



Faciliter Open Science Training for European Research

Etablir un plan de gestion des données dans le cadre d'un projet européen



Aurore Cartier
Magalie Moysan
Nathalie Reymonet

Plan de l'intervention

1. Les **enjeux** d'une bonne gestion : le **cycle de vie** de la données
2. Les **outils** d'une bonne gestion : le **plan de gestion** de données
3. **Rédiger** un plan de gestion de données
 - *Structures et attendus*
 - *Outils et guides en ligne*
 - *Compétences et acteurs*



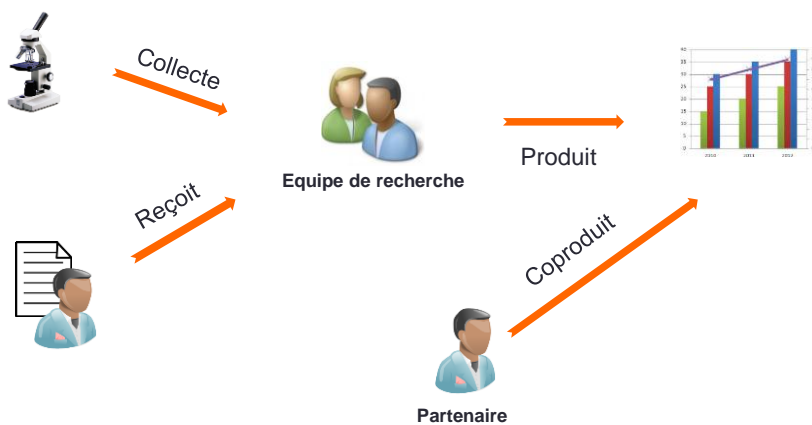
Les enjeux d'une bonne gestion

Le cycle de vie de la donnée



Qu'est ce qu'une donnée de recherche ?

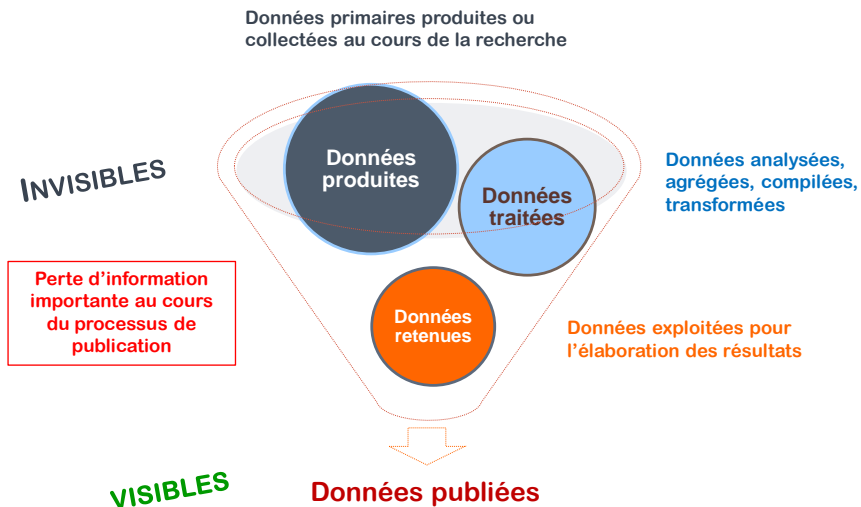
La donnée est une information factuelle enregistrée sur un **support**, **produite** ou **collectée**, selon divers **procédés** au cours d'un **processus**.



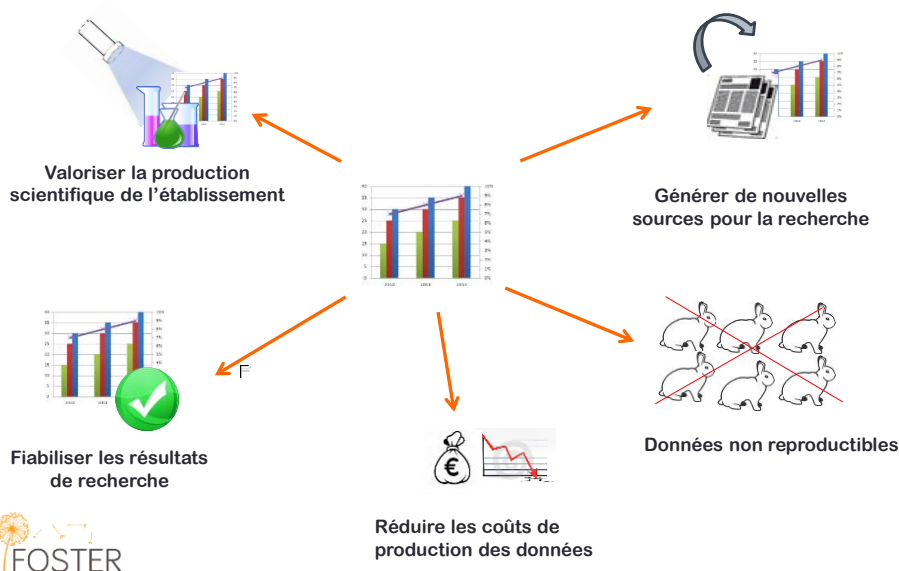
Types de données	Définitions	Valeur et spécificité	Exemples
Données d'observation (Observational datas)	Données obtenues en temps réel	Souvent uniques et irremplaçables	Données atmosphériques, d'enquêtes, échantillons, neuro-image
Données expérimentales (Experimental datas)	Données obtenues en laboratoire à partir d'équipements spécifiques	Reproductibles mais à des coûts dissuasifs	Séquence de génome, chromatographie, spectres RMN
Données de simulation (Simulation datas)	Données générées à partir de modèles test	Métadonnées et modèles ont une valeur supérieure aux résultats	Modèles climatiques, modèles économiques
Données dérivées ou compilées (Derived or compiled data)	Données obtenues par compilations ou traitement des données brutes	Reproductibles mais à des coûts dissuasifs	Texte et data mining, bases de données compilées, modèles 3D
Données de référence ou données canoniques (Reference or canonical datas)	Collections statiques ou organiques de jeux de données validées	Données publiées ou qui ont fait l'objet d'une curation	Banque de données sur le génome, structure chimiques, portail de données spatiale



Que deviennent les données aujourd'hui ?



Les enjeux de la diffusion et de la conservation des données



Quelles attentes pour les données dans H2020 ?



Détectables / citables

- Identifiant pérenne (DOI, Handle)
- Indexées

Fiables et intelligibles

- Validées
- Décrites, documentées
- Normalisées

Interopérables

- Entrepôts ouverts
- Protocoles moissonnables

Accessibles

- En ligne sur le web
- Gratuites

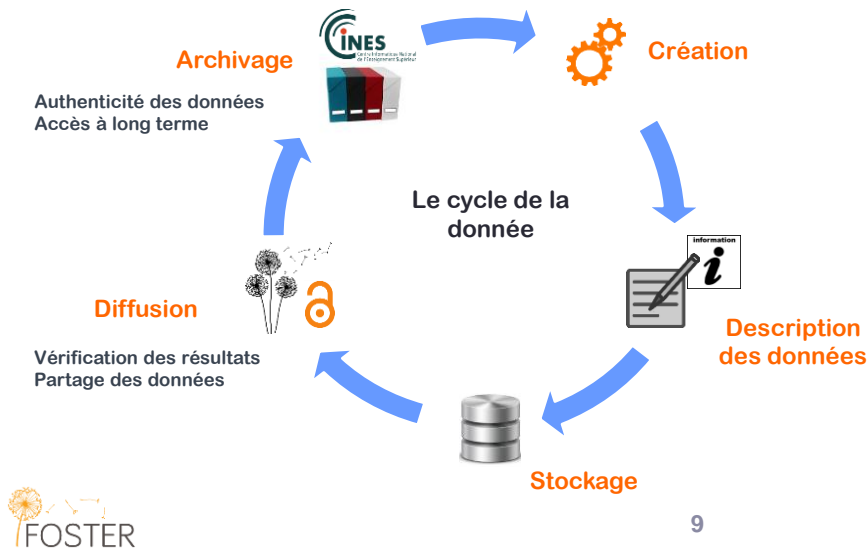
Métadonnées

Réutilisables

- Licence libres
- Formats non propriétaires
- Archivage pérenne







Documenter les données pour les rendre réutilisables







Les initiatives internationales pour l'ouverture des données

- **Politiques d'ouverture des données**

-  Université d'Harvard (Etats-Unis)
-  Université d'Edimbourg (Royaume-Uni)
-  Université de Göttingen (Allemagne)
-  Université de Leiden (Pays-Bas)

- **Data management plans**

-  U.S. National Science Foundation (Etats-Unis)
-  U.S. Department of Energy (Etats-Unis)
-  U.K. Research Councils (Royaume-Uni)
-  The Netherlands Organisation for Scientific Research (Pays-Bas)

Les **outils** d'une bonne gestion

Le plan de gestion de données dans Horizon 2020



Les attendus de l'Open research data pilot



- **produire un plan de gestion des données**
 - *Draft* à remettre dans les 6 mois suivant l'acceptation du projet
 - anticiper la dissémination des données générées au cours de la recherche
 - Prévoir des mises à jour au minimum à mi-parcours et au rapport final
- **déposer les données dans un entrepôt de données**
 - Concerne les données et métadonnées nécessaires à la validation des résultats présentés dans les publications
 - les autres données et métadonnées mentionnées dans le plan de gestion de données
- **documenter les résultats**
 - Fournir les informations et outils nécessaires à la validation et à la réutilisation des résultats et jeux de données déposées
 - Préciser les conditions et les perspectives d'utilisation et de réutilisation des données déposées



Le plan de gestion de données (DMP) dans Horizon 2020

Définition

*A data management plan **describes the data management life cycle for all data sets that will be collected, processed or generated** by the research project. It is a document outlining **how research data will be handled** during the research project, and even after the project is completed, **describing what data will be collected, processed or generated and following what methodology and standards, whether and how data will be shared and/or made open, and how it will be curated and preserved.***



[Guidelines on Data Management in Horizon 2020](#)

Le plan de gestion de données (PGD) ou « data management plan » (DMP)



Un plan de gestion de données est un document formel précisant la manière dont seront produites, traitées, décrites, diffuser et conserver les données au cours et à l'issue du projet.

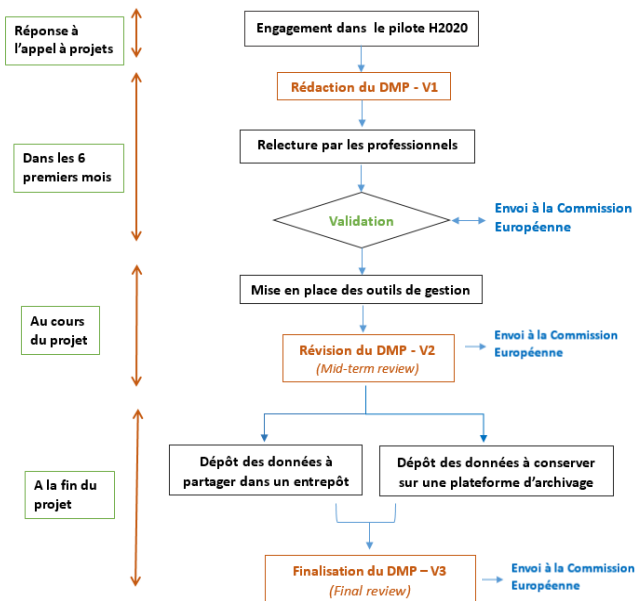
Il permet de :

- décrire le cycle de gestion et de vie des données produites, collectées ou générées tout au long du projet de recherche.
- anticiper les questions de gestion qui surviennent au cours d'une recherche et les conditions d'une conservation et diffusion future des données.



Les étapes de rédaction d'un DMP

Le DMP formalise au sein d'un document **unique, prospectif et évolutif**, des informations utiles au suivi du projet et à la bonne gestion des résultats obtenus, auparavant dispersées entre divers acteurs ou documents.



Le DMP dans le montage des réponses aux appels à projet

Project's proposal submission and evaluation

- ✓ Where relevant, applicants must provide a short, general outline of their policy for data management [...]
- ✓ The described policy should reflect the current state of consortium agreements regarding data management and be consistent with those referring to exploitation and protection of results.
- ✓ The data management section can be considered also as a checklist for the future and as a reference for the resource and budget allocations related to data management.



[Guidelines on Data Management in Horizon 2020](#)

Exemple de prévision de DMP

HORIZON 2020
PROGRAMME DE RECHERCHE ET INNOVATION
FINANÇÉ PAR L'UNION EUROPÉENNE



We wish to opt out of the Pilot on Open Research Data in Horizon 2020.

Yes

No

[†] According to article 43.2 of Regulation (EU) No 1209/2013 of the European Parliament and of the Council, of 11 December 2013, laying down the rules for participation and dissemination in "Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020)" and repealing Regulation (EC) No 1906/2006.

Data management activities

The use of a [Data Management Plan \(DMP\)](#) is required for projects participating in the [Open Research Data Pilot in Horizon 2020](#), in the form of a deliverable in the first 6 months of the project.

All other projects may deliver a DMP on a voluntary basis, if relevant for their research.

Are data management activities relevant for your proposed project?

Yes

No

A Data Management Plan will be delivered
(Please note: Projects participating in the Open Research Data Pilot **must** include a Data Management Plan as a deliverable in the first 6 months of the project).

Data Management is part of a Work Package.



Rédiger un DMP

Structures et attendus



Les principaux champs d'un DMP

- **Section 1** : Informations relatives au projet
- **Section 2** : Responsabilité des données : répartition des rôles
- **Section 3** : Ressources nécessaires à la mise en œuvre du PGD
- **Section 4** : Jeux de données
 - 4.1 Description du jeu de données
 - 4.2 Stockage, accès et sécurité des données – *au cours du projet*
 - 4.3 Métadonnées : documentation et organisation des données
 - 4.4 Dissémination des jeux de données – *à l'issue du projet*
- **Section 5** : Sélection et archivage – obligatoire que les données soient diffusées ou non

Valable pour l'ensemble du projet

Propre à chaque jeu de données

Valable pour l'ensemble du projet

Un jeu de données est un ensemble de données techniquement homogènes ou intellectuellement cohérentes.



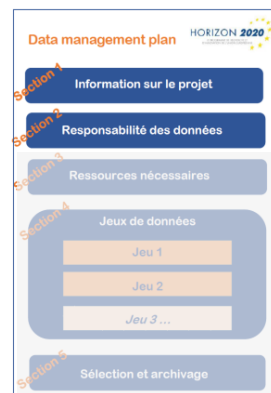
A l'échelle du projet

Section 1 : les informations relatives au projet

- **Objectifs** : informer sur le contexte administratif et scientifique du projet de recherche auquel est lié le PGD
- **Principaux champs** :
 - Identifiant de l'appel à projets
 - Thématiques,
 - Objectifs du projet

Section 2 : responsabilité des données

- **Objectifs** : identifier la ou les personne(s) qui seront en charge l'application et de la mise à jour du DMP tout au long du projet
- **Principaux champs** :
 - Responsable des données
 - Propriété des données (*renvoi vers Accord de consortium*)



A l'échelle du projet

Section 3 : ressources nécessaires

Objectifs : Estimez les compétences, ressources et coûts nécessaires à la mise en œuvre du DMP : gestion, curation, formation et archivage

Principaux champs :

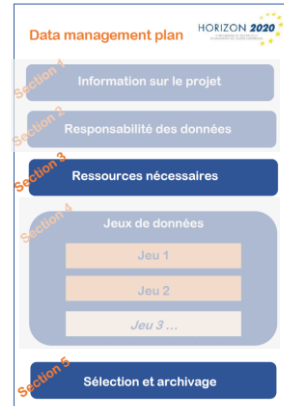
- Matériel
- Personnel
- Formation
- Coûts

Section 5 : Sélection et archivage des données

▪ **Objectifs** : sélectionner et prévoir l'archivage à long terme des données ayant vocations à être conservées

▪ **Principaux champs** :

- Sort
- Sélection
 - Volume
 - Durée de conservation préconisée



A l'échelle du jeu de données (dataset)

Section 4.1 : Description des données

▪ **Objectifs** : présenter le type de données du jeu qui seront produites et reçues dans le cadre du projet

▪ **Principaux champs** :

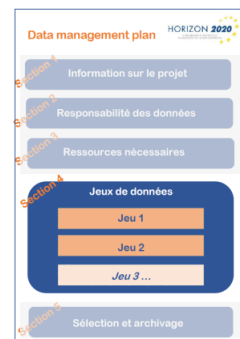
- Identifiant et nature du jeu de données
- Réutilisation de données existantes
- Méthode(s) de production
- Format des données

Section 4.2 : Stockage, accès et sécurité des données (au cours du projet)

▪ **Objectifs** : préciser modalités d'hébergement, de sauvegarde et d'accès aux données pendant la phase active du projet

▪ **Principaux champs** :

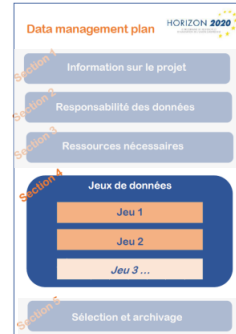
- Support et volumétrie prévisionnelle
- Type d'hébergement
- Garantie de confidentialité, intégrité, traçabilité et disponibilité des données
 - Gestion des accès
 - Condition d'échange et de partage



A l'échelle du jeu de données (*dataset*)

Section 4.3 : Métadonnées

- **Objectifs** : préciser la manière dont seront décrites et organisées les données produites ou reçues au cours du projet
- **Principaux champs** :
 - Standards et formats des métadonnées
 - Modes de production et responsabilités
 - Arborescence et règle de nommage
 - Documentation associée



Section 4.4 : Dissémination (à l'issue du projet)

- **Objectifs** : préciser les modalités et les éventuelles précautions éthiques, juridiques et techniques selon lesquelles les données seront diffusées
- **Principaux champs** :
 - Principe général de diffusion
 - Licence et potentiel de réutilisation
 - Choix de dépôt et de dissémination des données
 - Identification des jeux de données sensibles
 - Dérogation aux principes de diffusion et mesures de protection
 - Période d'embargo



Dépôt et partage des données : politiques et infrastructures

Trouver le bon entrepôt: **Databib**

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES



Les exceptions à la diffusion des données



Le chercheur peut justifier de la non-diffusion globale (*opt-out*) ou partielle de ses données pour diverses raisons :

- Si le projet ne collecte ou ne génère aucune donnée
- En cas d'incompatibilité avec l'exploitation industrielle et commerciale
- En cas d'incompatibilité avec des questions de sécurité
- Pour protéger des données personnelles
- Si la diffusion des données risque de compromettre l'objectif du projet
- Pour toute autre raison légitime faisant l'objet d'une explication



Ces exceptions ne dispensent pas de l'élaboration d'un plan de gestion des données...

... afin de rester cohérent avec son projet de recherche !



Exemple d'exception de diffusion



We wish to opt out of the Pilot on Open Research Data in Horizon 2020. Yes No

Please indicate the reason(s) for not being able to participate in the Pilot:

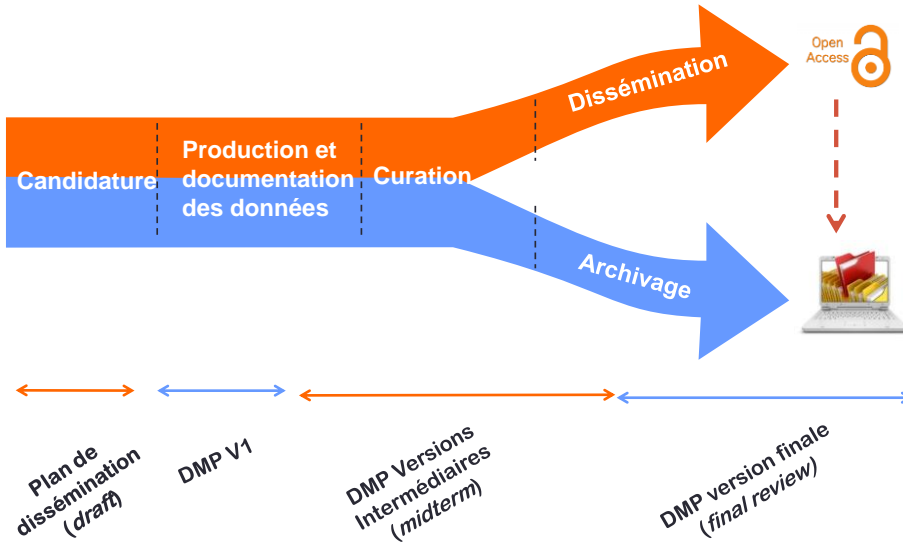
- | | |
|--|-------------------------------------|
| - the project does not generate any data | <input type="checkbox"/> |
| - to allow the protection of results (e.g. patenting) | <input type="checkbox"/> |
| - incompatibility with the need for confidentiality linked to security reasons | <input type="checkbox"/> |
| - incompatibility with privacy/data protection reasons | <input type="checkbox"/> |
| - achievement of the project's main aim would be jeopardised | <input type="checkbox"/> |
| - other legitimate reasons | <input checked="" type="checkbox"/> |

Please specify the reason:

Lack of time



Dissémination et conservation dans les projets H2020



27

Rédiger un DMP

Outils et guides en ligne





Le DMP Online du Digital Curation Center

Plan details | Initial DMP | Mid-term Review DMP | Final review DMP | Share | Export

This page gives you an overview of your plan. It tells what your plan is based on and gives an overview of the questions that you will be asked. [Edit plan details](#)

Plan name: My plan (Horizon 2020 DMP)
 ID: DMP_Foster_H2020_2015
 Grant number: 612425

Principal Investigator: [Plan details](#) | [Initial DMP](#) | [Mid-term Review DMP](#) | [Final review DMP](#) | [Share](#) | [Export](#)

This plan is based on:

You can give other people access to your plan here. There are three permission levels.
 Users with "read only" access can only read the plan.
 Editors can contribute to the plan.
 Co-owners can also contribute to the plan, but additionally can edit the plan details and control access to the plan.
 Add each collaborator in turn by entering their email address below, choosing a permission level and clicking "Add collaborator".
 Those you invite will receive an email notification that they have access to this plan, inviting them to register with DMPonline if they don't already have an account. A notification is also issued when a user's permission level is changed.

Collaborators

Email address	Permissions	
Aurore Cartier	Owner	
magalie.moysan@univ-paris-diderot.fr	Co-owner	Remove user access
nathalie.reymonet@univ-paris-diderot.fr	Co-owner	Remove user access



Plan details | **Initial DMP** | Mid-term Review DMP | Final review DMP | Share | Export

For each data set specify the following: (5 questions, 0 answered)

The DMP should address the points below on a dataset by dataset basis and should reflect the current status of reflection within the consortium about the data that will be produced

Data set reference and name: [Guidance](#) [Add comment](#)
 EC Guidance:

[Save](#)

Not answered yet

Data set description: [Guidance](#) [Add comment](#)
 EC Guidance:

Plan details | **Initial DMP** | **Mid-term Review DMP** | Final review DMP

Scientific research data should be easily: 1. Discoverable (1 question, 0 answered)

2. Accessible (1 question, 0 answered)

3. Assessable and intelligible (1 question, 0 answered)

4. Usable beyond the original purpose for which it was collected (1 question, 0 answered)

5. Interoperable to specific quality standards (1 question, 0 answered)

Initial DMP

Mid-term Review DMP

Final review DMP

Format: pdf [Export](#)





Le DMP Tool de la California Digital Library de l'Université de Californie

Use the search box or the A-Z links below to narrow down the list by plan title. Public DMPs are plans created using the DMPTool and shared publicly by their owners. They are not vetted for quality, completeness, or adherence to funder guidelines.

A - F G - L M - S T - Z All

Plan Title	Funder Template	Institution	Owner	Download
Inquiring into Engineering Faculty Assumptions about Students: A Profound Human Systems Intervention	NSF-ENG: Engineering	California Polytechnic State University-San Luis Obispo	Lizabeth Schiemer	
Late Season Productivity, Carbon, and Nitrogen Dynamics in a Changing Arctic	NSF-EAR: Earth Sciences	Oregon State University	Miguel Goni	
Language Endangerment and Revitalization at the Diachronal Level: A current analysis of Salasaca Kichwa (ISO 639-3 = qxt)	NSF-SBE: Social, Behavioral, Economic Sciences	Temple University	Robin Aronow	

[View All](#)



DMPTool is a service of the University of California Curation Center of the California Digital Library
Copyright © 2010-2015 The Regents of the University of California

[Privacy Policy](#) | [Accessibility Policy](#) | [Terms of Use](#) | [Contact Us](#) | [About](#)



Guide d'accompagnement des Universités : Paris Descartes et Paris Diderot



UNIVERSITÉ
PARIS DESCARTES
Service commun de la documentation

UNIVERSITÉ
PARIS
DIDEROT
Service commun de la documentation

Réaliser un plan de gestion de données

Ce document a été conçu afin d'accompagner les chercheurs et chargés de projets lors de la rédaction de plans de gestion de données (Data Management Plans, DMP). Sa structure s'appuie sur le modèle proposé par la Commission européenne dans le cadre d'Horizon 2020 et divers modèles de plans de gestion de données existants tels que celui de la National Science Foundation (NSF) ou de l'InterUniversity Consortium for Political and Social Research (ICPSR). Les champs requis par la Commission européenne sont signalés par un astérisque. Les exemples mentionnés dans ce document sont tous de guides existants.

Ce document constitue un guide de rédaction et non une liste de champs obligatoires.

10/19 janvier 2013

Ce document contient :

- Un rappel synthétique des exigences H2020 en matière de gestion et d'ouverture des données
- Une présentation des acteurs et échéances intervenant dans la rédaction du DMP
- Un modèle de DMP sous forme de grille avec exemples à l'appui

URL :

http://www.univ-paris-diderot.fr/DocumentsFCK/recherche/Realiser_un_DMP_V1.pdf



Tutoriels de l' INIST-CNRS



L'institut de l'information scientifique et technique (INIST) du CNRS propose des tutoriels sur le libre accès aux résultats de recherche dans Horizon 2020.

- Tutoriel 1 : principes généraux
- Tutoriel 2 : DMP – Présentation générale
- Tutoriel 3 : DMP – Description des données
- Tutoriel 4 : Standards et métadonnées
- Tutoriel 5 : DMP – Partage des données
- Tutoriel 6 : DMP – Archivage des données
- Tutoriel 7 : Déposer ses données dans un entrepôt

URL : <http://www.inist.fr/?-Tutoriels-multimedias-H2020->



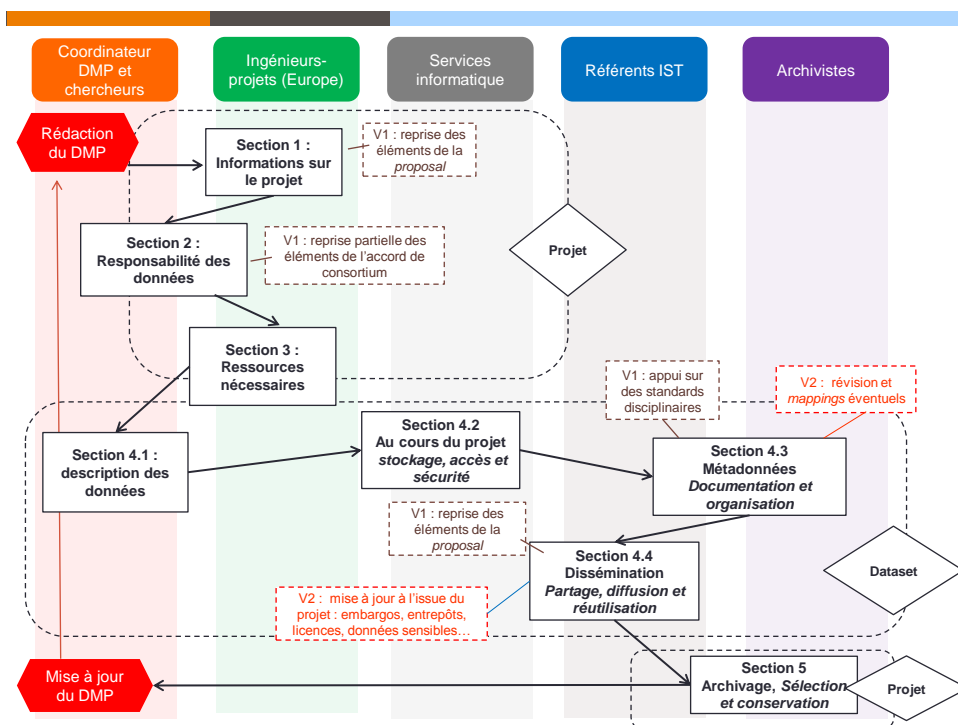
Rédiger un DMP

Compétences et acteurs



Compétences et savoir-faire nécessaires pour la rédaction d'un DMP

- ✓ Connaître les enjeux du droit et de la propriété intellectuelle en matière de production et de diffusion des données
- ✓ Repérer les principaux standards et formats de métadonnées disciplinaires pour la description des données
- ✓ Identifier les besoins en matière de gestion, stockage et sécurité des données et les coûts associés
- ✓ Pouvoir anticiper les coûts éligibles pour la mise en *open access* des publications et des données et leur conservation
- ✓ Connaître les principaux entrepôts d'archives ouvertes de publications et de données et leur fonctionnement
- ✓ Anticiper les coûts et besoins de la conservation pérenne des données, ses coûts et acteurs



Merci pour votre attention

Aurore Cartier – aurore.cartier@parisdescartes.fr

Magalie Moysan – magalie.moysan@univ-paris-diderot.fr

Nathalie Reymonet - nathalie.reymonet@univ-paris-diderot.fr



Références bibliographiques

- ANRT, CURIF. Séminaire conjoint « HORIZON 2020 : les enjeux et les conditions du succès » (25 juin 2014)
- European Commission . *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* (déc. 2013)
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- European Commission. *Guidelines on Data Management in Horizon 2020* (déc. 2013)
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- Open Access in Horizon 2020 - EC funded projects. *Briefing paper for Project Coordinators* .
<https://www.openaire.eu/guide-for-project-coordinators-in-horizon-2020/document-details>
- Commission européenne. *Recommandation de la commission du 17.7.2012 relative à l'accès aux informations scientifiques et à leur conservation* (17.7.2012) C(2012) 4890 final
http://medoanet.sciencesconf.org/conference/medoanet/pages/recommendation_access_and_preservation_scientific_information_fr_copie.pdf
- OCDE. *Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics* (2007) <http://www.oecd.org/fr/science/sci-tech/38500823.pdf>
- Commission européenne. *La charte européenne du chercheur* (2005)
http://ec.europa.eu/euraxess/pdf/brochure_rights/eur_21620_en-fr.pdf

