



FAQ - Données de la recherche, par le Méta-Atelier de la Donnée en Nouvelle-Aquitaine

Julien Baudry, Camille Desiles, Anna Loeff, Marc Odunlami, Anaïs Schmitt

► To cite this version:

Julien Baudry, Camille Desiles, Anna Loeff, Marc Odunlami, Anaïs Schmitt. FAQ - Données de la recherche, par le Méta-Atelier de la Donnée en Nouvelle-Aquitaine. Urfist de Bordeaux. 2026. <hal-05554090>

HAL Id: hal-05554090

<https://hal.science/hal-05554090v1>

Submitted on 16 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

FAQ – Données de la recherche

Méta-Atelier de la Donnée en Nouvelle-Aquitaine

Dernière mise à jour en janvier 2026

Groupe « Communication au public » du Méta-Atelier de la Donnée (coordonné par l'URFIST de Bordeaux)

Membres : Julien Baudry (Université Bordeaux Montaigne), Camille Desiles (CNRS / Les Afriques dans le Monde), Anna Loeff (Université de Bordeaux), Marc Odunlami (Université de Pau et des Pays de l'Adour), Anaïs Schmitt (La Rochelle Université)

À propos du document

Ce document propose une *Foire aux questions (FAQ)* de base pour la communication sur les *données de la recherche*. Il peut être réutilisé et adapté par les établissements membres du *Méta-Atelier de la Donnée en Nouvelle-Aquitaine*.

Cette FAQ a été réalisée entre **2024 et 2025**, à partir de plusieurs ressources nationales :

- [SoMSH – Loire](#)
- [DOREMITI – Université de Strasbourg](#)
- [Réseau CADOR – Université Lyon 3](#)
- [CNIL – Recherche scientifique hors santé](#)
- [DMP OPIDoR](#)
- [DORANUM – Données de la recherche](#)
- [GTSO Couperin – Bonnes pratiques et définitions](#)
- [Université de Lorraine – FAQ Données de la recherche](#)

Le nombre de questions a été limité afin de se concentrer sur les aspects les plus **génériques**, indépendamment des disciplines ou des structures. Les **codes et logiciels** ne sont pas abordés ici.

Citation recommandée

Julien Baudry, Camille Desiles, Anna Loeff, Marc Odunlami, Anaïs Schmitt, FAQ – Données de la recherche, Méta-Atelier de la Donnée en Nouvelle-Aquitaine, 2025.

Licence



Ce document est mis à disposition selon les termes de la licence [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

Vous êtes libre de **partager**, **adapter** et **utiliser** ce contenu, y compris à des fins commerciales, à condition de **créditer l'auteur**.

© 2025 — Méta-Atelier de la Donnée en Nouvelle-Aquitaine (MAD-NA).

Table des matières

1	Questions générales	5
1.1	Qu'est-ce qu'une donnée de recherche ?	5
1.2	Une activité de recherche génère-t-elle toujours des données ?	5
1.3	Est-on obligé de rendre publiques ses données de recherche ?	5
1.4	Dans le cadre d'un projet international ou avec un partenaire privé, comment faire si les règles diffèrent ?	6
2	Stockage, partage, archivage et <i>data papers</i>	7
2.1	Stockage, archivage, partage... comment s'y retrouver ?	7
2.2	Où stocker mes données pendant le projet de recherche ?	7
3	Entrepôt de données	9
3.1	Qu'est-ce qu'un entrepôt de données ?	9
3.2	Entrepôt, plateforme, portail, catalogue, infrastructure : quelle différence ?	9
3.3	Est-on obligé d'utiliser un entrepôt de données ? / Pourquoi déposer ses données ?	9
3.4	Comment préparer mes données pour les rendre plus faciles à partager et réutiliser ?	9
3.5	Le partage des données dans un entrepôt permet-il l'archivage à long terme ?	10
3.6	Comment choisir son entrepôt ?	10
3.7	Peut-on mettre un embargo sur ses jeux de données ?	11
3.8	Quelle licence choisir ?	11
4	Archivage	12
4.1	Qu'est-ce que l'archivage ?	12
5	<i>Data Papers</i>	13
5.1	Qu'est-ce qu'un <i>data paper</i> ?	13
6	Plan de Gestion de Données (PGD)	13
6.1	Généralités sur les PGD	13
6.1.1	Qu'est-ce qu'un PGD ?	13
6.1.2	Qui rédige un PGD ?	13
6.1.3	Quand doit-on le rédiger ?	13
6.1.4	Pourquoi mettre en place un PGD même sans financement spécifique ?	13
6.1.5	Quels sont les bénéfices d'un PGD ?	14
6.1.6	Comment identifier les produits de recherche à inclure dans un PGD ?	14
6.2	Plateformes et modèles de PGD	14
6.2.1	Qu'est-ce que DMP OPIDoR ?	14
6.2.2	Quels sont les avantages d'utiliser une plateforme ?	14
6.2.3	Quel modèle choisir pour rédiger un PGD ?	14
6.2.4	Où trouver des exemples de PGD ?	15
6.3	Contenu d'un PGD	15
6.3.1	Quels sont les éléments clés d'un PGD ?	15
6.3.2	Comment décrire les métadonnées dans un PGD ?	15
7	Identifiants pérennes	15
7.1	Pourquoi les données doivent-elles disposer d'un identifiant pérenne ?	15
7.2	Comment obtenir un DOI ?	16
7.3	Comment citer mes données ?	16

8 Aspects juridiques	17
8.1 Suis-je propriétaire de mes données de recherche ?	17
8.2 Traiter des données personnelles en recherche : comment procéder ?	17

1 Questions générales

1.1 Qu'est-ce qu'une donnée de recherche ?

Source retravaillée : [So MSH](#), [OCDE](#)

« Les données de recherche sont des enregistrements factuels (chiffres, textes, images, sons) utilisés comme sources principales pour la recherche scientifique et généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche. » — *OCDE*

Une donnée de recherche est généralement de forme **numérique**, mais certaines définitions s'élargissent **au-delà du numérique**.

Voir, par exemple, [cette note](#)

1.2 Une activité de recherche génère-t-elle toujours des données ?

Source retravaillée : [GTSO Couperin](#)

Les **données de recherche** peuvent revêtir une multitude de formes : **quantitatives**, **qualitatives**, relevés de terrain sous forme de fichiers tabulés, entretiens enregistrés, fichiers **audio** ou **vidéo**, etc. Il arrive que, pour certains projets, le terme même de « *données* » paraisse inadapté.

Dans un tel cas, il faut remettre en perspective les **matériaux** sur lesquels s'appuie la recherche : prises de notes, brouillons d'article, listes bibliographiques structurées, corpus d'images ou de textes consultés sur le Web, etc.

Toutes ces informations ne sont pas nécessairement destinées à être **publiées** ni **accessibles à toutes et à tous**, mais elles constituent la **matière première de la recherche**, qui peut être décrite dans un **Plan de Gestion de Données (PGD)**.

Astuce

Même non diffusés, **notes**, **brouillons**, **corpus d'images**, **bibliographies structurées** constituent la **matière première** de la recherche et peuvent être décrits dans un **Plan de Gestion de Données (PGD)**.

1.3 Est-on obligé de rendre publiques ses données de recherche ?

Source retravaillée : [SoMSH](#)

Cette question s'envisage en considérant **la règle générale** et **les exceptions**.

La **règle générale** est que, juridiquement, les **données de recherche** sont considérées comme des **documents administratifs** :

produites dans le cadre d'une mission de service public, leur communicabilité est définie par le **Livre III du Code des Relations entre le public et l'administration (CRPA)**.

L'**ouverture des données** est une **obligation légale** pour toute recherche financée sur **fonds publics**.

Les **exceptions** permettent ensuite d'en cadrer l'application :

l'obligation ne concerne que les **documents finalisés**, et non les **documents préparatoires** ou les

carnets de laboratoire, dont la diffusion reste à l'appréciation des chercheuses et chercheurs.

L'ouverture peut être restreinte lorsque les données :

- Concernent des **secrets protégés par la loi** ;
- Comportent des **données personnelles ou sensibles** ;
- Sont soumises à un **droit d'auteur** ;
- Proviennent d'une recherche **cofinancée par des fonds privés** ;
- Relèvent d'une **Zone à Régime Restrictif (ZRR)**.

En ZRR, le processus de partage des données doit être le même que pour un article scientifique.

Voir la note : [Partage des données en ZRR \(Université de Poitiers, 2025\)](#)

! Principe directeur

« Aussi ouvertes que possible, aussi fermées que nécessaire. »

1.4 Dans le cadre d'un projet international ou avec un partenaire privé, comment faire si les règles diffèrent ?

Les pratiques de **gestion**, de **conservation** et de **diffusion** des données peuvent diverger selon les pays ou les partenaires.

Tant que chacune des parties **comprend les pratiques de l'autre**, cela ne pose pas de difficulté majeure :

le **Plan de Gestion de Données (PGD)** permet précisément de **documenter** des modalités différentes de stockage, de partage ou de diffusion.

Il n'a pas vocation à imposer une méthode unique, mais à **explicitier les règles de chaque acteur**.

Par ailleurs, l'**accord de consortium** entre partenaires du projet précise la **répartition des responsabilités** et peut **autoriser ou restreindre** certaines dispositions du PGD.

💡 Bonnes pratiques

- Vérifier que les clauses de confidentialité et de propriété intellectuelle sont compatibles avec le partage des données ;
- Décrire clairement les règles de diffusion dans le **PGD commun** ;
- S'appuyer sur le **contrat de consortium** pour formaliser les modalités de gestion.

2 Stockage, partage, archivage et *data papers*

Sources retravaillées : SoMSH, Recherche Data Gouv, LabRADoR – Sorbonne Université

2.1 Stockage, archivage, partage... comment s'y retrouver ?

Les dispositifs utilisés pour **gérer les données de recherche** — les sécuriser, les travailler collectivement, les préserver ou les diffuser — diffèrent selon :

- le **moment du cycle de vie** de la donnée (production, traitement, diffusion, conservation),
- et le **degré d'ouverture** envisagé.

Ces notions (*stockage*, *archivage*, *diffusion*, *exposition*, etc.) recouvrent donc des **étapes et des besoins distincts**, pour lesquels il est essentiel de choisir **l'outil approprié**.

Comprendre ces distinctions permet d'assurer une gestion cohérente et conforme aux exigences **FAIR**.

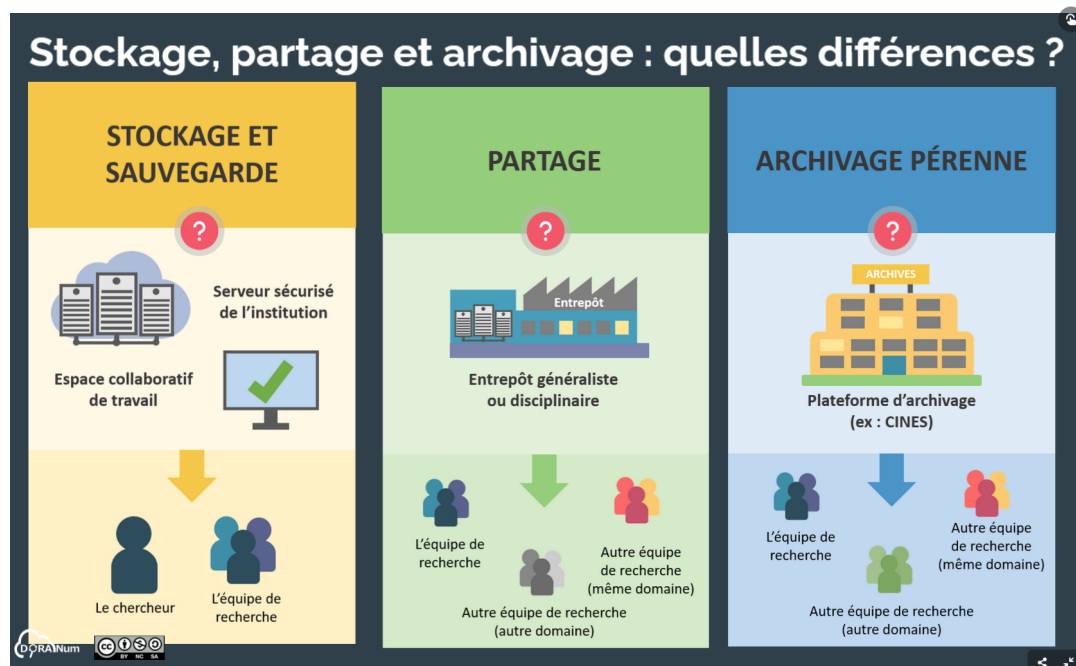


Figure 1: **Stockage, partage et archivage : quelles différences ?**

Source : « Stockage, partage et archivage, quelles différences », DORANUM.

<https://view.genially.com/5ddfe6579cf2560f66396c92>

2.2 Où stocker mes données pendant le projet de recherche ?

Pendant la durée du projet, il est essentiel d'adopter de **bonnes pratiques de stockage** afin d'assurer :

- la **sécurité** des fichiers,
- leur **accessibilité** pour l'ensemble des collaborateurs,
- ainsi que la **préservation de leur intégrité, intelligibilité et lisibilité**.

Cela implique notamment :

- l'utilisation de **supports fiables** (éviter les supports fragiles comme les clés USB) ;
- la **préférence pour des solutions institutionnelles** (serveurs ou clouds gérés par l'établissement) ;
- l'exclusion des plateformes **non conformes au RGPD** (Google Drive, Dropbox, etc.) ;
- le **chiffrement des fichiers sensibles** et une **gestion rigoureuse des droits d'accès** lorsque les données sont personnelles ou confidentielles.

Résumé des bonnes pratiques

- Éviter les supports fragiles (ex. : clés USB) ;
- Proscrire les plateformes non conformes au RGPD (ex. : Google Drive, Dropbox) ;
- Utiliser des solutions institutionnelles ;
- Chiffrer les fichiers sensibles.

La règle du 3-2-1

Pour garantir la sécurité et la pérennité des données, appliquez la **règle du 3-2-1** :

- **3 exemplaires** : créez deux copies de votre fichier original ;
- **2 supports** : stockez vos copies sur au moins deux supports différents ;
- **1 copie à distance** : conservez au moins une copie hors site (par exemple sur un serveur distant institutionnel).

Bon à savoir

Chaque établissement ou structure de recherche peut proposer **ses propres solutions de stockage**.

Avant de choisir un service ou un outil, **renseignez-vous sur les recommandations institutionnelles** afin d'assurer la conformité et la sécurité de vos données.

3 Entrepôt de données

Sources retravaillées : [Research Data Alliance \(RDA\)](#) · [Comité pour la Science Ouverte \(CoSO\)](#)
· [DORANum](#) · [COOPIST – CIRAD](#) · [SoMSH](#)

3.1 Qu'est-ce qu'un entrepôt de données ?

Un **entrepôt de données de recherche** (*data repository*) est une **base de données soutenue par une infrastructure** qui héberge, conserve et rend accessibles des données de recherche en respectant des principes de gestion rigoureux (**FAIR**).

Une fois déposés, les jeux de données se voient attribuer un **identifiant pérenne** et une **licence de réutilisation**, et sont **décrits par des métadonnées** permettant de les retrouver facilement.

Il existe de nombreux entrepôts, aux périmètres plus ou moins larges.

Le chercheur souhaitant déposer et partager ses données doit se tourner **en priorité vers un entrepôt disciplinaire**.

Si ce dernier n'existe pas, il peut se tourner vers **l'entrepôt de sa tutelle** ou **Recherche Data Gouv**, au sein duquel de plus en plus d'établissements disposent d'un espace institutionnel.

(Source : [FAQ GTSO Couperin](#))

3.2 Entrepôt, plateforme, portail, catalogue, infrastructure : quelle différence ?

- **Entrepôt de données** : permet l'hébergement, la gestion et la curation des données.
Exemples : [Recherche Data Gouv](#), [Nakala](#), [Quetelet-Progedo-Diffusion](#), [EaSy Data](#), [SEANOE](#).
- **Portail ou Catalogue** : point d'accès centralisé vers différentes ressources et outils.
Recense et expose des métadonnées issues d'entrepôts et d'autres bases de données, mais n'héberge généralement pas directement les jeux de données.
Exemples : [Catalogue Recherche Data Gouv](#), [Isidore](#), [GBIF](#), [OBIS](#).
- **Infrastructure de recherche** : ensemble structuré de ressources et de services facilitant la recherche et la gestion des données à grande échelle.
Exemples : [Recherche Data Gouv](#), [Progedo](#), [Huma-Num](#), [Data Terra](#), [ILICO](#).

3.3 Est-on obligé d'utiliser un entrepôt de données ? / Pourquoi déposer ses données ?

Rendre accessibles ses données dans un entrepôt est l'une des **pratiques majeures de la Science Ouverte**.

Cela contribue à la **transparence des processus de recherche** et à la **reproductibilité des résultats**.

Partager ses données permet à d'autres équipes de les réutiliser sans avoir à les régénérer, assurant ainsi **un gain de temps** et **une meilleure utilisation des fonds publics**.

Déposer ses données dans un entrepôt assure leur **préservation**, leur **visibilité**, leur **accès** et facilite leur **partage**, **réutilisation** et **citation** grâce à l'attribution d'un **identifiant numérique unique et pérenne** (souvent un DOI).

3.4 Comment préparer mes données pour les rendre plus faciles à partager et réutiliser ?

Il est essentiel de penser, dès le début du projet, au **cycle de vie des données**.

Elles ne seront pas traitées de la même façon selon qu'elles doivent être partagées en **cercle restreint** ou en **accès public**.

Plusieurs bonnes pratiques sont à adopter pour garantir la **réutilisabilité** et la **traçabilité** des données, y compris pour le chercheur lui-même après plusieurs années.

Bonnes pratiques à adopter

- Dans un document, décrire précisément les données et la méthodologie employée pour les produire (nom et version du logiciel, méthodologie scientifique, sources...). En cas de partage, ce document doit accompagner les données (par exemple, un fichier nommé *readme*).
- Veiller à utiliser des **formats et standards ouverts**, communément utilisés par la communauté.
- **Structurer** les fichiers selon une arborescence claire et transparente.
- **Nommer** les fichiers selon une convention cohérente.
- Fournir une **référence bibliographique** si les données sont liées à une publication.
- S'assurer que **les données personnelles** ne sont pas conservées ni transmises.

La rédaction d'un **Plan de Gestion de Données (PGD)** est un bon moyen de s'assurer que ces pratiques sont partagées et appliquées par l'ensemble du projet.

3.5 Le partage des données dans un entrepôt permet-il l'archivage à long terme ?

Le but d'un **entrepôt** n'est pas d'archiver les données à long terme.

Un entrepôt permet la **préservation** et la **diffusion** des données, et en assure l'**accessibilité** pour une durée d'environ **dix ans**.

L'**archivage pérenne** nécessite des infrastructures dédiées, comme les **services d'archivage institutionnels** ou **nationaux**.

En France, le **CINES** (Centre Informatique National de l'Enseignement Supérieur) propose un service d'archivage électronique pour les **données scientifiques**, **patrimoniales** ou **administratives**.

3.6 Comment choisir son entrepôt ?

Il existe plusieurs catégories d'entrepôts : **internationaux**, **nationaux** ou **institutionnels** ; **génériques**, **disciplinaires** ou **thématiques**.

Les **entrepôts disciplinaires** sont à privilégier lorsqu'ils existent.

Sinon, les **entrepôts institutionnels** ou **nationaux** constituent une bonne alternative (ex. : *Recherche Data Gouv* en France).

Il est recommandé de **suivre les pratiques de sa communauté scientifique** et d'utiliser des entrepôts **disciplinaires** ou **reconnus par son institution**.

Guides méthodologiques utiles

- [DORANum](#)
- [re3data – Répertoire international d'entrepôts de données](#)
- [FAIRsharing](#)
- [Guide Recherche Data Gouv](#)

Critères de sélection

Vérifiez que l'entrepôt choisi est :

- adapté au **type** et à la **taille** des fichiers à déposer ;
- conforme aux **recommandations du bailleur**, de l'institution ou de la revue ;
- reconnu par la **discipline** et la **communauté scientifique** (certains sont certifiés *CoreTrustSeal*) ;
- attribuant automatiquement un **identifiant pérenne** (DOI) ;
- garantissant la **pérennité d'accès**, le **partage** et la **réutilisation** ;
- **gratuit** ou avec des coûts de dépôt acceptables ;
- proposant des **modalités d'accès adaptées** (libre, restreint, sur demande, différé, sous embargo) ;
- offrant une **licence de diffusion appropriée** (*voir section ci-dessous*).

Le **Collège des Données de la Recherche du CoSO** a publié une note détaillant une **méthodologie d'identification des entrepôts thématiques de confiance** ainsi qu'une **liste non exhaustive** de ces entrepôts qui figure à [cette adresse](#).

3.7 Peut-on mettre un embargo sur ses jeux de données ?

Il est possible de **restreindre l'accès** aux données pendant le projet ou au moment du dépôt, à condition de choisir un entrepôt qui le permet.

L'**embargo** peut être **total** pour une durée donnée (avec ouverture automatique ultérieure), ou **partiel**, en accordant par exemple un **accès temporaire à des relecteurs** dans le cadre d'une évaluation d'article.

3.8 Quelle licence choisir ?

Sources retravaillées : [FAQ DoReMITI](#) · [Etalab](#) · [Légifrance](#)

Conformément au **Décret n° 2017-638 du 27 avril 2017** relatif aux licences de réutilisation gratuite des informations publiques,

les données publiques doivent être placées sous :

- la **Licence Ouverte Etalab**, ou
- la **Licence ODbL**.

La **Licence Ouverte Etalab** autorise la réutilisation, la reproduction, la modification et la redistribution des données, y compris à des fins commerciales,

sous réserve de **mentionner le nom du producteur** et la **date de dernière mise à jour**.

Elle est **compatible** avec toute licence exigeant la **mention de paternité**, notamment la **licence CC BY**

([Creative Commons Choisir sa licence](#)).

4 Archivage

Source : [Recherche Data Gouv](#)

4.1 Qu'est-ce que l'archivage ?

Dans le cadre de la **gestion des données de recherche**, la notion d'**archivage** est souvent mal comprise, car elle est parfois confondue avec des concepts voisins tels que le **stockage**, le **partage** ou la **publication** des données.

Pourtant, l'archivage a une **définition et des finalités bien spécifiques**. Il s'agit d'un processus distinct, essentiel pour garantir à la fois la **conservation** et la **conformité réglementaire** des données.

i Qu'est-ce que l'archivage n'est pas ?

- **L'archivage n'est pas du stockage** : Le stockage consiste à enregistrer une information sur un support physique (disque dur, bande magnétique, serveur de laboratoire, serveur institutionnel, etc.) pour y accéder de façon individuelle.
Il vise un usage immédiat ou à court ou moyen terme, tandis que l'archivage garantit la **pérennité** et l'**intelligibilité** des données.
- **L'archivage n'est ni du partage ni de la publication** : ...Lorsque des données sont déposées dans un entrepôt comme *Recherche Data Gouv*, l'objectif est de les rendre accessibles à la communauté scientifique, en favorisant leur **réutilisation** ou leur **valorisation**.
L'archivage, lui, s'inscrit dans une **démarche patrimoniale et réglementaire**.

L'**archivage des données de recherche** peut être défini comme l'ensemble des pratiques visant à **préserver les données dans le temps**, en garantissant leur **intégrité**, leur **authenticité** et leur **intelligibilité**, pour des besoins de **preuve**, de **mémoire** ou d'.

En France, le **Code du patrimoine (article L211-1)** rappelle que l'archivage est une **obligation légale** pour l'ensemble des documents, y compris les données issues de la recherche publique.

Les données archivées à long terme doivent être versées dans un **service public d'archives** sous le contrôle scientifique et technique de l'administration des archives.

« La définition juridique des archives est beaucoup plus large que l'acception commune qui consiste à considérer les "archives" comme des documents anciens.

Elle recouvre tous types de documents et de données, quel que soit leur forme et leur support.

Ainsi, sont des archives, dès leur création, une infinité de types de données de recherche (carnets ou photos de terrain, enregistrements d'entretiens, bases de données, algorithmes...).

Leur gestion se fait donc en concertation avec l'ensemble des acteurs tout au long du cycle de vie de la donnée afin, par exemple, d'assurer leur accessibilité ou de bien conserver celles qui ne peuvent pas être ouvertes immédiatement.

L'archivage est un ensemble de méthodes, processus et outils mis en œuvre pour gérer la conservation et l'utilisation des documents et informations à court, moyen et long termes.

À ce titre, les services d'archives d'établissement proposent aux chercheurs, en complément des circuits de publication en accès ouvert (RDG, HAL, Zenodo, Nakala...), des solutions de conservation et de communication des données de recherche qui permettent d'assurer la conformité avec les différentes réglementations (Code du patrimoine, RGPD, intégrité scientifique...) pour constituer le patrimoine scientifique de demain. »

Contrairement au **simple stockage**, l'archivage vise la **conservation durable** des données, dans le respect des réglementations (**Code du patrimoine**, **RGPD**, etc.).

L'archivage est une **obligation légale en France** :

il concerne **tous types de données, dès leur création**.

Il est **conseillé** de prendre contact avec les **archives de son établissement** et/ou avec le **CINES**.

5 Data Papers

Source retravaillée : [Université de Strasbourg – DoReMITI](#)

5.1 Qu'est-ce qu'un *data paper* ?

Le **data paper** (en français : *publication de données*) est une **publication scientifique, validée par les pairs**.

Il est rédigé autour d'un **jeu de données rendu accessible**, qu'il **décrit** et **documente** : conditions d'acquisition, usages potentiels, contexte de production.

Les *data papers* sont généralement publiés dans des **revues spécialisées** dédiées à ce type d'articles.

6 Plan de Gestion de Données (PGD)

6.1 Généralités sur les PGD

Sources : [DMP OPIDoR](#) & [DORANUM](#)

6.1.1 Qu'est-ce qu'un PGD ?

Un **PGD** est un **document évolutif** qui décrit comment les données de recherche (produites ou réutilisées) sont **collectées, organisées, stockées, partagées** et **préservées** tout au long du cycle de vie d'un projet.

6.1.2 Qui rédige un PGD ?

Tout chercheur impliqué dans un projet de recherche peut mettre en place un PGD.

Il est généralement rédigé par le **porteur du projet**, mais peut aussi être confié à un **gestionnaire de données**.

Une personne coordonne la rédaction et les mises à jour, mais cela demande l'implication de **toute l'équipe**.

6.1.3 Quand doit-on le rédiger ?

Le PGD doit être rédigé **dès la soumission du projet**.

C'est un **document vivant**, mis à jour régulièrement tout au long du projet.

6.1.4 Pourquoi mettre en place un PGD même sans financement spécifique ?

Un PGD formalise et structure la gestion des données d'une équipe, d'un laboratoire ou d'une institution, facilitant la **gestion**, le **partage** et la **traçabilité** des données.

Sa rédaction est **conseillée même en l'absence d'obligation** de la part d'un financeur.

6.1.5 Quels sont les bénéfices d'un PGD ?

Un PGD permet de :

- **Anticiper et planifier** la gestion des données en amont du projet ;
- **Faciliter** le partage et la préservation des données ;
- **Optimiser** le suivi des obligations juridiques et éthiques (RGPD, droits d'auteur, etc.) ;
- **Garantir** la qualité, la transparence et la reproductibilité des résultats de recherche ;
- **Accroître** la valeur scientifique des données et encourager la collaboration.

Un bon PGD contribue à la **rigueur scientifique**, à la **visibilité** des résultats et à la **gestion responsable** des données, ce qui est exigé par de nombreux **organismes financeurs**.

6.1.6 Comment identifier les produits de recherche à inclure dans un PGD ?

Le PGD couvre **toutes les productions issues d'un travail scientifique**, autres que les publications, qu'elles soient **créées** ou **réutilisées**, **brutes** ou **traitées**.

Cela inclut notamment :

- Les **jeux de données** ;
- Les **codes sources**, **scripts** ou **logiciels développés** ;
- Les **protocoles expérimentaux** ;
- Les **modèles**, **ontologies**, **notebooks** ;
- Les **documents de travail**, **questionnaires**, etc.

6.2 Plateformes et modèles de PGD

6.2.1 Qu'est-ce que DMP OPIDoR ?

DMP OPIDoR (*Data Management Plan pour l'Optimisation du Partage et de l'Interopérabilité des Données de Recherche*) est un **outil national** développé par l'**Inist-CNRS**, largement utilisé par la recherche française.

Il permet de créer des **PGD en ligne**, de façon **simple**, **collaborative** et **interopérable**.

Accès à [cette adresse](#)

6.2.2 Quels sont les avantages d'utiliser une plateforme ?

- Travail **collaboratif** facilité et sécurisé ;
- Accès à plusieurs **modèles** de PGD ;
- **Exportation** sous différents formats (PDF, texte...) ;
- **Interopérabilité** avec d'autres outils (par ex. *OASIS*, outil de suivi scientifique ANR) ;
- **Recommandé** par de nombreux organismes de recherche.

6.2.3 Quel modèle choisir pour rédiger un PGD ?

Le choix du modèle dépend des **exigences du financeur**, de l'**organisme de recherche** ou des **spécificités disciplinaires** du projet.

Le modèle commun « **Science Europe – modèle structuré** » est recommandé par le **réseau Science Ouverte** et plusieurs **agences de financement françaises** :

ADEME, ANR, ANRS-MIE, Anses, FRM, INCa, etc.

Ce modèle est disponible directement sur [DMP OPIDoR](#).

6.2.4 Où trouver des exemples de PGD ?

Des exemples sont disponibles sur :

- [DMP OPIDoR](#) → rubriques [Exemples de PGD](#) et [PGD publics](#)
- [Zenodo](#)
- [HAL](#)
- [Corpus de PGD disciplinaires](#) proposé par [RDA France](#)

6.3 Contenu d'un PGD

6.3.1 Quels sont les éléments clés d'un PGD ?

Un PGD contient en général **six sections principales** :

1. **Description** des données et collecte ou réutilisation de données existantes
2. **Documentation** et qualité des données
3. **Exigences légales et éthiques**, code de conduite
4. **Traitement et analyse** des données
5. **Stockage et sauvegarde** des données pendant le projet
6. **Partage et conservation à long terme**

6.3.2 Comment décrire les métadonnées dans un PGD ?

Les **métadonnées** doivent suivre les **standards disciplinaires**, garantir le **bon référencement** et la **réutilisation** des données.

Pour aller plus loin, voir :

- [DORANUM](#)
- [DMP OPIDoR](#).

7 Identifiants pérennes

Sources : [DoReMITI](#), [IRD Data](#), [DORANUM](#)

7.1 Pourquoi les données doivent-elles disposer d'un identifiant pérenne ?

Les données de recherche doivent disposer d'un **identifiant pérenne** — souvent un **DOI** (*Digital Object Identifier*) — afin d'être :

- **Facilement trouvables** sur le Web ;
- **Visibles et accessibles** au même titre que les publications ;
- **Citées et reliées** aux articles ou à d'autres produits de recherche.

7.2 Comment obtenir un DOI ?

Les deux principales agences d'enregistrement sont :

- [DataCite](#)
- [Crossref](#)

Lors du dépôt dans un **entrepôt de données**, un DOI est souvent **attribué automatiquement** (ex. : *Recherche Data Gouv*).

7.3 Comment citer mes données ?

Les entrepôts proposent en général un **format de citation** prêt à l'emploi.

Si ce n'est pas le cas, veuillez à inclure les éléments essentiels :

- **auteur**,
- **titre**,
- **date**,
- **identifiant**,
- **licence**.

[DataCite](#) propose également des **modèles de citation normalisés** lors de l'attribution d'un DOI.

8 Aspects juridiques

Sources retravaillées :

- Université Rennes 2 – Comprendre la propriété des données dans un contexte d'ouverture
- Bonnes pratiques de gestion des données de la recherche – OPIDoR (PDF)
- Protection des bases de données – Union européenne
- Université Rennes 2 – Respecter le régime des données personnelles et sensibles
- Guide des données personnelles et RGPD – MSH Lorraine (2024)
- Université de Pau et des Pays de l'Adour – PGD et données de la recherche
- SupDPO – Recommandations chercheurs (version 1.2-1-17)

8.1 Suis-je propriétaire de mes données de recherche ?

Un chercheur n'est **pas systématiquement titulaire du droit de propriété intellectuelle** sur les données qu'il produit — ce droit permettant d'en décider les modes de **reproduction** et de **diffusion**.

À retenir

- Pour qu'un chercheur soit **titulaire du droit d'auteur** sur les données qu'il a produites, celles-ci doivent présenter un **caractère original** (les choix reposant sur un savoir-faire technique ou des règles formelles ne sont pas protégés) et **porter la personnalité de l'auteur**.
Les **personnels BIATSS**, de par leur statut différent, **cèdent automatiquement le droit d'exploitation** des données à leur employeur.
- Si les données **ne sont pas originales** et sont intégrées à une **base de données**, c'est le **régime du droit des bases de données** qui s'applique (*droit sui generis*).
Dans ce cas, le titulaire des droits est le **producteur de la base**, généralement **l'établissement ayant financé la recherche**.
- La **Loi pour une République numérique (2016)** établit que les **données de recherche financées sur fonds publics** sont des **données administratives** qui doivent être **ouvertes par défaut**, sauf exceptions :
 - application de **droits de propriété intellectuelle** ;
 - **secret-défense** ou **données sensibles** ;
 - présence de **données personnelles** ou **couvertes par le secret professionnel**.

En cas de **projet partenarial**, la question des **droits et devoirs sur les données** doit être **anticipée** et formalisée, par exemple dans un **accord de consortium**.

8.2 Traiter des données personnelles en recherche : comment procéder ?

Une **donnée à caractère personnel** désigne toute **information** permettant d'**identifier directement ou indirectement** une personne physique.

Dans le cadre de vos travaux de recherche, vous pouvez être amené·e à **traiter des données personnelles**, et êtes donc tenu·e de **respecter la réglementation** en vigueur (*RGPD*).

 Bonnes pratiques à suivre

- Si votre recherche implique des **données à caractère personnel**, le **traitement doit être inscrit** dans le **registre des activités de traitement** de votre université ou de la structure désignée (notamment pour les **UMR**).
- En cas de doute, **contactez le Délégué à la Protection des Données (DPO)** de votre établissement.
- Assurez-vous d'appliquer les principes fondamentaux du **RGPD** : minimisation, transparence, consentement, sécurité et durée de conservation adaptée.