

# Comment anticiper et planifier la gestion des données de mes recherches ? Le Plan de Gestion de Données

*Parfait, mon projet est accepté, on fonce ! Mesurons, pesons, interviewons, modélisons, calculons, photographions... oups, plus de place sur mon disque dur, ni sur le serveur et nous avons tous une version différente des fichiers, impossible de savoir laquelle est la plus à jour... Nous aurions dû anticiper pour prévoir assez de stockage et définir comment gérer au mieux toutes les données prévues, bref : nous aurions dû faire un plan de gestion des données !*

## I. A quoi sert le Plan de Gestion de Données ?

Le Plan de Gestion de Données (PGD) (ou Data Management Plan (DMP)) est un document synthétique qui aide à organiser et anticiper toutes les étapes du cycle de vie des données. Il explique comment seront gérées les données d'un projet ou d'une structure, depuis leur création, collecte ou acquisition jusqu'à leur partage, leur ouverture et leur archivage.

Le **cycle de vie des données** désigne les différentes étapes de traitement des données au cours d'un projet de recherche, depuis leur collecte jusqu'à leur réutilisation.

Le PGD permet de prévoir, le plus en amont possible, les différentes mesures nécessaires pour la gestion des données tout au long du cycle, notamment les **coûts humains et financiers**, les **obligations réglementaires ou éthiques** en matière de confidentialité, ou au contraire d'ouverture et de communicabilité, ainsi que **l'avenir de ces données après la fin d'un projet**.

Le PGD facilite la **gestion des besoins ou des imprévus** (perte/panne d'ordinateur ou de support de stockage, convention de nommage, d'organisation des arborescences plus ou moins complexes de fichiers, de dossiers, etc.) en définissant les procédures adaptées pour éviter la **perte** ou **l'inexploitabilité** des données.

[A l'échelle des projets](#), de plus en plus d'**agences publiques de financement de la recherche** demandent aujourd'hui ce document.

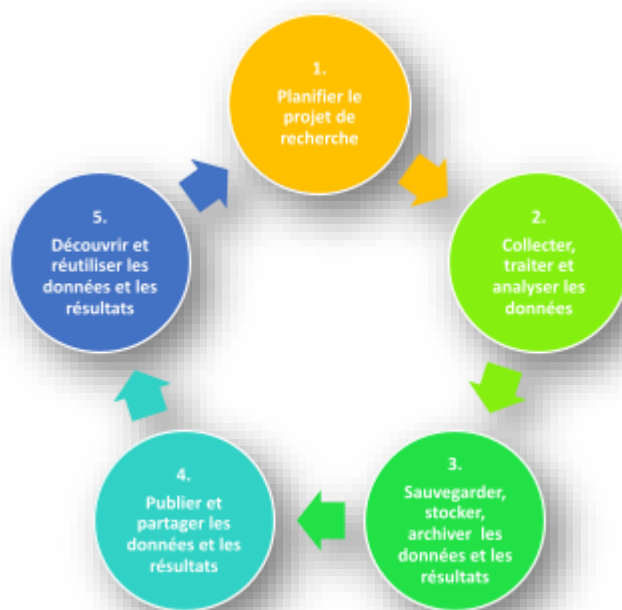
Grâce à son rôle de planification et d'anticipation, ce document est aussi salubre pour la gestion de plus long terme **des données d'une structure** (unité de recherche, équipe, laboratoire, plateforme...). Dans le cadre d'une structure, le PGD peut servir de référence pour la rédaction des PGD des projets qui se mèneront en impliquant cette structure. Le PGD de structure peut également servir de ressource informative et de guide pour les usagers, les collaborateurs et les nouveaux arrivants.

## II. De quoi se compose le Plan de Gestion de Données ?

Les questions à traiter dans un PGD sont les suivantes :

### Gouvernance, logistique et ressources :

- Identification du projet de recherche ou de la structure (titre, participants, financeurs, etc.)



Adapté d'après [University of Sidney](#) et [URFIST de Paris](#).

- Identification de la ou des personnes qui seront formellement responsables de la gestion des données du projet et de la rédaction du PGD;
- Ressources (humaines, financières, techniques, etc.) qui devront être mobilisées pour mettre en œuvre le PGD.

**Identification des données :**

- Nature des données collectées (sources extérieures, etc.) ou créées ;
- Modalités de collecte ou de création des données ;
- Modalités de contrôle de la qualité des données ;
- Nature des [métadonnées](#) ou des documents qui accompagneront les données (pour les rendre intelligibles et réutilisables).

**Mesures [légal](#), [réglementaires](#) et [éthiques](#) :**

- Identification du/des responsable(s) de la mise en œuvre des mesures nécessaires ;
- Identification des données concernées par des obligations de confidentialité ou d'ouverture ;
- Mesures prises ou à prendre (par exemple, dans le cadre de la collecte de données personnelles, les mécanismes prévus pour garantir le maintien de la confidentialité des individus dans le cadre du [RGPD](#)).

**Infrastructures et procédures :**

- Modalités de [stockage des données](#) ;
- Modalités d'accès, de partage (travail collaboratif) et de sécurisation des données ;
- Modalités de [gestion des données](#) (convention de nommage, indexation, etc.) ;
- Modalités de [sauvegarde des données](#).

**Avenir des données en fin de projet :**

- Identification des données qui devront être conservées et/ou ouvertes ;
- Modalités de conservation et d'archivage sur le long terme ;
- Modalités d'ouverture des données ([entrepôts](#), [datapapers](#)... ) ;
- Eventuelles [restrictions](#) à l'ouverture des données.

### III. Quand rédiger le Plan de Gestion de Données ?

Le PGD est un **document itératif** : il est mis à jour à différentes étapes de l'évolution du projet de recherche ou de la structure.

Dans le cadre d'un projet de recherche, une première version, dressant les éléments de cadrage générique, est produite en début de projet. Dans le cadre de [projets financés par des agences](#) telles que la Commission Européenne ou l'Agence Nationale de la Recherche (ANR) en France, cette première version est demandée **6 mois après le lancement du projet**.

Des **versions actualisées** peuvent ensuite être produites en cours de projet (évolution de la nature des données, des stratégies de partage et d'ouverture, etc.) ; dans le cas de recherches financées par des agences, elles peuvent être demandées au cours d'évaluations périodiques, et la version finale du PGD fait de plus en plus partie des livrables demandés par ces agences.

**Dans le cadre d'une structure**, le PGD est d'autant plus utile qu'il est mis à jour à chaque évolution des procédures, des infrastructures, des moyens et des décisions prises concernant les données de la structure. Inversement, le PGD de structure permet d'identifier les évolutions et prises de mesures nécessaires pour optimiser la gestion des données.

## IV. Comment rédiger un Plan de Gestion de Données ? Modèles, trucs et astuces

Un PGD se présente sous la forme d'un **formulaire à remplir**, fréquemment organisé sous forme de questions.

Compte tenu de la diversité des sujets traités et de la nécessité de consensus entre les acteurs du projet ou de la structure autour des procédures et orientations définies pour la gestion des données, le PGD est un **document collaboratif** amené à être rédigé ou enrichi à plusieurs. La rédaction de sa première version est ainsi le bon moment pour identifier les sujets à instruire, les ressources à allouer, les éventuelles difficultés à anticiper, et les points d'accord à trouver (procédures de travail, etc.). C'est également le moment d'identifier les personnes les plus compétentes à consulter ou impliquer dans la rédaction de chacune des thématiques du PGD.

Plusieurs [modèles de plans de gestion de données](#) sont proposés par les agences de financement et les organismes de recherche. Ces modèles couvrent en général les mêmes sujets, sous des formes qui peuvent varier, et en intégrant des éléments plus spécifiques relevant de politiques institutionnelles.

Ces modèles peuvent être renseignés directement sous **traitement de texte ou tableur**, mais il existe également des **applications** facilitant la gestion et le renseignement collaboratif de ces documents, comme [DMPOpidor](#), développé par l'INIST (CNRS).

AgroParisTech propose [deux trames de PGD \(sur DMPOpidor\)](#) pour les projets de recherche et les structures, en français et en anglais.

### Pour aller plus loin :

Une [checklist](#) pour vérifier qu'on a bien complété son PGD.

Une aide pratique pour estimer le [coût de la gestion des données](#).

### Exemples de PGD :

[PGD publics](#) dans DMP OPIDOR

PGD de [projets européens H2020](#)

Exemple de [PGD en sciences de l'environnement](#) (projet IMPRINT)