

Se familiariser avec les plans de gestion de données de la recherche en 6 points

1. Pourquoi gérer les données de la recherche ?
2. Définir un plan de gestion de données (PGD), ou *Data Management Plan (DMP)*
3. Exemple de trame d'un plan de gestion de données
4. Le modèle français/anglais du plan de gestion de données FAIR d'Horizon 2020
5. Un autre modèle en français de plan de gestion de données
6. Des modèles de plans de gestion de données en anglais

1. Pourquoi gérer les données de la recherche ?

Gérer des données de recherche fait partie du projet de recherche. La gestion des données de recherche (ou *Research data management, RDM*) requiert une organisation, une planification et un suivi rigoureux tout au long du projet et au-delà pour assurer la pérennité, l'accessibilité et la réutilisation de ces données.

La gestion des données de recherche répond aux objectifs suivants :

- elle accroît l'efficacité de la recherche, en facilitant l'accès et l'analyse des données par le chercheur qui a conduit la recherche ou par tout autre chercheur ;
- elle assure la continuité de la recherche par la réutilisation des données, tout en évitant la duplication des efforts ;
- elle favorise la diffusion élargie et accroît l'impact ; des données de recherche correctement formatées, décrites et identifiées gardent une valeur à long terme ;
- elle permet d'assurer l'intégrité de la recherche et la validation des résultats. Des données de recherche exactes et complètes permettent de retracer les processus qui ont conduit aux résultats ;
- elle réduit le risque de perte et renforce la sécurité des données par l'utilisation de dispositifs de stockage robustes et adaptés ;
- elle accompagne l'évolution actuelle de la publication, où les données étayant les résultats présentés par exemple dans un article scientifique doivent être documentées et rendues accessibles via un entrepôt de données ;
- elle satisfait aux conditions de financement du projet par les bailleurs, qui demandent de plus en plus l'ouverture de ces données pour qu'elles soient accessibles librement et gratuitement ;
- elle atteste de votre responsabilité ; en gérant vos données de recherche et en les rendant disponibles, vous faites preuve d'une utilisation responsable du financement public de la recherche.

2. Définir un plan de gestion de données (PGD), ou *Data Management Plan* (DMP)

La gestion des données de recherche est facilitée par la définition d'un plan de gestion de données (ou *Data Management Plan*, DMP). Un plan de gestion de données est un document formel explicitant la façon dont vous obtenez, documentez, analysez et utilisez vos données à la fois au cours de votre recherche et une fois le projet terminé. Il décrit dans le détail les méthodes et processus de création, de fourniture, de maintenance, de conservation et de protection des données.

3. Exemple de trame d'un plan de gestion de données

La trame de plan présentée ci-dessous est une adaptation de : **Checklist for a Data Management Plan**. V.4.0. Edinburgh, UK: Digital Curation Centre (DCC), 2013. Téléchargeable au format PDF à <http://www.dcc.ac.uk/resources/data-management-plans>, cette trame correspond au formulaire en ligne développé par **Digital Curation Centre's DMPonline tool** (<https://dmponline.dcc.ac.uk/>).

Informations administratives

- nom et identifiant du projet
- description du projet
- agence(s) de financement
- nom et identifiant éventuel du responsable principal de projet
- contact pour les données de projet
- date de la 1^{re} version
- date de la dernière mise à jour
- politiques associées au projet, incluant les instructions ou recommandations de l'agence de financement et de l'institution

Collection de données

- description des données, incluant le type de données, le format et le volume
- jeux de données existants qui seront utilisés
- méthodes de collecte et de création des données
- système d'organisation, de nommage et de gestion des répertoires et des fichiers
- processus d'assurance qualité mis en œuvre

Documentation et métadonnées

- informations nécessaires pour lire et interpréter les données
- organisation de la collecte et de la documentation
- standards de métadonnées adoptés

Ethique, cadre légal

Ethique

- détails de l'accord de conservation et de partage des données
- étapes pour la protection de l'identité des participants

- étapes pour la sécurité du stockage et du transfert de données sensibles

Cadre légal : droits de propriété intellectuelle et copyright

- nom de(s) propriétaire(s) des données
- licence(s) pour la réutilisation des données (par exemple, une des licences *Creative Commons* ou *Open Data Commons*)
- restrictions d'utilisation par les tierces parties
- délai requis pour le partage de données (embargo lié à la publication dans une revue ou délai d'application d'un brevet)

Stockage, sauvegarde, sécurité

Stockage et sauvegarde

- lieu de stockage des données
- plan de sauvegarde
- personne ou équipe responsable de la sauvegarde
- procédures de récupération

Sécurité

- risques et leur gestion
- dispositif d'accès
- dispositif éventuel pour le transfert sûr et intègre des données collectées sur le terrain

Sélection et conservation

- informations détaillées sur les données qui seront retenues, partagées et/ou conservées, et référence aux obligations contractuelles, légales ou réglementaires
- utilisations de recherche prévues des données
- durée de conservation des données au-delà du projet
- entrepôt ou archive de conservation des données et responsabilités associées
- temps et effort nécessaire à la préparation des données pour leur conservation et leur partage

Partage des données

- étapes à mener pour faciliter la prise de connaissance (*discovery*) des données par les autres
- conditions de restriction du partage des données et détails de leur application dans l'accord de partage de données
- mécanisme de partage de données (via un entrepôt, sur demande expresse ou tout autre processus)
- délai de publication
- procédure éventuelle d'obtention d'un identifiant persistant pour les données

Responsabilités et moyens

- nom de la personne responsable de la mise en œuvre du plan de gestion de données
- nom de la personne responsable de chaque activité de gestion des données
- équipements et logiciels requis (en addition à ceux existants fournis par l'institution)
- besoins additionnels d'expertise ou de formation

- charges imposées par les entrepôts de données

4. Le modèle français/anglais du plan de gestion de données FAIR d'Horizon 2020

(version française) **Lignes directrices pour la gestion des données FAIR dans Horizon 2020. Version 3.0.** Commission européenne (traduction française : INIST-CNRS, France), 26 juillet 2016, 13 p.

(version anglaise) **Guidelines on FAIR Data Management in Horizon 2020: Version 3.0.** European Commission, 26 July 2016, 12 p.

<http://www.horizon2020.gouv.fr/cid82025/le-libre-acces-aux-publications-aux-donnees-recherche.html>

Ces lignes directrices aident les bénéficiaires du programme Horizon 2020 à rendre leurs données de recherche trouvables, accessibles, interopérables et réutilisables (ou **FAIR** : *Findable, Accessible, Interoperable and Reusable*). Elles visent à promouvoir auprès des porteurs de projets une saine gestion des données comme part essentielle des bonnes pratiques de la recherche.

Elles sont accompagnées d'un modèle de plan de gestion de données s'inspirant du concept FAIR et composé d'une série de questions : Quel est l'objet de la collecte/génération des données et son rapport avec les objectifs du projet ? Quels types et formats de données le projet va collecter/générer ? Allez-vous réutiliser des données existantes et de quelle manière ? Quelle est l'origine de ces données ? Quel est le volume prévu des données ? A qui pourraient-elles être utiles ?

5. Un autre modèle en français de plan de gestion de données

Réaliser un plan de gestion de données « FAIR » : guide de rédaction. V2. Université Paris Diderot, Université Sorbonne Paris Cité, Université Paris Descartes. 2018, 33 p.

https://archivesic.ccsd.cnrs.fr/sic_01690547v2

Ce guide de plan de gestion de données tient compte du cycle de vie des données, de leur création dans le cadre d'un projet de recherche jusqu'à leur archivage. Il intègre les attendus de la Commission européenne dans le cadre d'Horizon 2020 et du FAIR data management. Il s'appuie également sur divers modèles tels que celui de la [National Science Foundation \(NSF, Etats-Unis\)](#) ou de [l'Interuniversity Consortium for Political and Social Research \(ICPSR, Université du Michigan\)](#).

La trame du PGD proposé est également accessible via l'interface web **DMP OPIDoR** mise à disposition des chercheurs de l'Enseignement supérieur et de la recherche française. Elle comprend 3 phases à renseigner en fonction de l'avancement du projet : 1) Contexte, acteurs et ressources ; 2) Stockage, partage, protection et diffusion au cours du projet ; 3) Dissémination et archivage à l'issue du projet.

6. Des modèles de plans de gestion de données en anglais

Digital Curation Centre's DMPonline tool - <https://dmponline.dcc.ac.uk/>

DMPonline (2018) est une interface web développée par Digital Curation Centre (DCC, Edinburgh Royaume-Uni) pour la rédaction en ligne de plans de gestion de données.

Après enregistrement gratuit et connexion, l'utilisateur répond via un formulaire en ligne à une série de questions correspondant à la trame du plan de gestion de données fourni par DCC. L'ensemble des réponses saisies en mode texte peut être exporté en format web ou en texte.

La trame de plan *Checklist for a Data Management Plan*, qui correspond au formulaire en ligne, est téléchargeable au format PDF à <http://www.dcc.ac.uk/resources/data-management-plans/checklist>. Elle est découpée en 8 sections : 1) Informations administratives ; 2) Collection de données ; 3) Documentation et métadonnées ; 4) Ethique et cadre légal ; 5) Stockage, sauvegarde, et sécurité ; 6) Sélection et conservation ; 7) Partage des données ; 8) Responsabilités et moyens.

University of California's DMPTool - <https://dmptool.org/>

DMPTool (2018) est une interface web développée par l'University of California Curation Center (Etats-Unis) pour rédiger en ligne des plans de gestion de données selon différents modèles.

Après enregistrement gratuit et connexion, l'utilisateur sélectionne un modèle (*Template*) parmi plus de 30 modèles basés sur les exigences d'agences de financement américaines ou britanniques comme [National Institutes of Health \(NIH\)](#), [National Science Foundation \(NSF\)](#), [U.S. Department of Agriculture \(USDA\)](#), [Digital Curation Centre \(DCC\)](#).

Après la saisie guidée par des questions, le plan peut être exporté aux formats HTML, PDF et TXT. Il est organisé en 5 sections : 1) Types de données produites ; 2) Données et standards de métadonnées ; 3) Politiques d'accès et de partage des données ; 4) Politiques pour la réutilisation et la redistribution des données ; 5) Plans d'archivage et de conservation.

ESRC Data management plan: guidance for peer reviewers - <https://esrc.ukri.org/funding/guidance-for-peer-reviewers/>

Un guide (2013, 4 p., format PDF) est fourni par l'Economic and Social Research Council (ESRC, Royaume-Uni) pour les relecteurs et rapporteurs de projets impliqués en début et fin de cycle de vie d'un projet.

Des conseils aident à évaluer la qualité d'un plan de gestion de données et son adéquation avec le projet en recherche économique ou sociale auquel il est associé. Une attention est portée aux données sensibles, à la confidentialité, à l'éthique, à la responsabilité, à la sécurité et à la propriété intellectuelle.

NERC Data Management planning - <https://nerc.ukri.org/research/sites/data/dmp/>

Un plan de gestion de données (*Full Data Management Plan*, 16 August 2012, 2 p., format .doc) est proposé par le [Natural Environment Research Council \(NERC\)](#), Royaume-Uni) aux porteurs de projet pour les jeux de données (*datasets*) de valeur à long terme destinés à être rendus disponibles dans des centres de données NERC pour leur archivage et leur réutilisation.

Le plan comporte 9 sections : 1) Information sur le projet ; 2) Centre de données désigné (*Data Center*) ; 3) Rôles et responsabilités ; 4) Activités liées à la production de données ; 5) Gestion du projet ; 6) Description des jeux de données ; 7) Qualité des données ; 8) Besoins de services particuliers liés à la complexité ou au volume des données ; 9) Information sur la plan de gestion des données.

Marie-Claude Deboin

Délégation à l'information scientifique et technique, Cirad

17 juillet 2014, mise à jour 3 mai 2018

Comment citer ce document :

Deboin, M.C. 2018. *Se familiariser avec les plans de gestion de données de la recherche*, en 6 points. Montpellier (FRA) : CIRAD, 5 p.
<https://coop-ist.cirad.fr/donnees-recherche>

*Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons : Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International, disponible en ligne : <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.fr> ou par courrier postal à : Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.
Cette licence vous permet de remixer, arranger, et adapter cette œuvre à des fins non commerciales tant que vous créditez l'auteur en citant son nom et que les nouvelles œuvres sont diffusées selon les mêmes conditions*