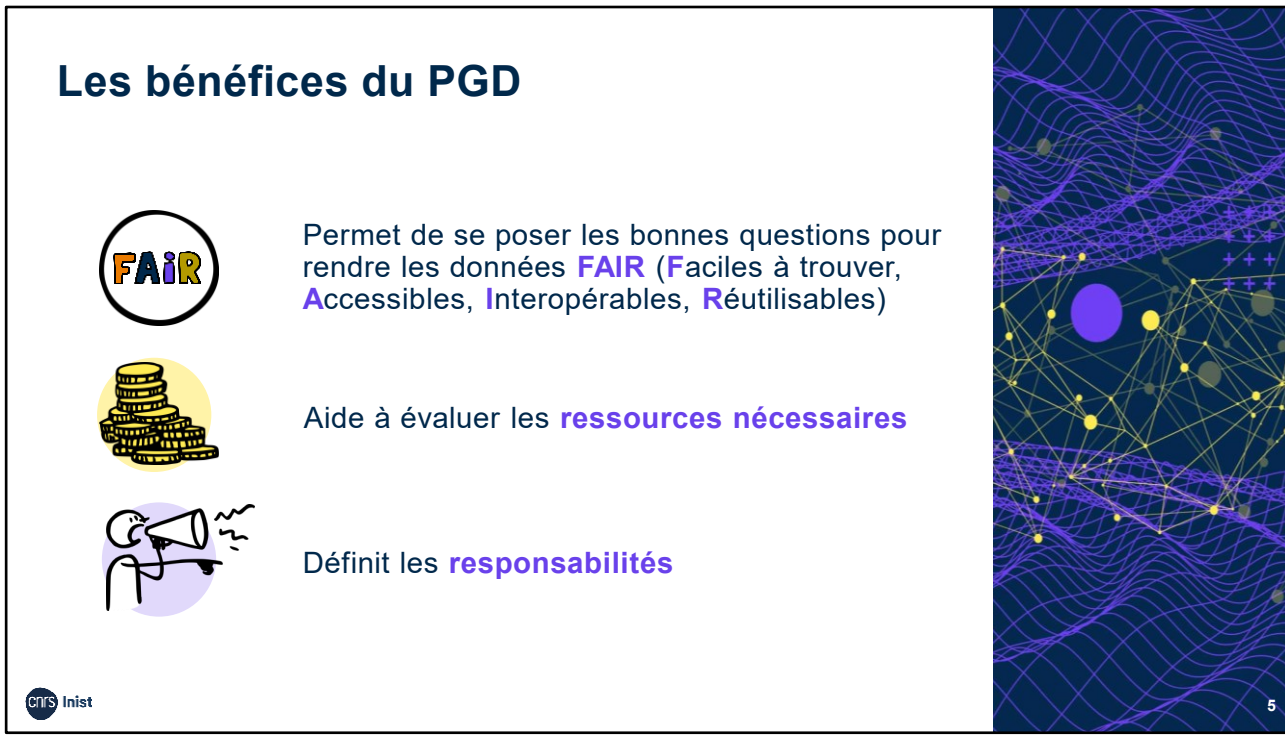


Le PGD (ou DMP) est un document de quelques pages, qui permet de définir et de mettre en place les actions indispensables pour gérer les données qui vont être collectées ou produites au cours d'un projet.
Il explique comment sont gérées les données depuis leur création ou collecte jusqu'à leur partage et leur archivage.
Le PGD est évolutif : sa rédaction démarre dès le début d'un projet, mais il va être mis à jour régulièrement en fonction de l'évolution du projet. Vous n'avez pas à répondre à toutes les questions du PGD dans sa phase initiale mais il est bon de réfléchir à tous les points abordés dans le PGD. Ce document est là pour faciliter la gestion des données de recherche de votre projet.
Surtout, le PGD aide à organiser et anticiper toutes les étapes du cycle de vie de la donnée.

Ressource :

- *DoRANum. Cours introductif sur le Plan de Gestion de Données (PGD).*

https://doi.org/10.13143/T3J4-VN03



L'élaboration d'un plan de gestion permet de se poser les bonnes questions afin d'obtenir des données FAIR, c'est-à-dire faciles à trouver, accessibles, interopérables et réutilisables.

Le PGD va aussi vous aider à réfléchir aux ressources nécessaires afin de mettre en œuvre les actions décrites dans le PGD :

- temps nécessaire à la préparation des données pour le stockage, le partage et l'archivage des données,
- coûts de matériel, rémunération des personnels,
- frais de stockage (serveurs dédiés, traitement, maintenance, sécurité, accès…), de partage (site web, publication…) et d'archivage des données.

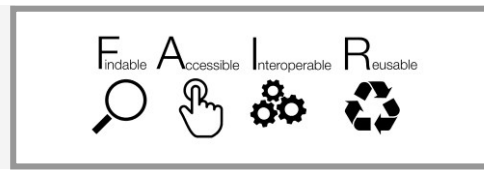
Dans le PGD, il faut désigner nominativement la ou les personne(s) responsable(s) de la gestion des données pour toutes les étapes du projet et au sein du partenariat s'il y a lieu :

- saisie des données,
- production des métadonnées,
- contrôle de la qualité des données,
- mise à jour du PGD.

Principes FAIR

4 principes à respecter pour garantir une utilisation optimale des données de recherche et des métadonnées associées, à la fois **par les hommes et par les machines**

- **F** (Findable) = Facile à trouver
- **A** (Accessible) = Accessible
- **I** (Interoperable) = Interopérable
- **R** (Reusable) = Réutilisable



Sangya Pundir [CC BY-SA 4.0]

Applicables tout au long du cycle de vie des données

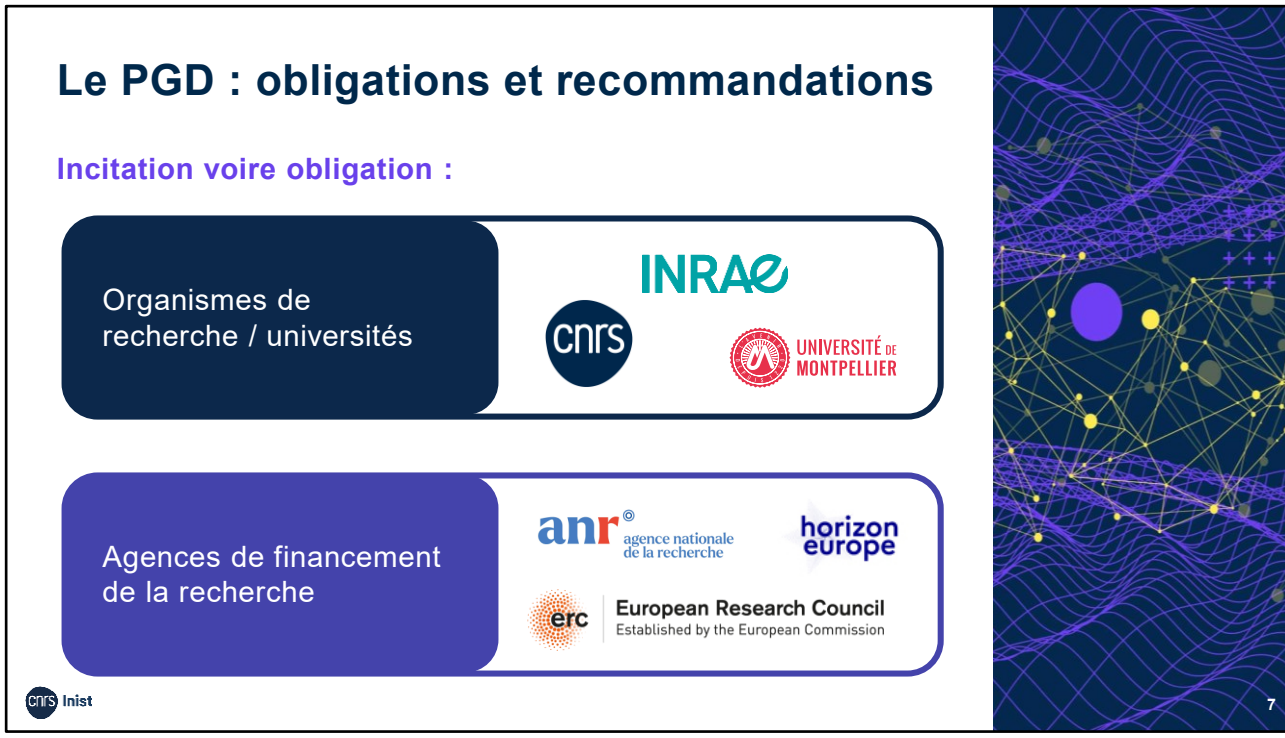
cnrs Inist

6

Les principes FAIR sont admis par les différentes communautés scientifiques au niveau international, ainsi que par les financeurs (ex : Commission européenne, ANR, etc.)

Ressource :

DoRANum. Les principes FAIR. https://doi.org/10.13143/Z7S6-ED26



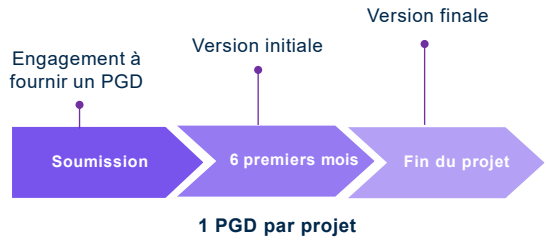
Le PGD répond aussi à une incitation voire à une obligation institutionnelle :

- Certains organismes de recherche et universités comme le CNRS, l'INRAE ou l'Université de Montpellier incitent ou encouragent leurs chercheurs à élaborer un PGD pour faciliter l'adoption de bonnes pratiques mais aussi permettre de répertorier les données produites au sein de l'institution.
- Pour d'autres, comme les agences de financement de la recherche, le PGD constitue un livrable qui devra être fourni dans les 6 mois qui suivent le démarrage scientifique du projet : ANR, Commission européenne (Horizon Europe), European Research Council…

Le PGD : obligations et recommandations

Exemple de l'anr agence nationale de la recherche

- Demande un PGD pour tous les projets qu'elle finance
- Le PGD est un livrable à fournir dans les 6 mois



cnrs Inist

8

C'est depuis 2019 que l'Agence Nationale de la Recherche demande l'élaboration d'un plan de gestion de données pour les projets qu'elle finance.

Elle demande :

- 1 PGD par projet.
- Lors de la soumission du projet : de s'engager à en fournir un.
- Qu'une première version soit livrée dans les 6 premiers mois du projet.
- Que la version finale du PGD soit livrée à la fin du projet.

Le PGD conditionne-t-il le versement de l'aide ?

« L’ANR vérifiera que les différentes versions du PGD seront bien transmises. Le dernier versement de l’aide est conditionné par la réception d’un PGD et de ses mises à jour à la date de fin des travaux scientifiques. » https://anr.fr/fr/lanr/engagements/faq-pgd/#:~:text=Le%20PGD%20conditionne%2Dt%2Dil,de%20fin%20des%20travaux%20scientifiques.

02

Contenu et rédaction

Outils de rédaction d'un PGD

Il existe des outils en ligne qui facilitent la rédaction.
En France, l'outil recommandé est :

DMP OPIDoR



cnrs Inist

10

DMP OPIDoR a été personnalisé par l'Inist-CNRS pour les équipes de recherche françaises et est l'outil de rédaction de PGD le plus utilisé en France.
Dans l'enquête menée en 2020/2021 par le GTSO Données de Couperin sur les services d'accompagnement aux données, 77% des répondants conseillent DMP OPIDoR.

Ressource

Féret, R., Bracco, L., Lehoux; Elise, & Augouvernaire, M. (2021). Enquête sur l'appui à la gestion des données de la recherche en service de documentation et d'information scientifique et technique. Zenodo. https://doi.org/10.5281/zenodo.5078504



Il n'existe pas de trame unique. Plusieurs modèles de PGD sont disponibles, mais le réseau Science Ouverte des agences de financement françaises (ADEME, ANR, ANRS-MIE, Anses, FRM, INCa) et de nombreux organismes de recherche recommandent l'utilisation du **modèle commun structuré** (disponible sur DMP OPIDoR).

Dans un PGD on retrouve les éléments suivants :

- Les informations administratives
- La description de vos différents jeux de données
- Tout ce qui concerne la documentation de vos données avec le renseignement des métadonnées, selon quel standard…
- Les procédures mises en place pour assurer la qualité des données
- Les exigences légales et éthiques
- La sécurité des données
- Les modalités de stockage et de sauvegarde sécurisée des données durant le projet
- Les modalités de partage : quels jeux seront partagés, dans quel entrepôt, selon quelles conditions…
- Éventuellement, les modalités d'archivage pérenne si certains jeux de

données le nécessitent

- La répartition des responsabilités
- L’évaluation des coûts.

À noter qu’il existe des spécificités pour la description de logiciels.

Un exemple de PGD : TRANSFUNERAIRE



DMP du projet "TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles "

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - Modèle de PGD (français)" fourni par Agence nationale de la recherche (ANR).

Renseignements sur le plan

Titre du plan	DMP du projet "TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles "
Langue	fra
Date de création	2020-06-26
Date de dernière modification	2020-07-22
Identifiant	

Renseignements sur le projet

Titre du projet	TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles
Résumé	Notre programme est réalisé par deux équipes d'anthropologues sociaux travaillant sur l'Europe (sous la direction d'Elisabeth Anstett) et sur l'Amérique-Latine (sous la direction de Valérie Robin Azevedo) ; il propose d'effectuer une analyse comparative des pratiques funéraires qui ont actuellement lieu en Europe et en Amérique-Latine, dans des contextes marqués par la violence extrême et la mort de masse. Il vise à éclairer leurs transformations les plus contemporaines mais aussi les bricolages rituels, religieux et politiques mis en œuvre pour faire avec des corps fragmentés ou une absence inéluctable de corps. Au-delà d'une interrogation sur l'adaptabilité des grammaires et des lexiques funéraires, et d'un questionnement sur l'impact des pratiques médico-légales, notre objectif est de saisir les enjeux de ces transformations et les modalités de reconfiguration de la relation aux défunts ré-inhumés. Dans ces situations mortuaires inédites, une compréhension empirique et précise des procédures matérielles et symboliques entourant le destin posthume des restes humains ou des corps absents, ambitionne ainsi de mieux cerner les mécanismes socioculturels de la résilience en temps de crise.
Sources de financement	• Agence nationale de la recherche (ANR) : ANR-19-CE27-0022

Produits de recherche :

1. Données collectées par les membres du programme Transfunéraire (Collection)
2. Données produites par les membres, dans le cadre du programme Transfunéraire (Collection)
3. Fichier croisant les informations relatives aux données collectées et produites (Texte)

DMP du projet "TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles "

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

Données collectées par les membres du programme Transfunéraire

1a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

Les données sont collectées et produites lors d'enquêtes ethnographiques réalisées sur le terrain à partir d'entretiens non directif et d'observations (données produites). Ces enquêtes permettent aussi de collecter des corpus documentaires (archives presse ou média, et archives personnelles).
La situation sanitaire mondiale induite par la crise du Covid-19 (maladie à coronavirus) a empêché que la plupart des enquêtes de terrain prévues à être réalisées en 2020 pour produire des données qualitatives inédites ne se fassent dans les délais prévus initialement.
Cette situation a amené les membres à réorienter leurs travail vers l'analyse de matériaux préexistants. Ces données sont issues de recherches antérieures financées par les institutions respectives des membres du projet, et dont ils conservent un usage libre. Ces données sont utilisées pour les analyses forgées par les chercheurs membres du programme de la même façon que les données collectées et les données produites.
Ces données peuvent être :

- soit issues de recherches antérieures des membres du programme.
- soit issues d'une collecte réalisée à distance du terrain, sur des sources numérisées disponibles en ligne (articles de presse, images photographiques ou filmées).

Ces données peuvent être : des photographies, des enregistrements sonores, des enregistrements vidéos, des documents écrits (articles de presse, documents d'archives, notes de terrain).

1b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

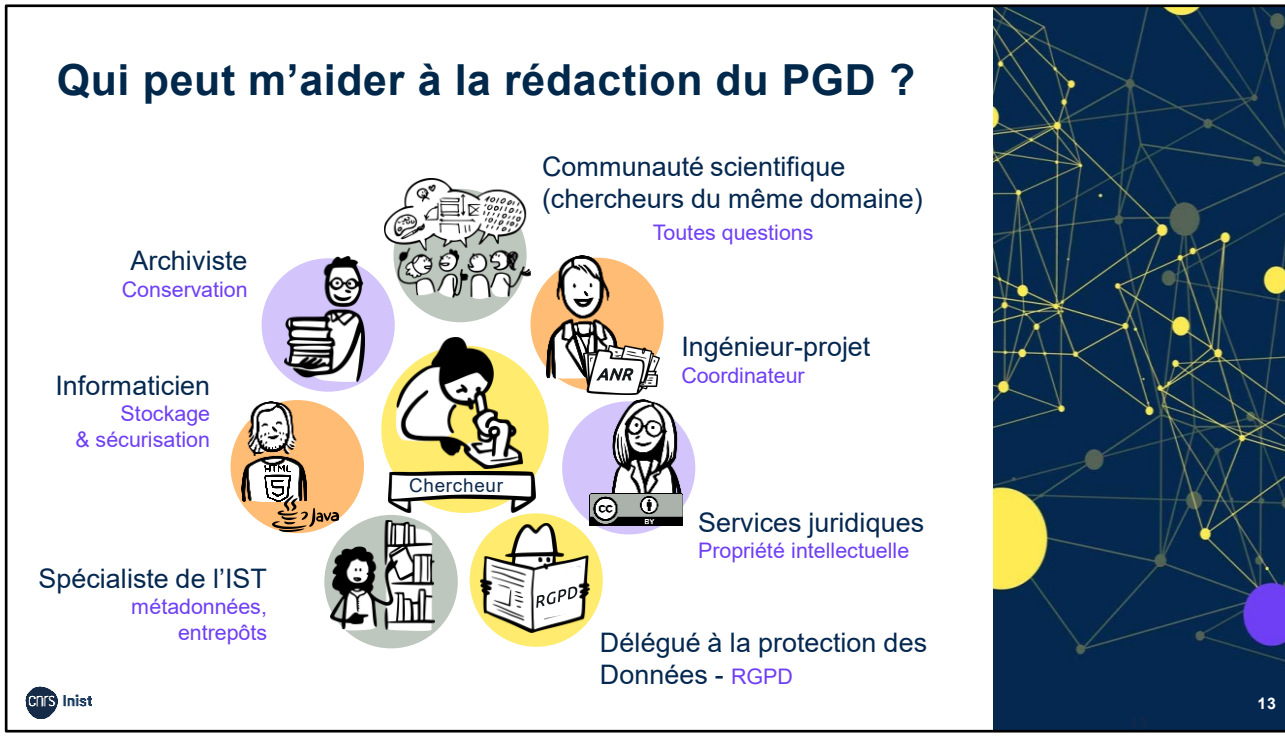
Les données collectées et produites peuvent être :
- Des images
- Des vidéos
- Des documents

cnrs Inist

12

Voilà un exemple d'un PGD rédigé avec DMP OPIDoR puis exporté en PDF. Vous pouvez voir la première partie avec les renseignements sur le plan et le projet puis la partie rédaction qui se présente sous forme de questions/réponses organisées en rubriques, qu'il faut compléter.

Pour y accéder : https://dmp.opidor.fr/plans/6619/export.pdf



L'équipe de recherche n'est pas seule face à la rédaction du PGD.

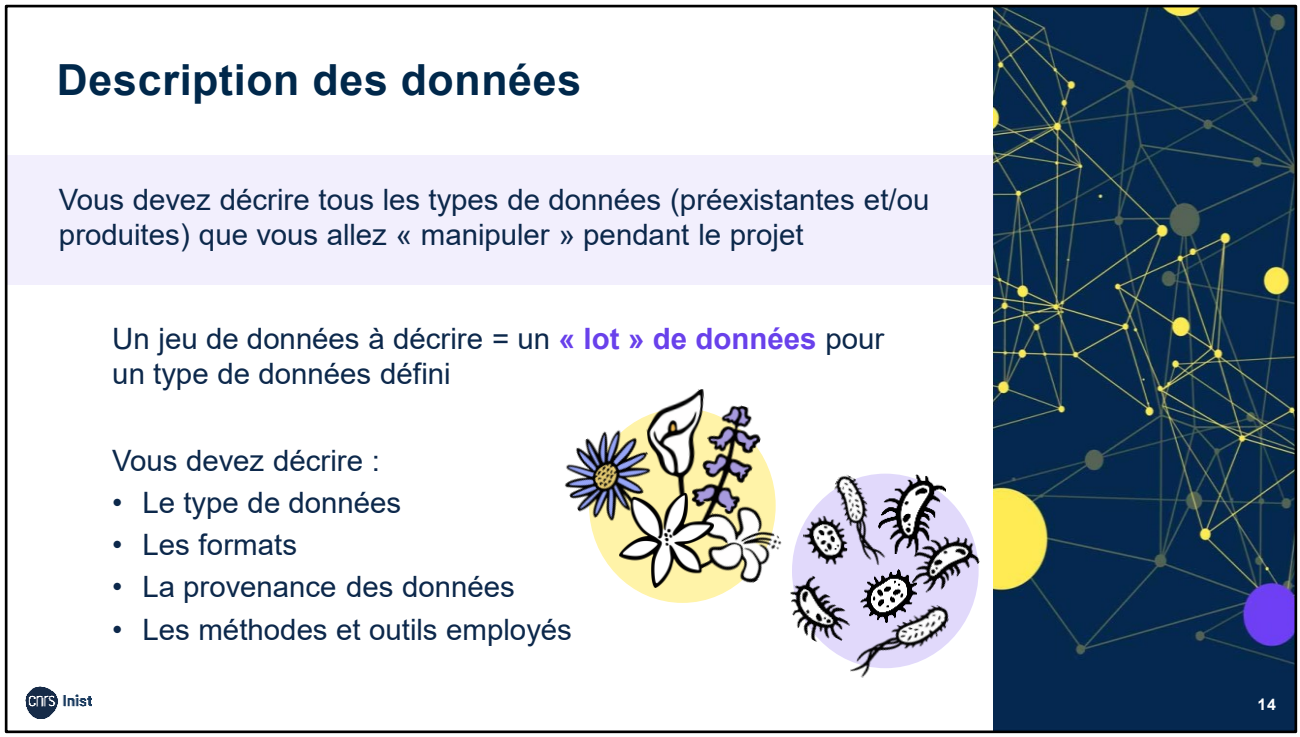
Des personnes ressources peuvent être sollicitées :

- **Les ingénieurs projet** peuvent aider à coordonner les actions autour du PGD mais aussi dans l'obtention d'un agrément, pour l'éligibilité des coûts...
- **Les services juridiques** peuvent notamment conseiller sur la propriété intellectuelle des données.
- **Les délégués à la protection des données** sont les interlocuteurs privilégiés pour les aspects RGPD (règlement général sur la protection des données).
- **Les spécialistes de l'Information Scientifique et Technique** peuvent notamment aider au choix des standards de métadonnées, conseiller sur les entrepôts mais aussi réaliser des alignements avec des référentiels existants. Les Ateliers de la donnée proposent aussi un accompagnement au niveau local.
- **Les informaticiens** sont les interlocuteurs privilégiés en ce qui concerne le stockage et la sécurisation des données, les aspects infrastructure et les coûts associés.
- **Les archivistes** peuvent aider à sélectionner les données pour la conservation, à définir les durées et les solutions techniques.
- **La communauté scientifique :** les chercheurs du même domaine peuvent être d'une grande aide grâce au partage d'expérience, par exemple dans le choix de

standards de métadonnées ou celui d'un entrepôt de données...

Ressources :

- *DoRANum. Les acteurs de la gestion et du partage des données de la recherche. https://doi.org/10.13143/1GJK-HS77*
- *Cat OPIDoR : recense les services dédiés aux données de la recherche en France. https://cat.opidor.fr/*
- *SOS-PGD : répertoire des Services Opérationnels de Soutien à la rédaction des Plans de Gestion des Données.15 janvier 2024. https://gtso.couperin.org/gtdonnees/sos-pgd/*
- *Liste des Ateliers de la donnée : https://recherche.data.gouv.fr/fr/ateliers-de-la-donnee*



Description des données : Quelles données (types, formats, volumes, qualité…) seront collectées ou produites, et comment ?

Il s'agit de la partie la plus scientifique du PGD, dans laquelle vous allez devoir décrire les données que vous allez « manipuler » durant le projet (données préexistantes et/ou produites pour le projet).

Il ne s'agit pas de décrire toutes les données du projet, mais uniquement de décrire des jeux de données en fonction de leurs formats, de leurs types.

Concrètement, vous devez décrire :

- Le type de données (données numériques, textuelles, images, audio, vidéos etc.),
- Les formats, ouverts de préférence,
- La provenance des données (collecte de nouvelles données, conversion ou transformation de données, partage/échange de données, achat de données etc.),
- Les méthodes et outils.

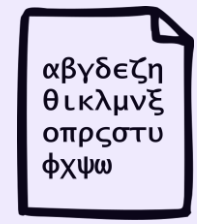
Ressources :

- *DoRANum. Définition(s) des données de la recherche. https://doi.org/10.13143/B4ZC-8E79*
- *DoRANum. L'origine et la description des données de la recherche. https://doi.org/10.13143/E9ZH-W908*

Description des données : exemple 1

Extrait du PGD public *Hospitam* *Lauriane Locatelli (ENS Lyon)* :

« Les données produites seront les suivantes :
- Documents textuels des textes antiques non lemmatisés et lemmatisés en langue grec (alphabet grec ancien avec accent, format Unicode) en .txt regroupés en dossiers par auteurs, c'est-à-dire 6 dossiers de maximum 30 Mo […] »



cnrs Inist

15

Exemple tiré du PGD public Hospitam :

On voit que le type de données, les formats (ouverts), les volumes sont bien décrits.

Pour information, la lemmatisation désigne un traitement lexical apporté à un texte en vue de son analyse.

Description des données : exemple 2



Extraits du PGD public *Transfunéraire*, *Clara Duterme (AMU)* :

« Les données collectées peuvent être :
- soit issues de recherches antérieures des membres du programme.
- soit issues d'une collecte réalisée à distance du terrain, sur des sources numérisées disponibles en ligne (articles de presse, images photographiques ou filmées). […] »

« Les données produites incluent :
- des photographies
- des enregistrements sonores
- des enregistrements vidéos
- des documents écrits tels que des notes de recherches, des schémas ou des cartes […] »

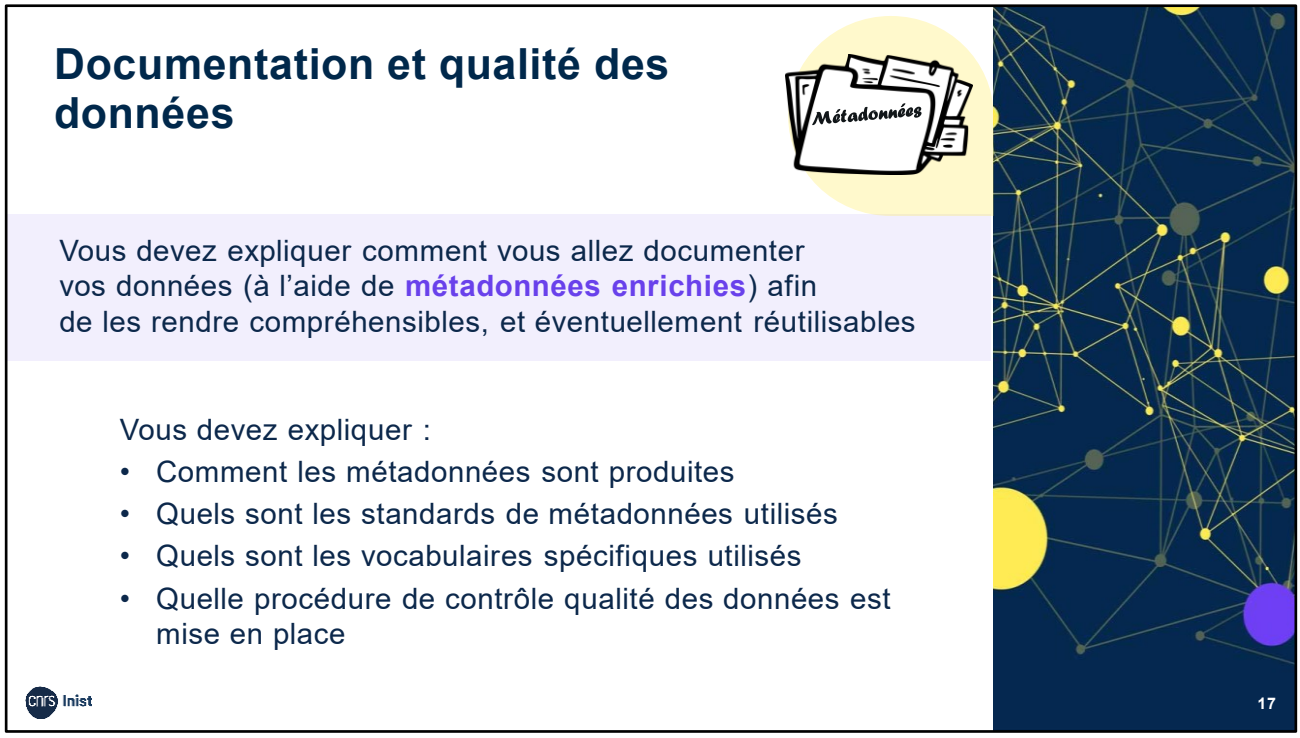
cnrs Inist

16

Exemple tiré du PGD public Transfunéraire :

Une bonne distinction est faite entre les données collectées (c'est-à-dire préexistantes) et les données produites.

L'origine des données collectées est mentionnée.



Documentation et qualité des données : quelles métadonnées, standards et formats ? Quel contrôle qualité ?

Comment les données seront identifiées, décrites, quels standards de métadonnées seront utilisés ?

Vous devez expliquer comment vous allez documenter vos données (à l'aide de métadonnées enrichies) afin de les rendre compréhensibles, et éventuellement réutilisables.

Rappel : les métadonnées **permettent de décrire plus précisément les données**. Elles facilitent la recherche et l'identification des données.

Elles peuvent donner des informations sur les conventions de nommage, les procédures et méthodes d'analyse, sur la définition de variables, d'unités de mesure etc…

Dans le PGD, vous devez expliquer :

- Comment les métadonnées sont produites ?
- Quels sont les standards de métadonnées utilisés ? Il est recommandé d'utiliser

des standards de métadonnées de votre discipline.

- Si une procédure de contrôle qualité des données est mise en place, afin d'assurer la qualité de la collecte. Préciser alors les processus comme la calibration, la répétition des échantillons ou des mesures, la capture standardisée des données, la validation de saisie des données, la revue par les pairs…

Documentation et qualité des données : exemple 1

Extrait du PGD public *IMPRINT*, *Jonathan Lenoir (CNRS) :*

« Chaque fichier de données sera accompagné au minimum d'un fichier texte (.txt) de type "Read_me.txt" indiquant les métadonnées telles que la liste des noms (header) des variables enregistrées, avec pour chaque variable (header), le type d'appareillage utilisé pour la prise de mesure, les conditions dans lesquelles les données ont été collectées, l'unité de la variable mesurée ou toutes autres informations permettant une réutilisation des données. […]

Le standard de métadonnées envisagé dans le cadre du projet est le standard "Ecological Metadata Language" (EML) […] »



cnrs Inist

18

Exemple tiré du PGD public IMPRINT :

Le fichier Read_me.txt permet de préciser des métadonnées complémentaires : noms de variables, type d'appareillage utilisé, conditions de collecte… Cela facilite la réutilisation des données. Le standard de métadonnées EML est un standard disciplinaire.

Ressources :

- *DoRANum. Métadonnées, standards, formats : Fiche synthétique. https://doi.org/10.13143/VBJS-6288*
- *DoRANum. Les standards de métadonnées : pourquoi et lequel ? https://doi.org/10.13143/Y5PY-W521*
- *Répertoires de standards de métadonnées :*
  - *DCC. Disciplinary Metadata. https://www.dcc.ac.uk/guidance/standards/metadata*
  - *RDA. Metadata Standards Catalog. https://rdamsc.bath.ac.uk/*

Documentation et qualité des données : exemple 2

Extrait du PGD public *IMPRINT, Jonathan Lenoir (CNRS) :*

« Concernant la qualité et la conformité de la collecte des données microclimatiques, il est prévu une phase d'intercalibration des capteurs HOBO UA-001-08, HOBO UA-001-64 et TMS4 en conditions contrôlées. L'installation des capteurs de T°C et d'humidité relative du sol, in-situ, suivra un protocole standardisé pour l'ensemble des sites étudiés. Concernant les données saisies sur le terrain (inventaires dendrométriques et floristiques) ou en laboratoire (détermination des espèces de carabes et analyses de sol éventuellement), elles seront validées par l'ensemble des scientifiques impliqués dans le projet. […] »



cnrs Inist

19

Exemple tiré du PGD public IMPRINT :

Les procédures de contrôle qualité et de validation sont très précises.

Documentation et qualité des données : exemple 3

Extrait du PGD public *ACTES*, *Dominique Merg Essadi (UNISTRA)* :

« La construction de la méthodologie a été soumise à :
- Des psychologues cliniciennes […]
- Des médecins exerçant en maternité […]
- Des pairs […]



L'ensemble des données ainsi recueillies seront analysées par une équipe pluridisciplinaire de chercheurs, mobilisant différentes méthodes afin de tenter de répondre à la question et de justifier l'hypothèse de recherche. »

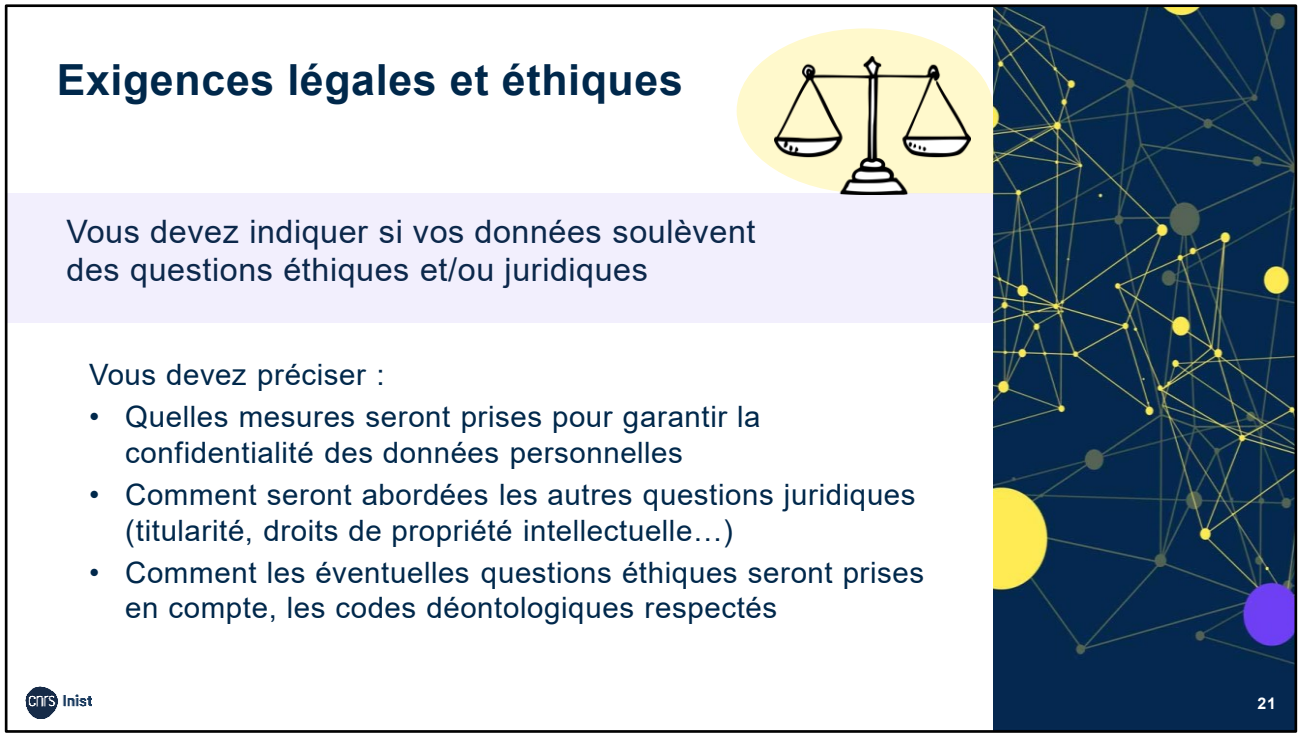
cnrs Inist

20

Exemple tiré du PGD public ACTES :

Ici, il est indiqué que la méthodologie de l'étude est contrôlée par les pairs et par des professionnels de santé qualifiés.

Quant à l'analyse des données recueillies, elle est effectuée par une équipe pluridisciplinaire de chercheurs selon différentes méthodes.



Exigences légales et éthiques : quelles questions éthiques et/ou juridiques les données soulèvent-elles ? Quelles éventuelles restrictions ?

Dans le PGD vous devez préciser :

- Quelles mesures seront prises pour garantir la confidentialité des données personnelles
- Comment seront abordées les autres questions juridiques (titularité, droits de propriété intellectuelle…)
- Comment les éventuelles questions éthiques seront prises en compte, les codes déontologiques respectés

Dans le plan, mentionnez l'existence d'éventuels comités d'éthique pour votre institution ou votre discipline.

Pour les questions RGPD, il existe dans de nombreux organismes/institutions un/une déléguée à la protection des données DPO ou DPD (en français).

Ressource

Liste des Organismes ayant désigné un(e) délégué(e) à la protection des données (DPD/DPO) : https://www.data.gouv.fr/fr/datasets/organismes-ayant-designe-un-e-delegue-e-a-la-protection-des-donnees-dpd-dpo/#/resources

Exigences légales et éthiques : exemple 1

Extrait du PGD public *Transfunéraire, Clara Duterme (AMU)* :

« Le programme de recherche « Transfunéraire » respecte le règlement général sur la protection des données (RGPD) de l'Union Européenne. Ses membres respectent les principes directeurs d'une recherche responsable, tels qu'énoncés dans le guide "Pratiquer une recherche intègre et responsable" publié par le Comité d'éthique du CNRS (COMET) en 2017.

[…] avant tout entretien en face à face, un formulaire de consentement sera signé par la personne enquêtée. Un consentement spécifique sera demandé pour les enregistrements audio. La signature du formulaire de consentement sera accompagnée d'une présentation claire des objectifs de la recherche, de la pratique de l'anonymisation et de l'usage éventuel des données à caractère personnel. »



cnrs Inist

22

Exemple tiré du PGD public Transfunéraire :

Ici, il est bien précisé que les règles sur la protection des données sont respectées. Un formulaire de consentement est proposé. Les objectifs de la recherche, de la pratique de l'anonymisation et de l'usage des données à caractère personnel sont présentés à la personne enquêtée.

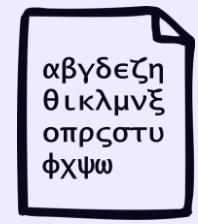
Ressources :

- *DoRANum. Aspects juridiques, éthiques, intégrité scientifique : Fiche synthétique. https://doi.org/10.13143/WZ7A-HH81*
- *DoRANum. Questions juridiques liées aux données de recherche : interview de Lionel Maurel. https://doi.org/10.13143/XJGM-HB78*
- *DoRANum. Webinaire des Tuto@mate « A qui appartiennent les données ? » - Présentation de Lionel Maurel. https://doi.org/10.13143/MNA5-YS10*
- *data.gouv.fr. Organismes ayant désigné un(e) délégué(e) à la protection des données (DPD/DPO). https://www.data.gouv.fr/fr/datasets/organismes-ayant-designe-un-e-delegue-e-a-la-protection-des-donnees-dpd-dpo/*

Exigences légales et éthiques : exemple 2

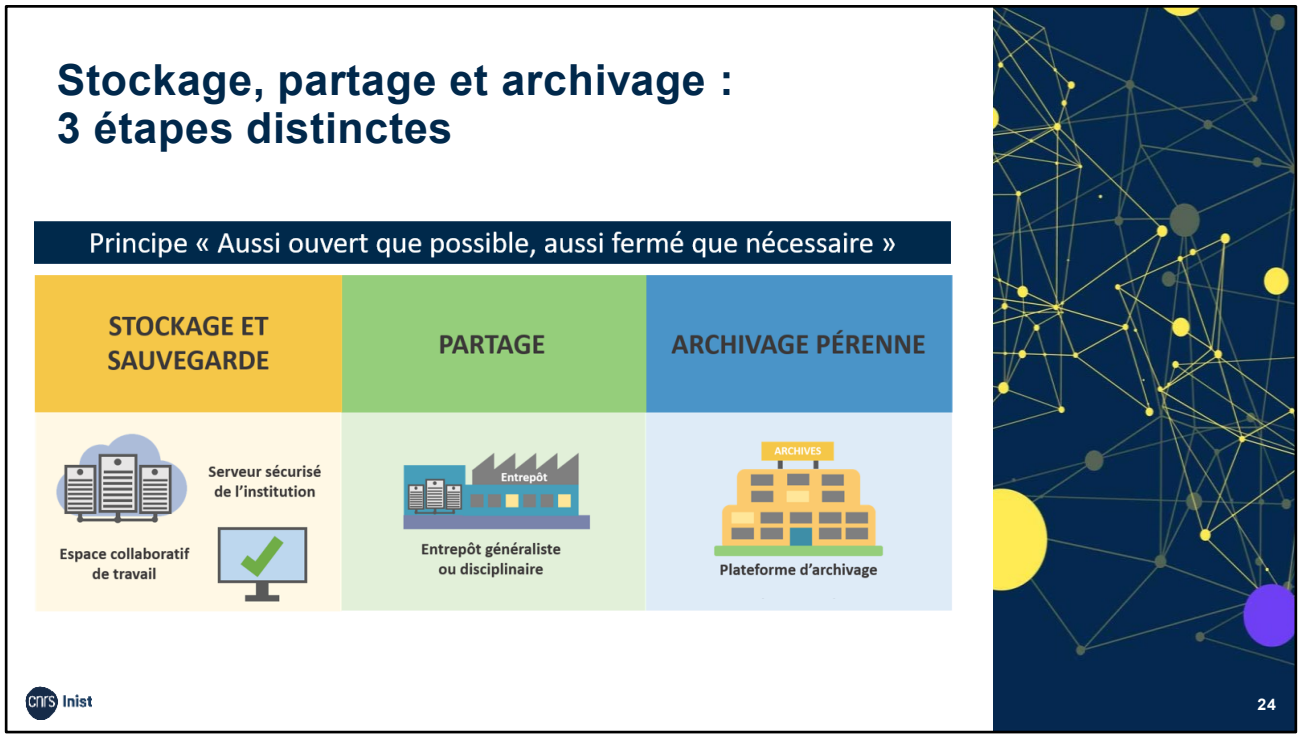
Extrait du PGD public *Hospitam*, *Lauriane Locatelli (ENS Lyon)* :

« Nos données sont des textes antiques libres de droit comme la bible ou le texte des auteurs antiques bruts, n'appartenant pas aux maisons d'éditions. Nous n'avons pas de données à caractère personnel. […] »



Exemple tiré du PGD public Hospitam :

Le projet n'est pas concerné par les exigences légales (traitement de données libres de droit et non sensibles) : les auteurs l'ont précisé dans la section.



- Le **stockage** et la **sauvegarde** sécurisés des données se font durant le projet. L'objectif est de garantir la sécurité des données et d'en **faciliter l'accès** pour l'ensemble des collaborateurs du projet.
- Le **partage** consiste à déposer les données dans un entrepôt de données afin de les rendre **accessibles facilement** et de permettre leur **réutilisation** par des chercheurs du même domaine ou d'un autre domaine, selon les principes FAIR et sur le court et le moyen terme. Le partage se fait souvent à l'issue du projet.
- L'**archivage pérenne** a pour objectif de conserver les données, d'en garantir l'**accès** et d'en préserver l'intelligibilité sur le long terme. Comme évoqué lors des journées RIP Data en octobre 2024*, le long terme est une période difficile à définir. C'est une période de temps suffisamment longue pour que des changements technologiques majeurs surviennent et mettent en péril la lisibilité des formats de fichiers.

* Journées dédiées à l'archivage, la sélection, la conservation et la suppression des données de recherche, organisées par le groupe de travail Données de la Recherche de la Mission pour les Initiatives Transverses et Interdisciplinaires du CNRS (DoReMITI)

Source :

Christine Hadrossek. Synthèse des journées d'étude RIP Data : Quelle sélection, conservation et suppression des données de recherche ?. 2024. https://hal.science/hal-04819987v1

Ressource :

DoRANum. Stockage, partage et archivage : quelles différences ?

https://doi.org/10.13143/5DAX-QP58



Stockage et sauvegarde des données : Comment les données seront-elles stockées et sauvegardées tout au long du processus de recherche ?

Vous devez expliquer quel processus de stockage des données sera mis en place durant le projet, afin d'en garantir l'accès pour l'ensemble des collaborateurs du projet.

Vous devez préciser :

- Quels seront les supports de stockage (serveur sécurisé de votre institution, espace de travail collaboratif, etc.) ?
- Quel système de nommage sera utilisé ? La fiabilité d'accès aux données pour l'ensemble d'une équipe passe par un système de nommage unique et précis des fichiers.
- Quelle sera la volumétrie des données ? Une estimation suffit.
- Comment sera organisée la sauvegarde des données ? À chaque point d'étape du projet, sélectionner les données à sauvegarder, à supprimer, définir leur durée de conservation.
- Comment seront gérées les versions ? Les différents états des données sont conservés en corrélation avec les différentes étapes de traitement ce qui permet de revenir à une version antérieure si besoin.

Stockage et sauvegarde : exemple 1



Extrait du PGD public *Transfunéraire*, *Clara Duterme (AMU)* :

« […] les données seront sauvegardées sur des cartes mémoires fournies aux participants (concerne les images photographiques et les films), […] chaque membre conserve une copie des données qu'il/ elle a produites sur son ordinateur personnel, protégé par mot de passe quand il s'agit de données à caractère particulièrement confidentiel. […] Les chercheurs peuvent également déposer leurs données sur les plates-formes numériques individuelles mises à disposition par leurs institutions pour en garantir la pérennité en cas de perte, ou procéder à leur sauvegarde sur les disques dur externes qui leur ont été fournis dans le cadre du programme de recherche […] »

cnrs Inist

26

Exemple tiré du PGD public Transfunéraire :

Dans cet exemple, les différents supports et lieux de sauvegarde sont bien précisés. Les données destinées à être partagées entre membres du programme sont stockées via une solution de stockage de type cloud sécurisé.

Stockage et sauvegarde : exemple 2

Extrait du PGD public *Hospitam*, *Lauriane Locatelli (ENS Lyon)* :

« Concernant la gestion du versioning, nous utiliserons un logiciel de gestion des versions : Git.
Git est un logiciel libre. »



cnrs Inist

Exemple tiré du PGD public Hospitam :

Nous voyons bien comment seront gérées les versions.

Stockage et sauvegarde : exemple 3

Extrait du PGD public *ACTES*, *Dominique Merg Essadi (UNISTRA)* :

« Le protocole [de stockage] des données a été soumis à la déléguée RGPD d'UNISTRA. Conformément à l'article L 1122-1 du Code de la Santé Publique (loi de Mars 2002 relative aux droits des malades).



Les données seront conservées pendant dix ans puis détruites. Si une participante à la recherche demande la destruction des données la concernant, ses données seront aussitôt détruites. »

cnrs Inist

28

Exemple tiré du PGD public ACTES :

Ici comme il est question de données personnelles, le protocole de stockage doit respecter le RGPD et être soumis à validation.

De plus, s'agissant également de données de santé, le code de la santé publique doit aussi être appliqué.



Stratégie de partage et d'ouverture des données : quelle est la politique d'accès, de partage et de réutilisation des données ?

Le PGD doit montrer que vous menez une réflexion à propos de l'ouverture et du partage de vos données, tout en gardant le principe **« aussi ouvert et FAIR que possible, aussi fermé que nécessaire »** au cœur de votre démarche.

Vous devez préciser :

- S'il y a une obligation de partage (c'est le cas par exemple pour certaines données géographiques).
- S'il y a une restriction ou interdiction de partage, par exemple pour des raisons éthiques, si les données sont confidentielles ou personnelles, soumises à une propriété de type intellectuel, commercial, ou pour des raisons de sécurité…
- Avec qui les données seront partagées (accès ouvert ou restreint) ? Sachant que même si vos données sont libres d'accès, vous pouvez décider quand même d'un embargo sur ces données avant de les rendre publiques.
- Dans quel(s) entrepôt(s) les données seront-elles déposées ? Avez-vous à disposition un entrepôt disciplinaire ou institutionnel ? Sinon vous pouvez déposer vos données dans l'entrepôt national Recherche Data Gouv.

- Quelles licences seront appliquées aux jeux de données ?
- Un identifiant persistant (ex : DOI) sera-t-il attribué ?

Partage et ouverture des données : exemple 1

Extrait du PGD public *Transfunéraire, Clara Duterme (AMU)* :



« Les données seront partagées à la fin du programme de recherche sur la plateforme Didomena, l'entrepôt de données de recherche de l'EHESS. (https://didomena.ehess.fr). Cette plateforme est dédiée aux sciences sociales et permet de partager et valoriser les données de recherche. Il n'existe pas de raison motivant un embargo, les données partagées seront immédiatement accessibles. […] »

cnrs Inist

30

Exemple tiré du PGD Transfunéraire :

Ici, on voit que l'entrepôt de données utilisé ainsi que les conditions d'accès aux données sont bien précisés.

Ressources :

- *DoRANum. Les critères pour choisir un entrepôt de données. https://doi.org/10.13143/ZQPB-9449*
- *DoRANum. Les licences de réutilisation dans le cadre de l'Open data et de la loi pour une République numérique. https://doi.org/10.13143/SSH2-ZD93*
- *Re3data (répertoire international d'entrepôts). https://www.re3data.org/*
- *Cat OPIDoR (répertoire des services français dédiés aux données de la recherche). https://cat.opidor.fr/*
- *Recherche Data Gouv : https://recherche.data.gouv.fr/fr*
- *Recherche Data Gouv, un entrepôt au service du partage et de l'ouverture des données de recherche : https://doranum.fr/depot-entrepots/recherche-data-gouv-un-entrepot-au-service-du-partage-et-de-louverture-des-donnees-de-recherche/*

Partage et ouverture des données : exemple 2

Extrait du PGD public *IMPRINT*, *Jonathan Lenoir (CNRS) :*

« Les données utilisées pour la publication des résultats seront […] publiées en libre accès au moment de la publication de l'article, ainsi que les scripts (lignes de codes) utilisés pour générer les résultats à partir des données. Ceci dans le but d'assurer la reproductibilité des résultats publiés. […] »



CNRS Inist

31

Exemple tiré du PGD IMPRINT :

Les codes sources utilisés sont partagés également.

Préservation à long terme des données



Vous devez dire si un plan de conservation des données sur le long terme a été établi

Vous devez préciser :

- Quelles sont les données (et les métadonnées associées) que vous avez sélectionnées pour l'archivage
- Quelle est la volumétrie prévue pour ces données
- Quelle sera la durée de conservation à long terme
- Quel budget est prévu pour l'archivage

cnrs Inist

32

Dans le PGD vous aurez à aborder aussi l'aspect **préservation à long terme des données** : quelle stratégie d'archivage ? Dans quelles conditions ? Pour quelles données ?

Toutes les données ne seront pas concernées par cette section.

Certains critères permettent de définir si l'archivage est nécessaire, par exemple : données coûteuses, non reproductibles, à haute valeur scientifique.

Lorsqu'on parle de « préservation à long terme », on parle d'archivage pérenne qui a pour objectif de conserver les données, et d'en garantir l'accès sur le long terme.

C'est différent du stockage durant le projet ou du partage dans des entrepôts de données.

Dans le PGD, vous devez indiquer :

- Quelles sont les données (et les métadonnées associées) que vous avez sélectionnées pour l'archivage ?
- Quel est la volumétrie prévue pour ces données ?
- Quelle sera la durée de conservation à long terme ?
- Quel budget est prévu ? L'archivage étant assez coûteux, il faut indiquer

quelles sont vos garanties de financement pour couvrir les frais. À noter que ce point en particulier peut faire l'objet d'une politique institutionnelle d'archivage (et plus largement une politique de gestion des données de recherche).

Préservation à long terme des données : exemple

Extrait du PGD public *Transfunéraire*, *Clara Duterme (AMU)*



« Le choix des données à conserver sur le long terme se fera en concertation commune entre les membres du programme ; la décision finale appartiendra aux deux coordinatrices du Programme, en fonction du caractère inédit des données, de leur valeur et de leur pertinence pour des démarches comparatives futures. […] »

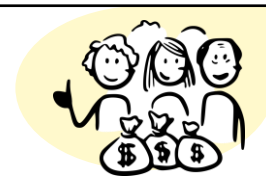
Exemple tiré du PGD public Transfunéraire :

Ici, on voit qu'au stade de rédaction de ce PGD, les questions d'archivage sont en cours de réflexion. Même si les actions ne sont pas encore bien définies, il est important de dire que la réflexion est en cours.

Ressources :

- *DoRANum. Stockage et archivage : fiche synthétique. https://doi.org/10.13143/0C4B-2743*
- *DoRANum. La minute stockage et archivage pérenne des données de recherche. https://doi.org/10.13143/77VJ-0011*
- *CINES. FACILE – Service de validation de formats. https://facile.cines.fr/*

Évaluer le coût de la gestion des données



Vous devez examiner et justifier soigneusement toutes les ressources nécessaires pour gérer et diffuser les données :

- Coût de production ou collecte des données
- Coût de préparation des données pour le partage/préservation

Exemples de coûts : frais de stockage, coût du matériel, temps de personnel, coûts de préparation des données pour le dépôt, frais d'entrepôt et d'archivage

À noter : certains coûts sont éligibles dans le cadre de projets financés

cnrs Inist

34

Vous devez examiner et justifier soigneusement toutes les ressources nécessaires pour gérer et diffuser les données.
Quelques exemples de coûts : frais de stockage, coût matériel, temps de personnel, coûts de préparation des données pour le dépôt, frais d'entrepôt et d'archivage.
Préciser les montants et comment ces coûts seront couverts.
Par exemple, pour les projets financés par l'ANR, toutes les dépenses liées à la gestion des données sont éligibles durant le projet : acquisition, collecte, stockage, personnel dédié à la gestion des données… En cas de recours à un tiers, le coût de stockage des données est admissible jusqu'à 5 ans après la date de fin scientifique du projet, sous réserve que le contrat avec ce tiers soit conclu avant la fin scientifique du projet.

Selon l'outil et le modèle de PGD choisis, les coûts peuvent être renseignés dans une section dédiée ou au fur et à mesure de la rédaction, dans

chaque partie du PGD.

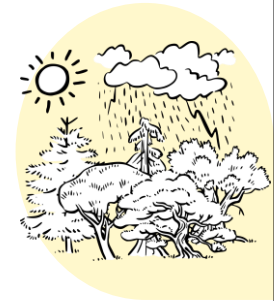
Ressources :

- *DoRANum. Le coût de la gestion des données. DoRANum. https://doi.org/10.13143/HCH2-H207*
- *OpenAIRE. Guides for Researchers. How to identify and assess Research Data Management (RDM) costs. https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs*
- *O'Connor Ryan, Delipalta Alexandra, Jones Sarah. What will it cost to manage and share my data? 21 mai 2020. https://doi.org/10.5281/zenodo.3837717*
- *EPFL Library. Cost Calculator for Data Management. https://costcalc.epfl.ch/*

Coûts de gestion des données : exemple 1

Extrait du PGD public *IMPRINT, Jonathan Lenoir (CNRS) :*

« Concernant les frais de stockage et de partage des données, nous utiliserons au maximum les plateformes de stockage gratuites, notamment pour les jeux de données de taille moyenne. Pour les jeux de données les plus volumineux (plus de 100 Go à plusieurs To), il sera éventuellement envisagé de recourir à des plateformes payantes pour la mise en libre accès (cf. Scientific Data). Pour cela, une partie du budget IMPRINT sera potentiellement alloué à la publication en libre accès (max 2000 EUR). […] »



cnrs Inist

35

Exemple tiré du PGD public IMPRINT :

Les plateformes gratuites sont privilégiées. Les plateformes payantes sont envisagées pour des jeux de données volumineux et pour favoriser le libre accès.

Coûts de gestion des données : exemple 2



Extrait du PGD public *MIP-MAP*, *Annie Andrieux (UGA) :*

Type de coût (Titre)	Montant	Etape du cycle de vie
cout plateforme (facture plateforme ME IBS)	50000 EUR	Coûts liés à la collecte/production des données - **MAP6 structure**
Matériel informatique (ordinateurs et/ou logiciels danalyse d'image)	2000 EUR	Coûts liés au traitement/analyse des données - **MAP6s as MIP**
Matériel informatique (ordinateur et/ou logiciles analyse image)	2000 EUR	Coûts liés au traitement/analyse des données - **MAP6 visualization**
Frais de personnel (PhD salary)	100000 EUR	Coûts liés à la collecte/production des données - **MAP6 structure**
plateforme (frais de plateforme)	40000 EUR	Coûts liés à la collecte/production des données - **MAP6 visualization**
Stockage (disques durs)	1000 EUR	Coûts liés à la conservation à long terme des données - **MAP6 visualization**

cnrs Inist

36

Exemple tiré du PGD public MIP-MAP :

Dans cet exemple issu du modèle commun structuré, les coûts ont été renseignés au fur et à mesure de la rédaction du PGD et pour chaque jeu de données.

L'avantage de l'utilisation du modèle structuré est qu'il génère automatiquement un tableau récapitulatif de l'ensemble des coûts.



Dans le PGD, vous devez identifier les personnes impliquées dans la gestion des données du projet et indiquer leur(s) rôle(s).
Pour chaque personne, il faut mentionner : le nom, l'institution d'appartenance, un e-mail de contact, un identifiant persistant auteur (par exemple, un identifiant ORCID) lorsque cela est possible.
Selon le projet, les responsabilités peuvent également être définies par localisation, par work-package, par espèce traitée, etc.

Selon l'outil et le modèle de PGD choisis, les responsabilités peuvent être renseignées dans une section dédiée ou au fur et à mesure de la rédaction, dans chaque partie du PGD.

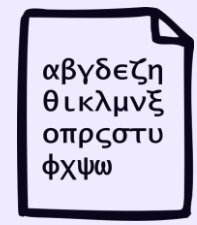
Ressource :

Dedieu, L. 2019. Rédiger un Plan de Gestion des Données en pratique. Montpellier (FRA) : CIRAD, 7 p. https://doi.org/10.18167/coopist/0066

Responsabilités : exemple 1

Extrait du PGD public *Hospitam* *Lauriane Locatelli (ENS Lyon)* :

« Le gestionnaire de données sera l'ingénieur recruté. […] La saisie des données et la production des métadonnées sera assurée par les chercheurs du projet. L'ingénieur s'occupera de la qualité des données, du stockage et de la sauvegarde ainsi que de l'archivage et du partage des données. Il sera responsable de la mise en œuvre du plan de gestion de données. […] »



Exemple tiré du PGD public Hospitam : Les rôles sont bien définis.

Responsabilités : exemple 2

Extrait du PGD public *Transfunéraire*, *Clara Duterme (AMU)* :

« Après la fin du programme de recherche, la responsabilité des données sera partagée par la Directrice de recherche au CNRS, coordinatrice du programme Transfunéraire, la Professeure à l'université de Paris, coordinatrice du programme Tranfunéraire et Clara Duterme (IGR, programme de recherche Transfunéraire). […] »



Exemple tiré du PGD public Transfunéraire : La responsabilité des données est précisée également pour l'après-projet.

Responsabilités : exemple 3

Extrait du PGD public *ACTES*, *Dominique Merg Essadi (UNISTRA) :*



Nom	Affiliation	Rôles
BACQUE Marie Frédérique	Institut de Psychologie UNISTRA	• Responsable des questions éthiques • Responsable juridique
MERG ESSADI Dominique - 0000000259485079	UNISTRA UR 3071 SULISOM	• Responsable de la qualité des données • Responsable des questions éthiques • Responsable du dépôt et de la diffusion des données • Responsable du plan • Responsable du traitement et de l'analyse des données • Responsable juridique […]
RESCH Véronique	CAE COOPRODUCTION Clairementdit.fr	• Coordinateur de projet • Personne contact pour les données • Responsable de la conservation à long terme des données • Responsable de la documentation des données • Responsable du traitement et de l'analyse des données
SILHAN Daria	UNISTRA UR 3071 SULISOM	• Responsable de la qualité des données

cnrs Inist

40

Exemple tiré du PGD public ACTES :

Avec le modèle commun structuré, il n'y a pas de rédaction d'un paragraphe dédié à la répartition des responsabilités.

L'ajout de contributeurs et la précision de leurs rôles se fait au fur et à mesure de la rédaction du PGD.

De la même manière que pour les coûts, un tableau récapitulatif des responsabilités est généré automatiquement.

03

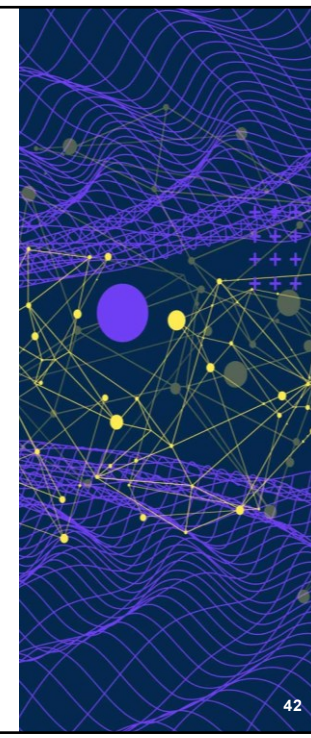
Présentation rapide de DMP OPIDoR

DMP OPIDOR, qu'est-ce que c'est ?

DMP OPIDoR → **https://dmp.opidor.fr/**

- Un outil **gratuit** développé par l'Inist-CNRS
- **Accessible** à l'ensemble de la communauté scientifique de l'ESR et à ses partenaires français ou étrangers
- **Collaboratif** : il permet les échanges entre les partenaires d'un même projet et les services d'accompagnement
- **Sécurisé** : hébergé sur un serveur à l'Inist-CNRS, en France
- **Évolutif** : l'outil évolue grâce aux retours des utilisateurs

CNRS Inist

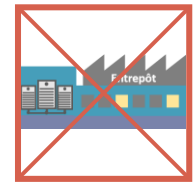


Pour vous guider dans la rédaction d'un PGD, l'Inist-CNRS propose un outil d'aide en ligne à l'élaboration de PGD.

DMP OPIDoR, c'est un outil :

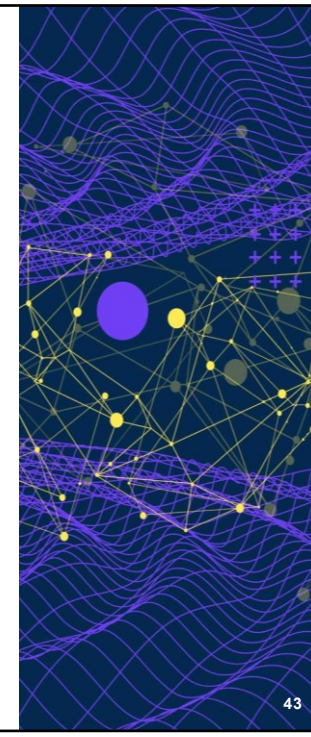
- Gratuit et développé par l'Inist-CNRS sur la base du code source DMP Roadmap (fusion du code DMP Online et DMPTool, outils britannique et américain),
- Accessible à l'ensemble de la communauté scientifique de l'ESR et à ses partenaires français ou étrangers,
- Collaboratif : il permet les échanges entre les partenaires d'un même projet et les services d'accompagnement,
- Sécurisé sur un serveur hébergé à l'Inist-CNRS, en France,
- Évolutif : l'outil évolue grâce aux retours des utilisateurs (chercheurs et personnels d'accompagnement).

À quoi ça sert ?



- Créer / rédiger un PGD à partir d'un **modèle**
- Bénéficier de **recommandations** de certains organismes
- **Collaborer** sur un PGD (partage du PGD, commentaires)
- Bénéficier d'une **assistance conseil** auprès des services d'appui de son organisme de recherche (s'il le propose)
- **Télécharger** son PGD sous différents formats (pdf, docx, html, JSON)

Attention : **DMP OPIDoR n'est pas un entrepôt de données ni une archive de PGD**



DMP OPIDoR permet de :

- Créer/rédiger un plan de gestion à partir d'un modèle.
- Bénéficier de recommandations, d'exemples, de certains organismes
- Partager votre plan avec vos collaborateurs, d'ajouter des commentaires.
- Bénéficier d'une assistance conseil auprès des services d'appui de votre organisme s'il le propose.
- Télécharger son PGD sous différents formats.

Attention, DMP OPIDoR n'est pas un entrepôt de données. Il ne sert pas à déposer des données de recherche mais à les décrire via la rédaction d'un PGD. DMP OPIDoR n'est pas non plus une archive de PGD.

Pour en savoir +



- Pour tout savoir sur le PGD :
DoRANum – Le Plan de Gestion de Données pas à pas

- Pour vous aider à améliorer le contenu de votre PGD :
DoRANum - Grilles de relecture de Plans de Gestion de Données



- DoRANum – Foire Aux Questions

- Pour retrouver toutes nos ressources pédagogiques : www.doranum.fr

Merci de votre attention

cnrs-contactpgd@cnrs.fr

contact-formation@inist.fr

Merci aux propriétaires de ces PGD de les avoir rendus publics ☺

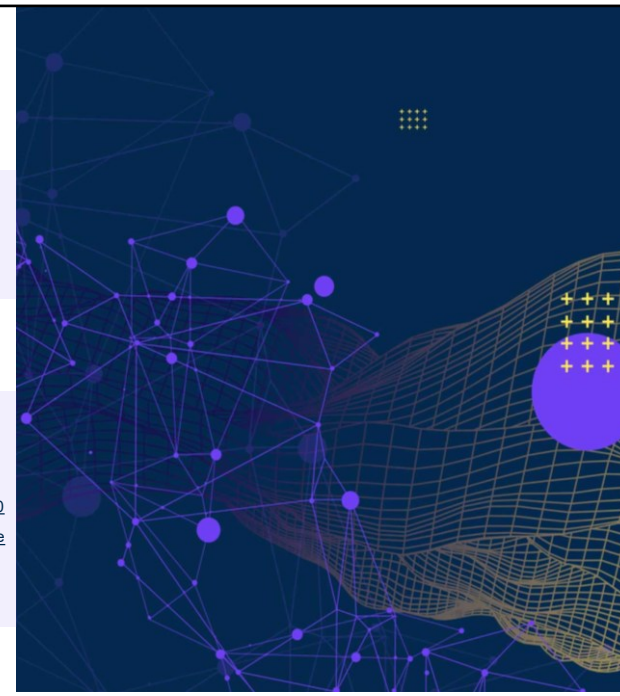
DMP Hospitam - Modèle ANR - ENS de Lyon - Lauriane Locatelli - 24/04/2020

DMP TRANSFUNERAIRE : approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles – Modèle ANR - Université d'Aix-Marseille - Clara Duterme - 22/07/2020

DMP IMPRINT – Modèle ANR – EDYSAN (CNRS) - Jonathan Lenoir – 12/03/2020

PGD ACTES – Modèle commun structuré – Université de Strasbourg – Dominique Merg Essadi – 04/03/2025

DMP MIP-MAP – Modèle commun structuré – Université Grenoble Alpes – Annie Andrieux – 08/07/2025



cnrs Inist DORANum

CC BY SA