



Créer un plan de gestion des données (DMP)

Joanna JANIK
CNRS/DIST

joanna.janik@cnrs-dir.fr



www.cnrs.fr
Direction de l'Information
Scientifique et Technique
3 rue Michel-Ange
75016 PARIS



Remerciements à :

Aurore Cartier, Magalie Moysan, Nathalie Reymonet (université Paris Diderot)

Marie Puren (INRIA)

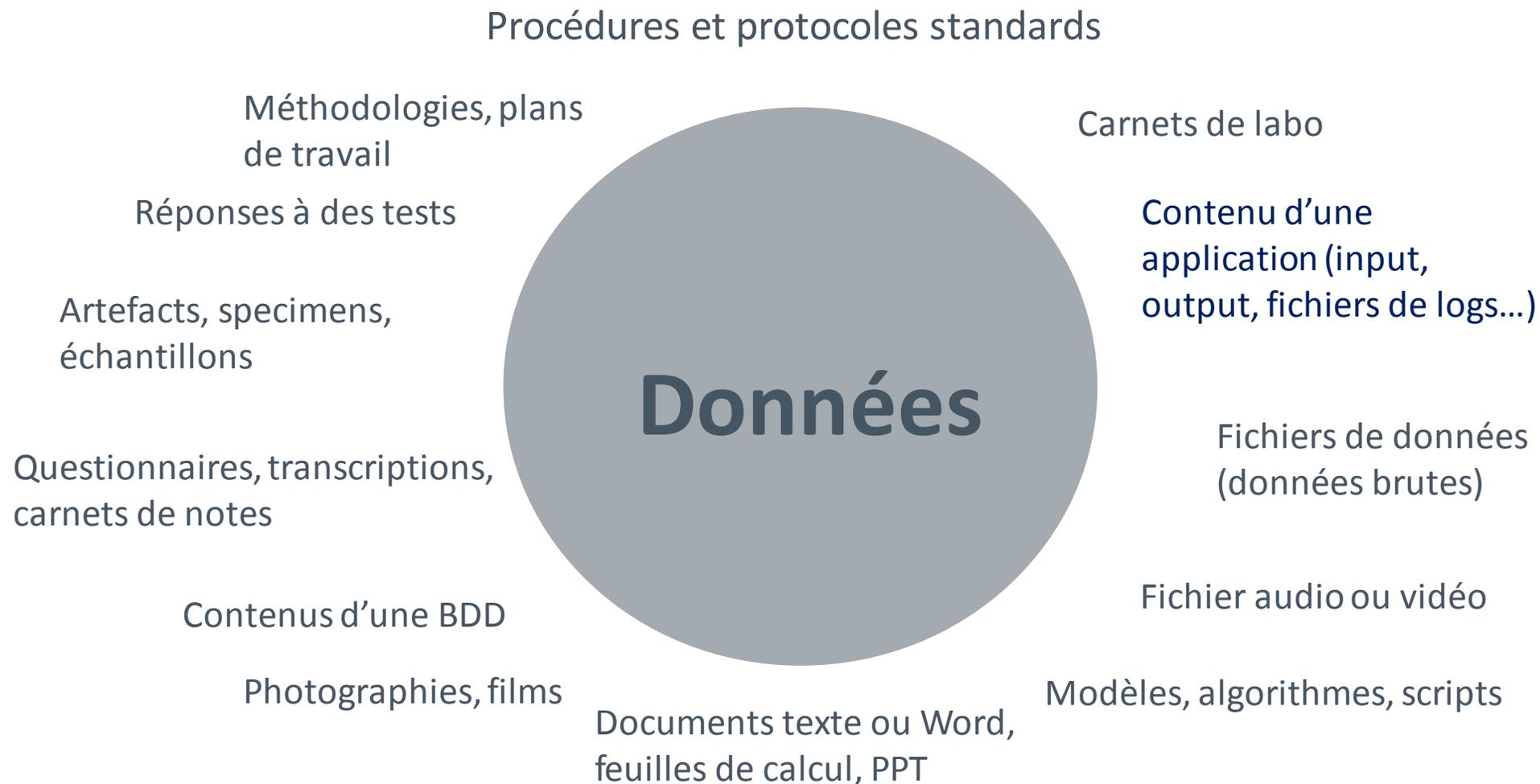
Emmanuelle Morlock (Cnrs lab. HiSoMA), Geneviève Romier (CNRS Institut de grilles et du Cloud)

Philippe Saby (université de Toulouse)





Les données de la recherche





Les données de la recherche

- ✓ des données qui aident à faire de la recherche
- ✓ des données qui peuvent être collectées, créées et analysées
- ✓ des données qui peuvent adopter des formats très différents



Les données de la recherche

Jeu de données ou dataset:

“agrégation, sous une forme lisible, de données brutes ou dérivées présentant une certaine ‘unité’, rassemblées pour former un ensemble cohérent”

« De l’Open data à l’Open research data : quelle(s) politique(s) pour les données de recherche ? »

Rémi Gaillard, mémoire Enssib 2014



Un nouveau modèle d'ouverture pour les données de la recherche

La **Science Ouverte** : « mouvement destiné à promouvoir plus de partage et de transparence » dans la science

- ✓ Promeut un nouveau modèle d'ouverture pour les données de la recherche.
- ✓ Principaux aspects:
 - La disponibilité et l'accès
 - La ré-utilisation et la redistribution
 - La participation de tous

Ouvert = «Quiconque peut librement avoir accès, utiliser, modifier et partager pour quelque motif que ce soit»

Open Knowledge International, ["The Open definition"](#)



Un nouveau modèle d'ouverture pour les données de la recherche



The Research Data Lifecycle

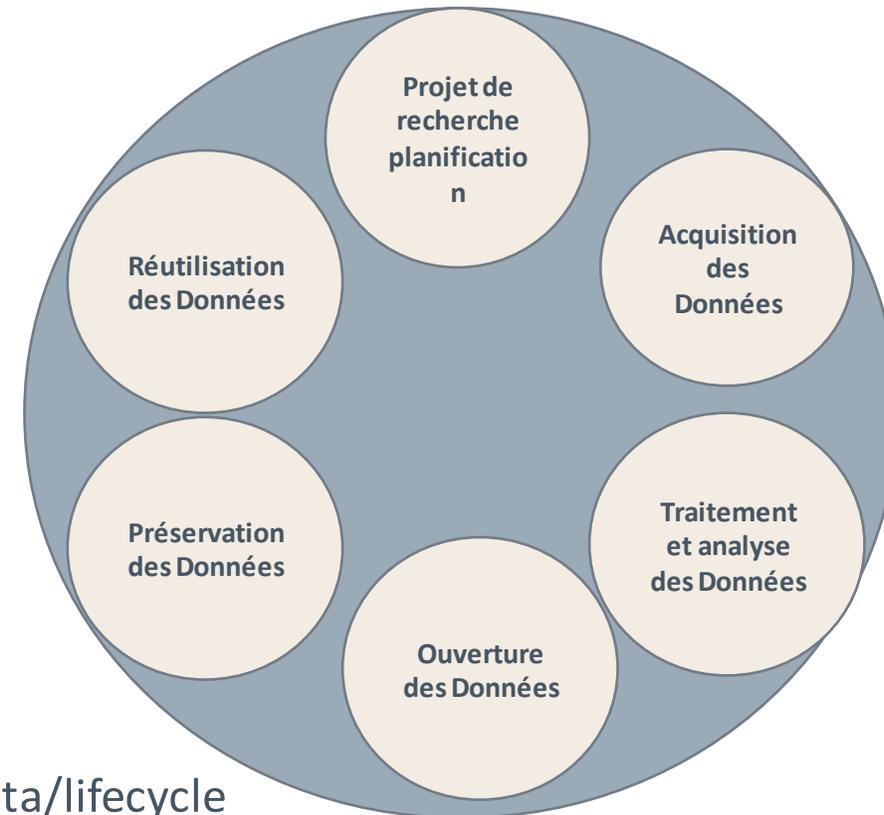


Schéma adapté de UK Data Service 2017

<https://www.ukdataservice.ac.uk/manage-data/lifecycle>





Partager des données, mais comment?

En gérant vos données !

- ✓ créer un **Plan de Gestion des Données**
 - avec la demande de financement
- ✓ **stocker** des données de recherche durant le projet
- ✓ **partager** les données avec les collègues autorisés
- ✓ **déposer** ces données dans un **entrepôt pérenne** (les rendre disponibles)



Soyons « FAIR »



FAIR Data

Findable : Métadonnées descriptives riches - Identifiants pérennes

Accessible : Utilisation de protocoles libres, ouverts et standardisés pour communiquer les données – les métadonnées toujours accessibles

Interoperable : formats ouverts, vocabulaires contrôlés & standards de métadonnées

Re-usable : Droits explicités et licences d'utilisation claires





Plan de gestion des données définition

« Data Management Plan » (DMP)

=

« Plan de gestion de données » (PGD)

« DMP describes the data management life cycle for all datasets to be collected, processed or generated by a research project. »

[Guidelines on Data Management in Horizon 2020, 26 juillet 2016](#)

Depuis juillet 2016: tous les projets H2020 doivent fournir un DMP





Une Bonne pratique de recherche

*« La vocation d'un DMP n'est pas de contraindre la pratique par l'exigence de nouveaux éléments mais plutôt de formaliser au sein d'un document **unique, prospectif, descriptif et évolutif**, un ensemble d'éléments et d'informations, auparavant dispersés entre divers acteurs et divers documents, éléments utiles au suivi du projet et à la bonne gestion des résultats obtenus. »*

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





Plan de gestion des données définition

- ✓ Le DMP est un **document formalisé** ;
- ✓ Son contenu peut varier selon les besoins des projets et les demandes des agences de financement ;
- ✓ C'est avant tout une **feuille de route**, en ajustement constant ;
- ✓ Le DMP est centré sur les données et les **jeux de données (datasets)** du projet de recherche.

Le DMP est un livrable du projet



Plan de gestion des données pourquoi ?

Le DMP sert à définir comment les données du projet seront :



Le DMP contient :

- ✓ une description des **données** (au sens large) : métadonnées, formats... ;
- ✓ une description de la **politique associée aux données** (accès, diffusion, propriété intellectuelle, confidentialité...)

Le DMP participe à la **sécurisation** et à la **pérennisation** des données.



Plan de gestion des données pourquoi ?

Une bonne pratique de recherche qui permet de :

- ✓ Formaliser la gestion de son travail
- ✓ Gérer son investissement/budgétiser
- ✓ Donner de la visibilité aux données en montrant qu'elles peuvent être réutilisées
- ✓ Ecrire des data papers
- ✓ Gagner du temps par la suite - un DMP peut améliorer la gestion des données





Plan de gestion des données acteurs/responsabilités

✓ Chercheur

- Responsable des données, coordinateur du DMP
- Description des données et découpage en différents jeux de données

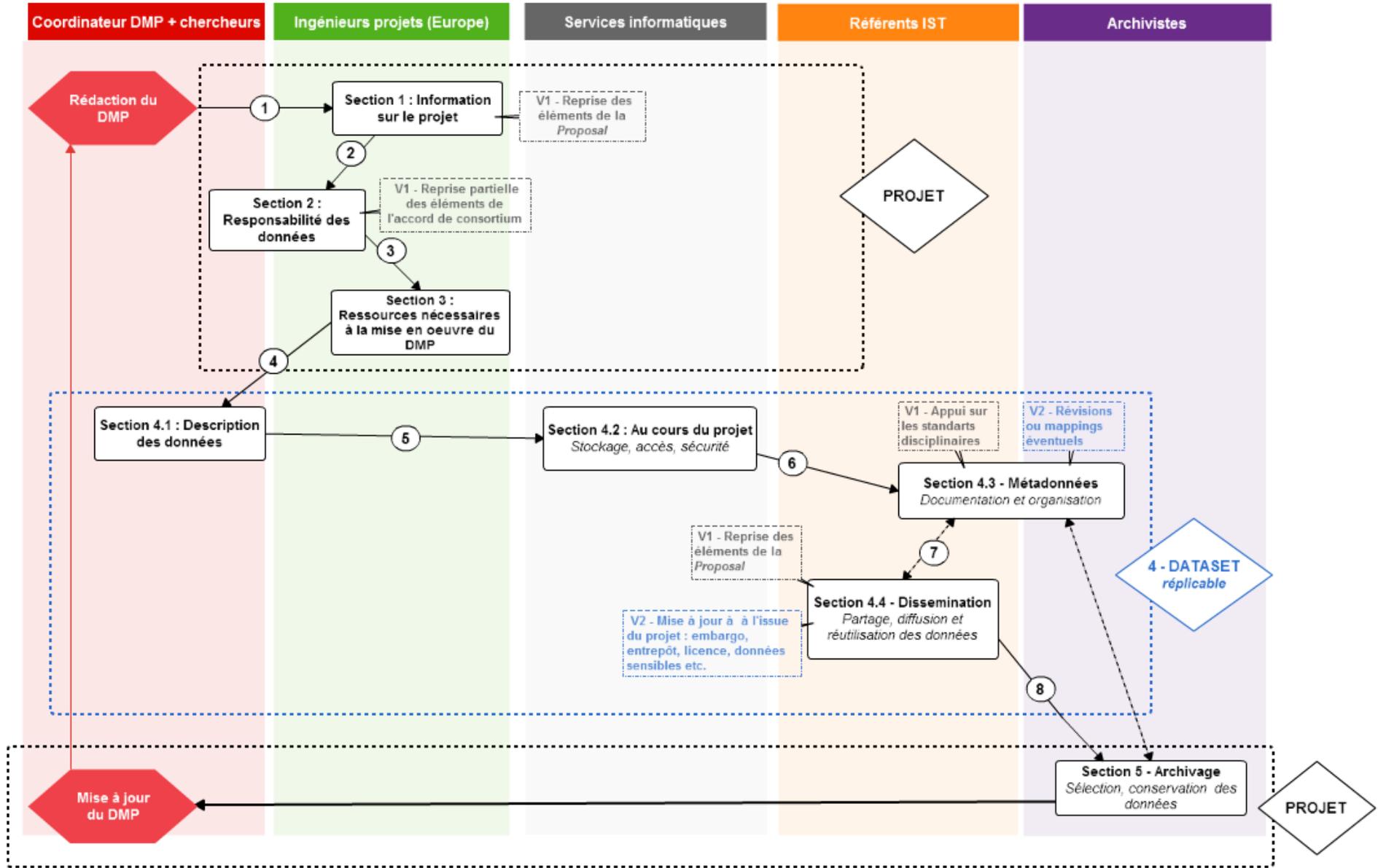
✓ Ingénieur-projet : coordonne les actions autour du DMP

✓ Informaticien : interlocuteur pour le stockage et la sécurisation des données, les aspects infrastructures et les coûts associés

✓ Spécialiste de l'IST : en appui au chercheur, propose des standards, des alignements (mappings) avec des référentiels existants

✓ Documentaliste et/ou archiviste : les mieux placés pour aider le chercheur à sélectionner les données pour la conservation, définir les durées et les solutions techniques







Plan de gestion des données

les 4 grandes parties

1. La description des jeux de données : quels types de données recueillies ou générées par le projet?
2. Les standards et métadonnées : quel(s) standard(s) avez-vous utilisé pour les métadonnées?
3. Le partage des données : comment ces données seront-elles exploitées, partagées, rendues accessibles?
4. L'archivage et la conservation des données : comment ces données seront-elles administrées et conservées?





Plan de gestion des données structure en 8 séctions

Séction 1-2 : Informations générales sur le projet

Séction 3 : Informations générales sur les données

Séction 4 : Sécurité, gestion de risques et éthique

Séction 5 : Jeux de données

Séction 6 : Principes de dissémination et protection

Séction 7 : Archivage des données

Séction 8 : Ressources nécessaires





Plan de gestion des données outils disponibles

✓ [DMP OPIDoR](#) : adaptation de DMPOnline par le CNRS

✓ [DMPonline](#) :

-créé par le Digital Curation Center (DCC)

-conforme aux recommandations de la Commission Européenne (intègre H2020)

-propose modèles et conseils

✓ [DMPTool](#):

-créé par l'Université de Californie

-modèle simplifié de DMP





Plan de gestion des données pour s'entraîner et savoir plus

✓ [CatOPIDoR](#)

✓ [DoRANum](#)

Bibliographie :

[European Commission Directorate-General for Research & Innovation H2020 Programme, Guidelines on FAIR Data Management in Horizon 2020, Version 3.0, 26 July 2016](#)

[Cartier Aurore, Moysan Magalie, Reymonet Nathalie, Réaliser un plan de gestion de données, Paris: Université Paris Diderot, Sorbonne Paris Cité, Université Paris Descartes, version 1, 9 janvier 2015](#)

[Inist,-CNRS Tutoriels “Le libre accès aux résultats de la recherche dans le cadre d’Horizon 2020”, juin 2016](#)

[DCC \(Digital Curation Center\), Checklist for a Data Management Plan. version 4.0, 2013, Edinburgh](#)

[Deboin, M.C. 2014. Découvrir des plans de gestion de données de la recherche, en 4 points., Montpellier \(FRA\) : CIRAD, 6 p. <http://url.cirad.fr/ist/donnees-de-la-recherche>](#)

