



Créer un plan de gestion des données (DMP)

Marie Puren

ANF – INIST – Vandoeuvre-Les-Nancy – 6 juillet 2017



Objectifs

A la fin de cet atelier, les participants sauront:

- Définir un plan de gestion des données
- Identifier les éléments clés d'un DMP
- Créer un DMP

Introduction

« Une gestion des données efficace assure des bénéfices **tout au long et au-delà** de la durée de vie du projet de recherche [...] Pour le chercheur, la perception des données comme un instrument de recherche et un nouveau savoir, peut être à la source de **changements profonds**. Des données bien gérées mènent à des données de recherche de **meilleure qualité**, accroît leur **visibilité** mais aussi les bénéfices connexes liés à des **taux de citation améliorés**. »

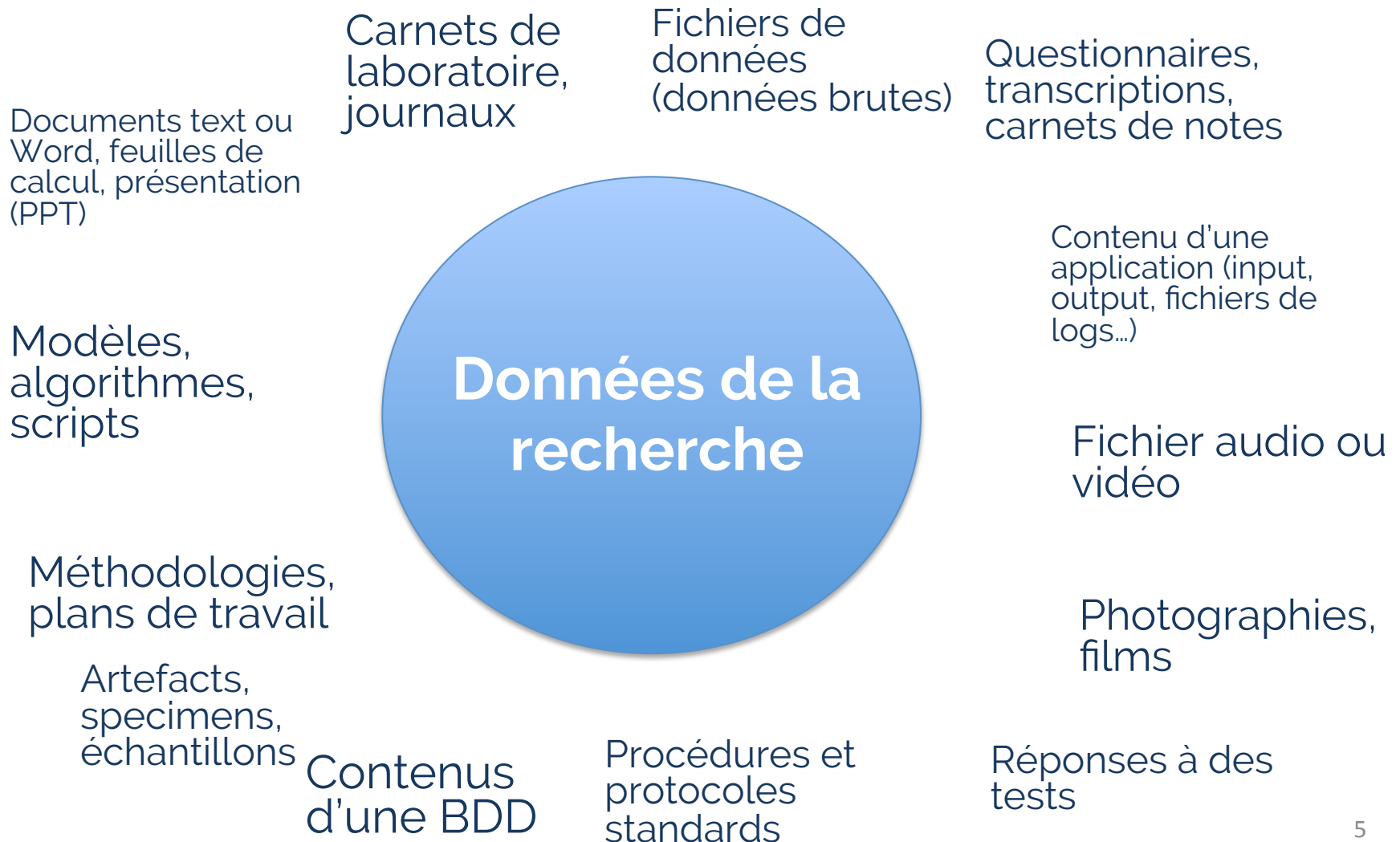
[How to develop RDM Services](#), Digital Curation Center

Les données de la recherche

« Les données de la recherche sont créées comme des résultats directs de “faire de la recherche”. En ce sens, cela exclut les supports destinés à l'enseignement et les documents administratifs (comme les fiches de poste, les mails ou les rapports financiers). Les données de la recherche peuvent être disponibles dans une infinie variété de formats. »

University of Bristol,
[“A brief guide to research data”](#)

Les données de la recherche

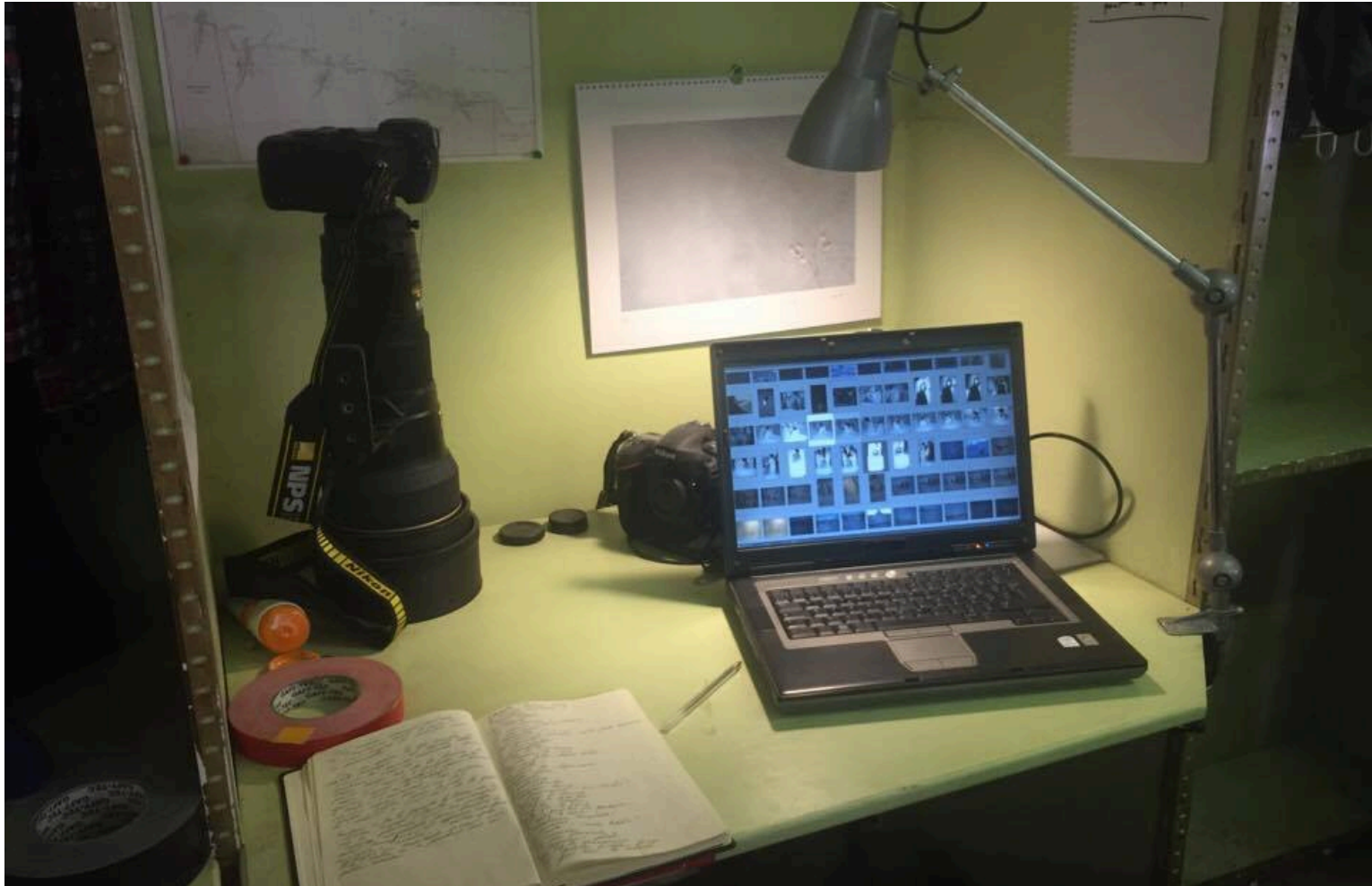


Les données de la recherche

On peut définir les données de la recherche de la façon suivante:

- 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



Exposition *Antartica*, 26 avril 2016-17 avril 2017, Musée des confluences, Lyon (<http://www.exposition-antarctica.com/>) 7

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”. Rémi Gaillard, [De l'Open data à l'Open research data : quelle\(s\) politique\(s\) pour les données de recherche?](#)
 - Série de données compilées dans plusieurs documents différents, ou série de données compilées dans un document unique - par exemple, dans un fichier CSV
 - en fonction de la discipline, mais aussi des projets et des objectifs de la recherche, “l'échelle à laquelle ces données assemblées acquièrent leur unité pour former un ‘jeu’” va beaucoup varier. (*Ibid.*)

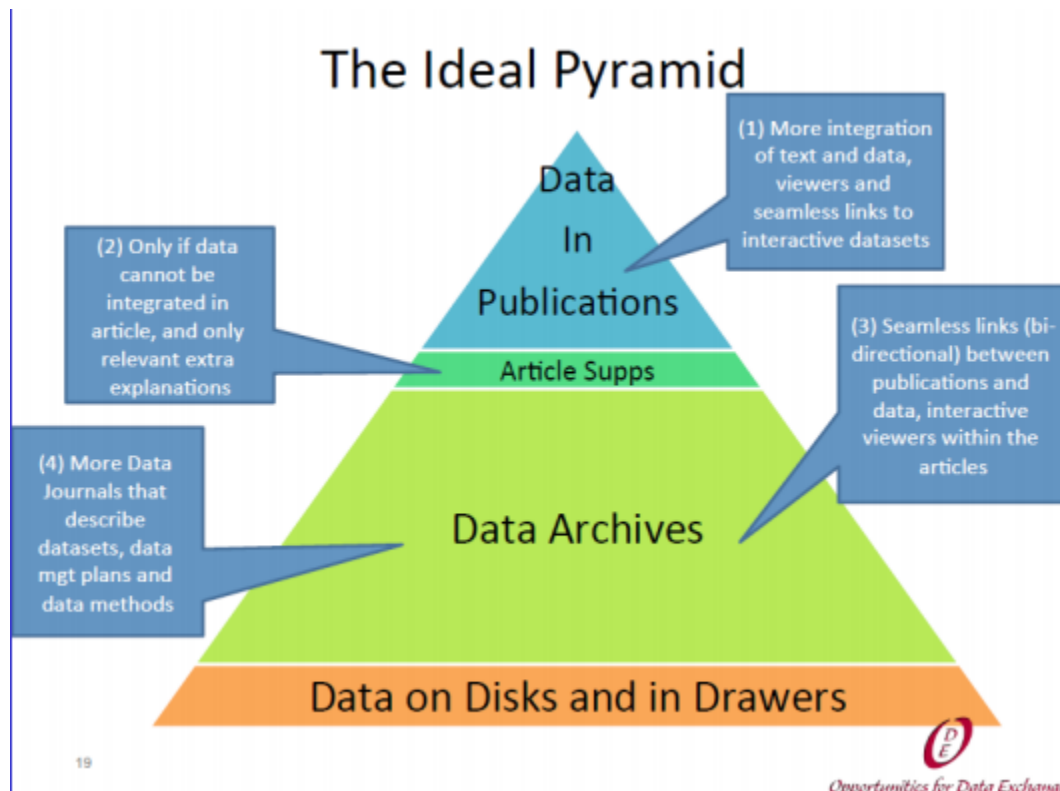
Une nouvelle demande

- Demande : données de recherche = accessibles
- Plus conservées dans des données personnels ou au fond d'un tiroir

Une nouvelle demande

La « pyramide idéale » :

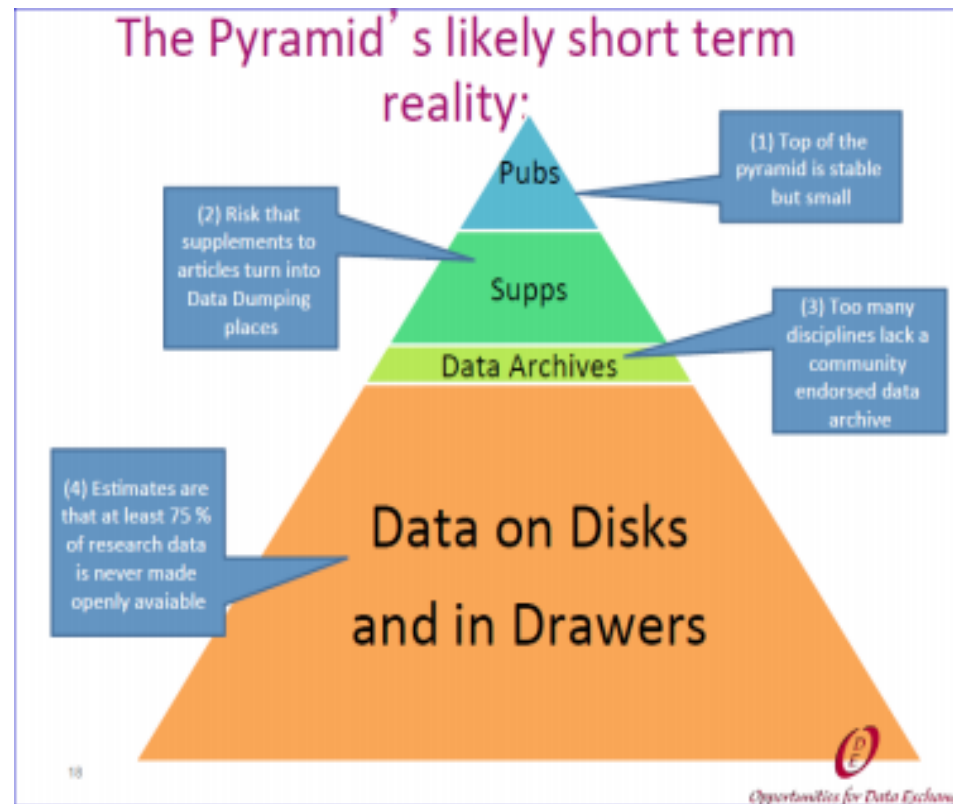
- très peu de données inaccessibles
- la grande majorité d'entre elles dans des archives et des articles



Une nouvelle demande

La « pyramide réelle » :

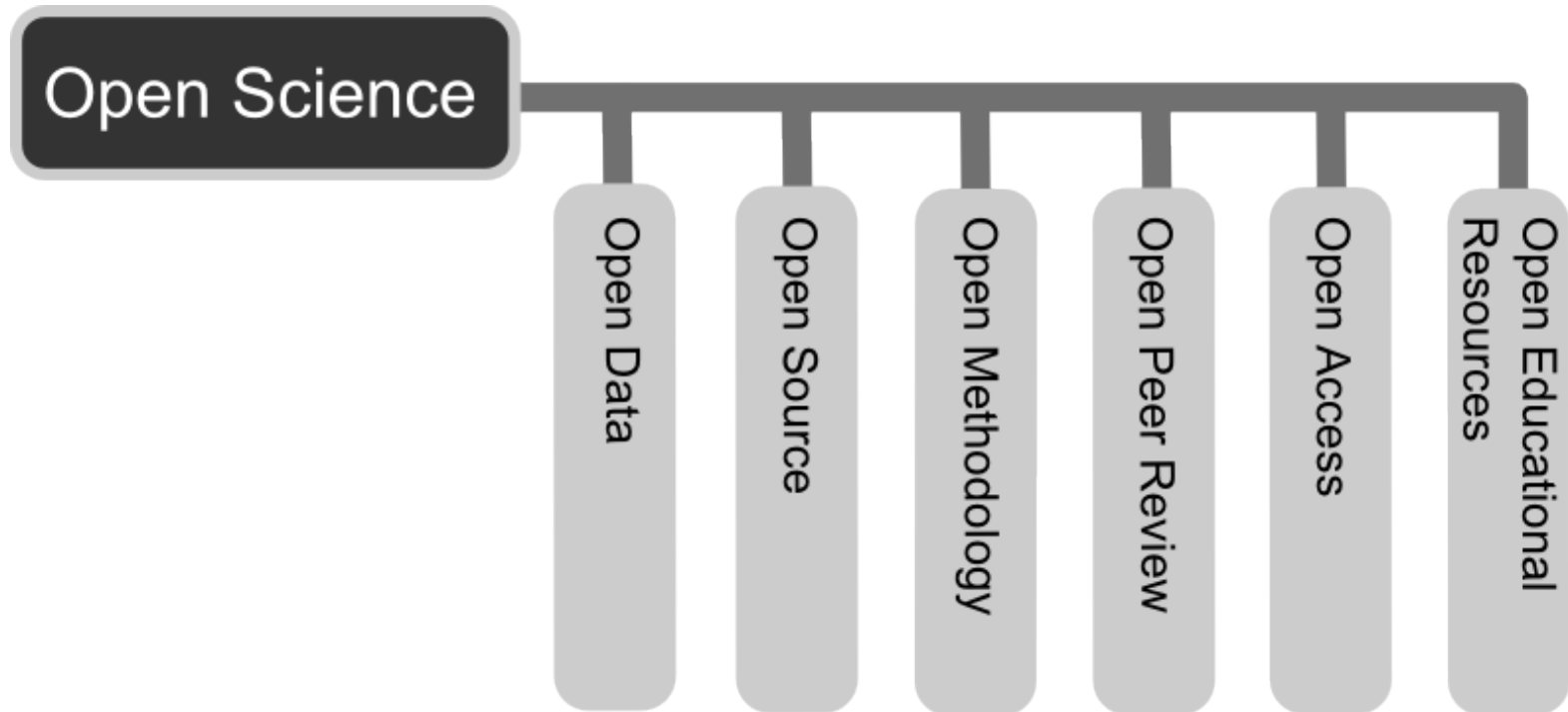
- Peu de données accessibles
- la grande majorité d'entre elles dans des « tiroirs » ou des disques durs



Un nouveau modèle d'ouverture pour les données de 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

Un nouveau modèle d'ouverture pour les données de recherche



Par [Andreas E. Neuhold](#), travail personnel- basé sur "[The taxonomy tree](#)", FOSTER (Facilitate Open Science Training for European Research)

Pour plus d'infos sur la Science ouverte: Michael Nielsen, [Reinventing Discovery: The New Era of Networked Science](#), Princeton University Press, 2011.

Un nouveau modèle d'ouverture pour les données de recherche

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 recherche

- 2003 : Déclaration fondatrice de Berlin en faveur de l'Open Access : les données de la recherche sont assimilées aux publications.
- 2006 : Déclaration sur l'accès aux données de la recherche financée par des fonds publics par l'OCDE. L'OCDE insiste sur la nécessité d'ouvrir les données produites dans le cadre de recherches financées sur fonds publics.

Un nouveau modèle d'ouverture pour les données de recherche

- 2007 : Le Conseil scientifique du Conseil européen de la recherche (CER) publie des recommandations demandant la mise en accès libre des résultats de recherches financées par le CER, dans un délai de six mois suivant leur publication, dans des archives ouvertes disciplinaires appropriées (research repositories) ou des archives ouvertes institutionnelles (institutional repository).
- Fin 2009 : Création de l'infrastructure OpenAIRE (Open Access Infrastructure for Research in Europe) pour les publications scientifiques en libre-accès

Un nouveau modèle d'ouverture pour les données de recherche

- Fin 2011, OpenAIREplus lui succède et étend son périmètre aux données de la recherche.
- 2012 : Recommandation de la Commission européenne aux États-membres “relative à l'accès aux informations scientifiques et à leur conservation”. Elle pose comme principal objectif qu'avant 2014, “des politiques de libre-accès aux articles et données scientifiques [soient] établies dans tous les États membres à tous les niveaux pertinents”. Cette ambition rejoint l'idée de la création d'un espace européen de la recherche, qui se veut une “innovation ouverte”.

Un nouveau modèle d'ouverture pour les données de recherche

- 2013 : Open Research Data Pilot : mise en place d'une politique européenne en matière de libre-accès.
- 2015: Le Plan d'action de l'ANR et son *Appel à projets générique 2016* encourage les chercheurs à promouvoir leurs résultats en accès ouvert.
- 2016: Extension de l'Open Research Data Pilot à tous les projets Horizon 2020

Soutenue par des initiatives européennes et nationales

Horizon 2020 – programme européen pour la Recherche et l'Innovation

The Pilot on Open Research Data (ORD Pilot)

« Le pilote ORD s'applique essentiellement aux **données nécessaires pour valider les résultats** présentés dans des publications scientifiques. **D'autres données** peuvent aussi être fournies par les bénéficiaires sur la base du **volontariat**, comme mentionné dans leurs Plans de Gestion des Données ».

[H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#),
Version 3.0, 26 July 2016, p.3.

Soutenue par des initiatives européennes et nationales

Extension du pilote ORD en juillet 2016

« La Commission européenne a permis l'accès aux et à la réutilisation des données de la recherche générées par les projets Horizon 2020 grâce au Pilote ORD. Comme mentionné dans le programme de travail 2017 (2017 work program), le pilote est étendu pour couvrir tous les domaines thématiques [du programme H2020] [...] En étendant le pilote, l'Open access devient le paramètre par défaut pour les données de recherche générées dans le cadre d'Horizon 2020. »

[H2020 Programme Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#), Version 3.0, 26 July 2016, p.8.

Partager des données, mais comment?

Comment puis-je faire?



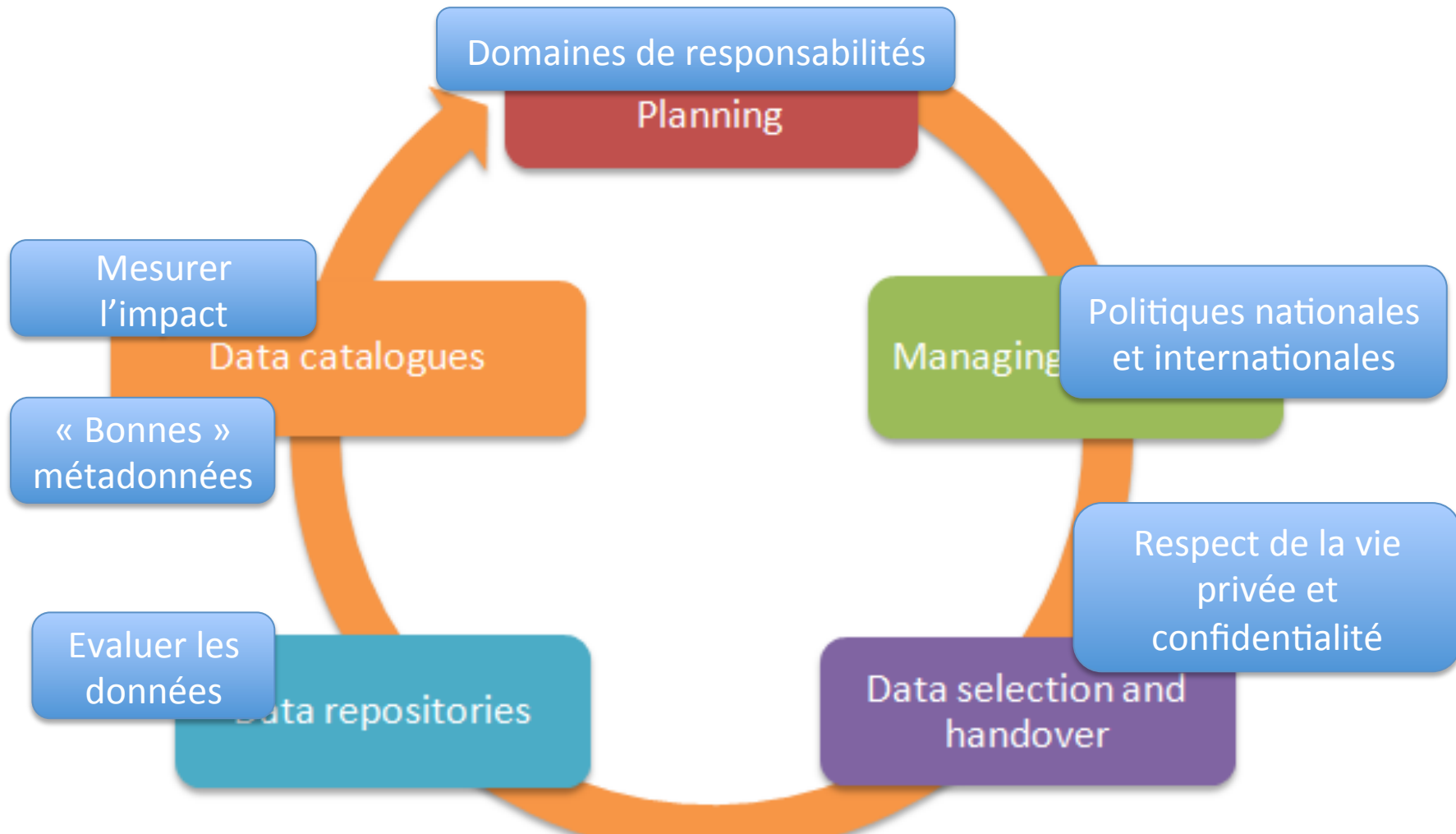
En gérant vos données

La gestion des données implique généralement:

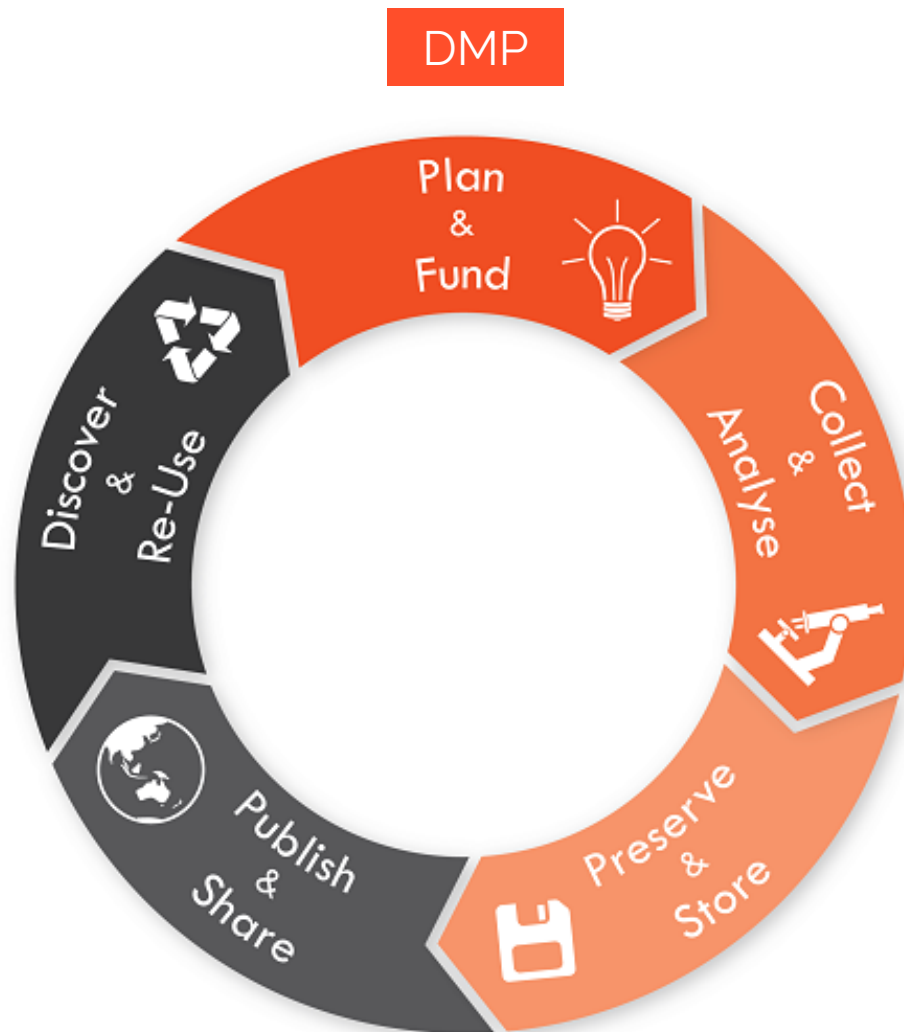
- De créer un **Plan de Gestion des Données** (ou **DMP – Data Management Plan**), soumis avec la demande de financement, afin d'expliquer comment les données seront gérées pendant et après l'achèvement du projet ;
- De stocker des données de recherche durant le projet et de les partager avec les collègues autorisés ;
- A la fin du projet, de déposer ces données dans un entrepôt pérenne, et de les rendre disponibles pendant plusieurs années.

Research data management (or RDM)

Developing RDM services



En créant un plan de gestion des données



Créer un plan de gestion des données

Avec le stockage des données et le dépôt de celles-ci dans un entrepôt pérenne, créer un plan de gestion de données est l'une des étapes les plus importantes dans la gestion des données de la recherche.

Créer un plan de gestion des données

Définition

« Les plans de gestion des données sont des éléments-clé d'une bonne gestion des données. Un plan de gestion des données décrit le cycle de vie de la gestion des données qui seront recueillies, traitées et/ou générées [...]. »

[H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#),
Version 3.0, 26 juillet 2016, p.4.

Créer un plan de gestion des données

Définition

Un document formel qui décrit:

- Les données que vous espérez acquérir ou générer au cours du projet de recherche,
- Comment vous allez gérer, décrire, analyser et stocker ces données,
- Quels mécanismes vous utiliserez à la fin de votre projet pour partager et préserver ces données.

Créer un plan de gestion des données

Définition

- Une feuille de route régulièrement mise à jour ;
- Un document normalisé ;
- Son contenu varie en fonction des exigences du projet et des demandes des agences de financement ;
- Il se concentre sur les données et jeux de données recueilli(e)s, créé(e)s, analysé(e)s par le projet de recherche.

Créer un plan de gestion des données

Définition

DMP = livrable du projet, mais pas un document technique

- Il matérialise la politique du projet en matière de données ;
- Il résume les buts et les actions qui seront mis en place ;
- Il répond aux exigences des financeurs.

Créer un plan de gestion des données

Pourquoi?

L'étape principale de la gestion des données de recherche

« Gestion » active des données numériques (=> le « M » de DMP):

- Une gestion continue (sauvegarde, migration, conversion...), tout au long du cycle de vie des données ;
- Un plan d'action en termes de qualité des données, de faisabilité techniques et de viabilité financière.

Créer un plan de gestion des données

Pourquoi?

L'étape principale de la gestion des données de recherche

Gestion des données ≠
« Intendance » (stewardship) des données

=

Optimiser des ressources dans un but spécifique

- En identifiant et en rendant visibles les actions à conduire ;
- En planifiant les étapes-clé, les dates butoir et les périodes cruciales.

Créer un plan de gestion des données

Pourquoi?

L'étape principale de la gestion des données de recherche

Gestion active et curation numérique

« Les activités de curation de données permettent de faciliter la découverte et la récupération de données, de maintenir la qualité des données, de leur ajouter de la valeur et d'en fournir pour de futures réutilisations. Ce nouveau champ inclut la représentation, l'archivage, l'authentification, la gestion, la préservation, la récupération, et l'utilisation. »

Créer un plan de gestion des données

Pourquoi?

L'étape principale de la gestion des données de recherche

Curation numérique ≠ Stockage de données

« Le stockage de données se résume à garder les données existantes et d'assurer qu'on pourra y accéder en fonction des besoins. Il n'implique pas nécessairement des **pratiques de rafraichissement ou de migration de format** (essentiel pour maintenir les données dans une forme utilisable), ni des pratiques de curation **comme l'amélioration des données par l'addition de métadonnées, ou la migration d'un standard vers un autre**. La curation de données va donc beaucoup plus loin que le stockage des données. »

[Digital Humanities Data Curation](#)

Créer un plan de gestion des données

Pourquoi?

Les « FAIR Data »

Un DMP « aide les bénéficiaires du programme Horizon 2020 à rendre leurs données de recherche 'findable' (trouvable), 'accessible' (accessible), 'interoperable' (interopérable) et 'reusable' (réutilisable) (FAIR) [...]. »

[H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#),
Version 3.0, 26 July 2016, p.3.

Créer un plan de gestion des données

Pourquoi?

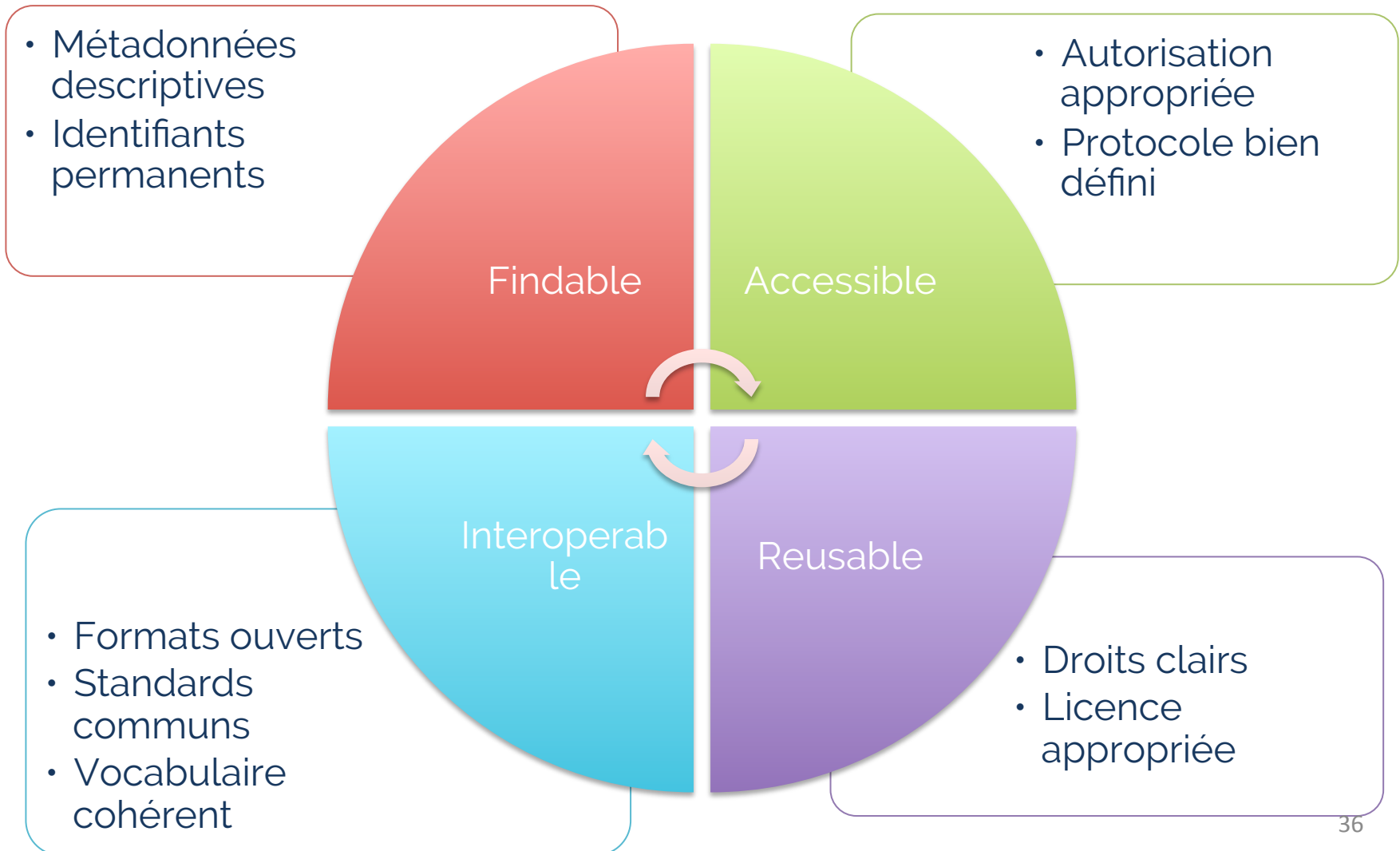
FAIR Data

- Janvier 2014: une rencontre est organisée par le Netherlands eScience Center et le Dutch Techcentre for the Life Sciences (DTL) au Lorentz Center à Leyde.
- **Les principes FAIR:** “les producteurs et les consommateurs de données – à la fois machines et humains – peuvent plus facilement découvrir, avoir accès, interopérer et réutiliser intelligemment, avec des citations adéquates, les vastes quantités d’information qui sont générées par la science contemporaine et son usage intensif des données. » ([Force 11](#))

Créer un plan de gestion des données

Pourquoi?

FAIR Data



Créer un plan de gestion des données

Pourquoi?

Pour satisfaire les exigences du programme H2020

- Une première version du DMP (livrable) : dans les **6 premiers mois** du projet.
- A l'étape de la proposition ; fournir une **courte description** de la politique de gestion des données.
- Depuis juillet 2016: **tous les projets H2020** doivent fournir un DMP.

Créer un plan de gestion des données

Pourquoi?

Pour satisfaire les exigences du programme H2020

- Le DMP doit être régulièrement **mis à jour**.
- H2020 incite ainsi la publication d'un DMP mis à jour au milieu du projet (**Mid-term review DMP**)
- H2020 exige au minimum une nouvelle version du DMP avec les mises à jours nécessaires à la fin du projet (**Final review DMP**)

Créer un plan de gestion des données

Pourquoi?

Pour satisfaire les exigences du programme H2020

Exigences minimales (1^{ère} version du DMP):

- Une **description des données** qui seront générées ou collectées ;
- Les **standards** et les **métadonnées** qui seront utilisées ;
- Le **partage** des données ;
- L'**archivage et la préservation** (les principes FAIR).

Créer un plan de gestion des données

Pourquoi?

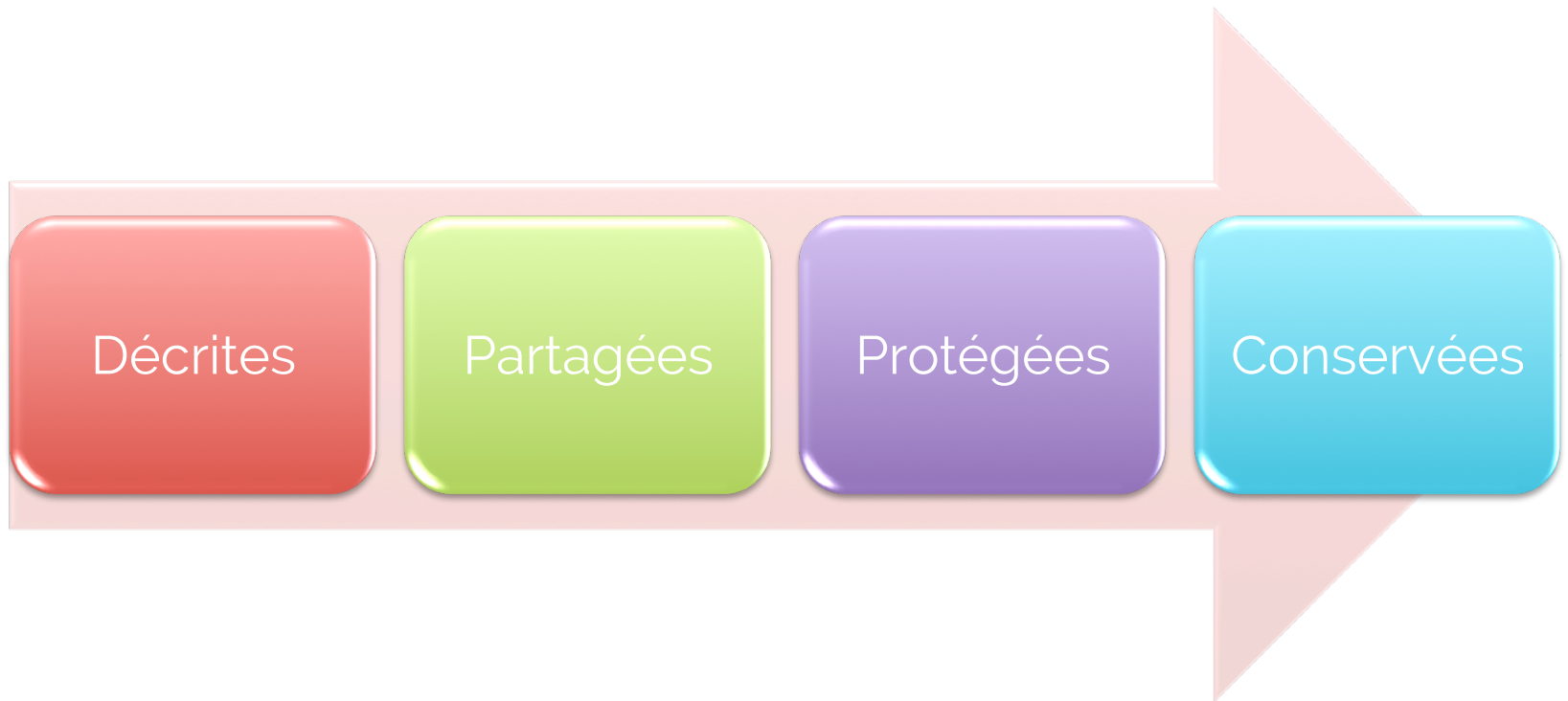
Parce que c'est une bonne pratique de recherche qui permet de

- 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 car un DMP peut améliorer la gestion des données

Créer un plan de gestion des données

En résumé

- Quand on crée un DMP, on définit comment les données du projet seront:



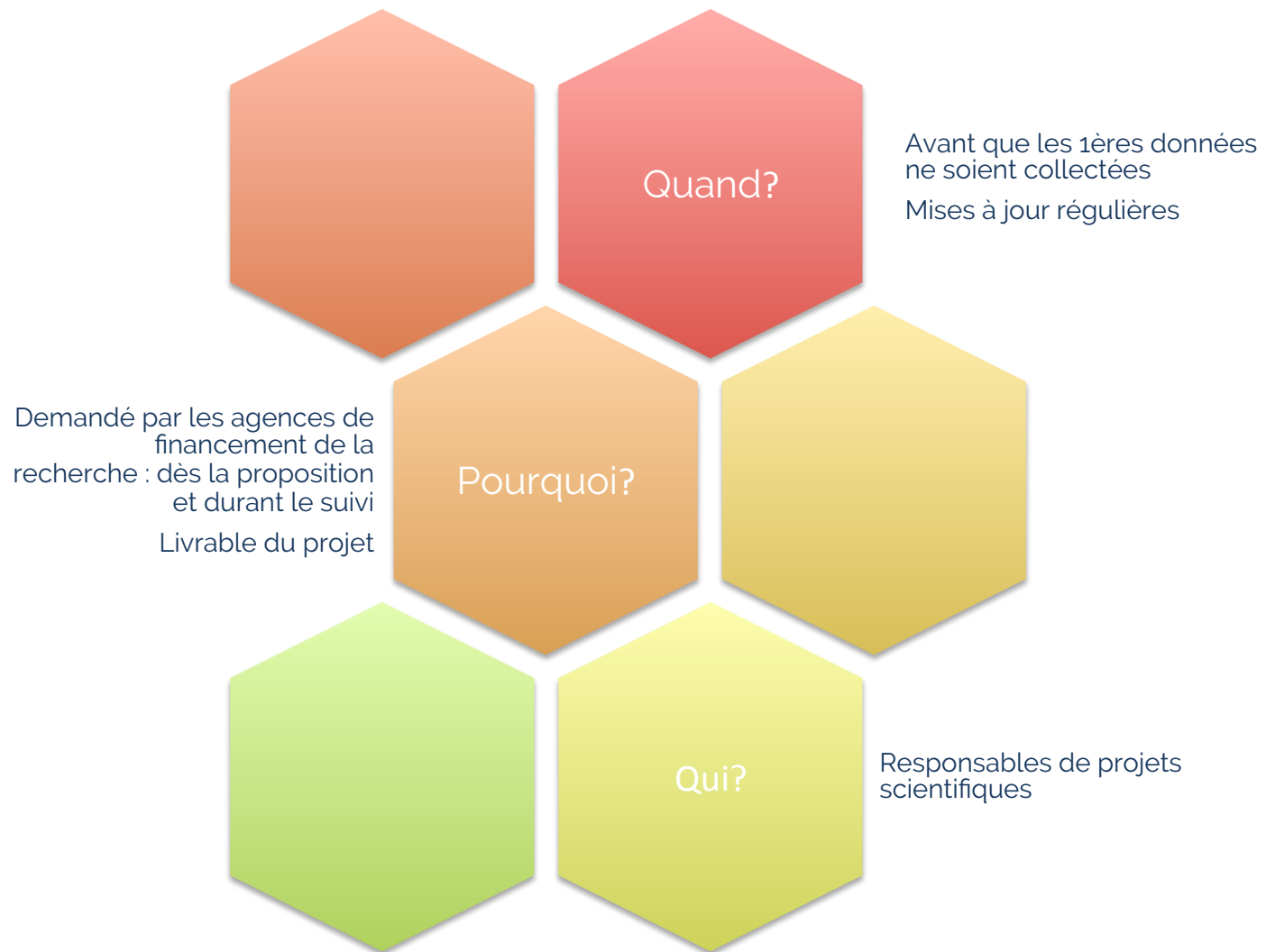
Créer un plan de gestion des données

En résumé

- Il contient
 - Une description du cycle de vie des données (pendant et après le projet)
 - Une description des données
 - Une description de la politique associée aux données (accès, diffusion, propriété intellectuelle...)
 - Un chiffrage des coûts
- Il participe à
 - La **sécurisation** et
 - La **pérennisation** des données

Créer un plan de gestion des données

Les premières questions



Créer un plan de gestion des données

Définir les responsabilités



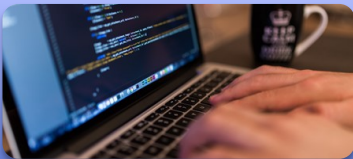
Chercheur

- 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



Spécialiste de l'IST

- Propose des standards, des alignements avec des référentiels existants



Documentaliste et/ou archiviste

- Aide le chercheur sélectionner les données, définir les durées de conservation et les solutions techniques

Créer un plan de gestion des données

Comment se structure un DMP?

Les informations administratives du projet

Il faut d'abord indiquer **l'identité** du projet :

- Son identifiant
- Son nom
- Une description (court résumé)
- Les agences de financement
- Le(s) responsable(s) du projet (nom, coordonnées, identifiants – identifiant [ORCID](#) par exemple)
- Le(s) responsables de la gestion des données (nom, coordonnées)
- Dates :
 - De la 1^{ère} version du DMP
 - De la dernière mise à jour

Créer un plan de gestion des données

Les grandes parties d'un DMP

Pour chaque jeu de données:

Il faut indiquer :

- Un nom
- Un identifiant

Créer un plan de gestion des données

Les grandes parties d'un DMP

Pour chaque jeu de données, il faut enfin respecter les 4 grandes parties d'un DMP:

1. La description des jeux de données : 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?

Créer un plan de gestion des données

1. La description des données

Quels seront les types de données générées ou recueillies par le projet?

- L'origine des données
- Leur nature et leur échelle
- Qui pourraient les utiliser?
- Appui d'une publication scientifique?
- Y a-t-il des données similaires?

Créer un plan de gestion des données

1. La description des données

Origine des données

- Source:
 - Collecte de jeux de données existants?
 - Création de jeux de données nouveaux?
 - Les deux?

Créer un plan de gestion des données

1. La description des données

Origine des données

- Préciser en détail **la méthodologie adoptée** pour recueillir ces jeux de données
 - Données existantes: accès à ces données (citation de la source, modalités d'accès, description synthétique)
 - Données nouvelles : processus de création ou le mode opératoire mis en place pour les obtenir

Créer un plan de gestion des données

1. La description des données

Description formelle et technique

- **Type de données** : source, forme, stabilité, volumétrie
- **Format**, ouverture, encodage, compression
- **Systeme d'organisation, de nommage et de gestion** des répertoires et des fichiers

Créer un plan de gestion des données

1. La description des données

Type de données

- Source :
 - données d'observation,
 - expérimentales,
 - de simulation,
 - dérivées ou compilées (ex. : fouille de texte)
- Forme:
 - données textuelles (notes, carnets de laboratoire)
 - données numériques (tableaux)
 - données audiovisuelles
 - modèles/ codes informatiques
 - données spécifiques liées à certains instruments

Créer un plan de gestion des données

1. La description des données

Type de données

- Stabilité des données :
 - données fixes : elles ne changent pas après avoir été recueillies
 - croissantes : de nouvelles données peuvent être ajoutées, mais les anciennes ne seront pas modifiées ou supprimées
 - révisables : de nouvelles données peuvent être ajoutées, et les anciennes peuvent être modifiées ou supprimées
- Versions: organiser, identifier et sauvegarder les différentes versions (V1, V2, V3...etc.)
- Volume : préciser le volume envisagé pour chacun des jeux de données:
 - Capacités de stockage?
 - Budget nécessaire à prévoir

Créer un plan de gestion des données

1. La description des données

Format, ouverture, encodage, compression

- Multitude de formats possibles, pour une multitude de types de données

Formats fermés	Formats ouverts
Excel (.xls, .xlsx)	Comma separated value (.csv)
Word (.doc, .docx)	Plain text (.txt) ou si le formatage est nécessaire PDF/A (.pdf)
PowerPoint (.ppt, .pptx)	Alternative Format PDF/A (.pdf)
Photoshop (.psd)	TIFF (.tif, .tiff)
Quicktime (.mov)	MPEC-4 (.mp4)

Créer un plan de gestion des données

1. La description des données

Systeme d'organisation, de nommage et de gestion des répertoires et des fichiers

- Quelle(s) convention(s) avez-vous choisi?

Exemple :

Titre_projet_Jeudedonnées1(identifiant)_aa
aammjj

Créer un plan de gestion des données

1. La description des données

Systeme d'organisation, de nommage et de gestion des répertoires et des fichiers

Quelques principes de base

- Harmoniser les dates
- Identifier les différentes versions
- Mettre en place des sauvegardes
- Organisation pertinente et cohérente des dossiers et sous-dossiers
- Choisir des noms homogènes et pertinents
- Supprimer des versions obsolètes si nécessaire

Créer un plan de gestion des données

2. Standards et métadonnées

Utilité des métadonnées

- Métadonnées : données sur les données
- Elles permettent de décrire des ressources numériques en précisant leurs propriétés
 - Contenu intellectuel (résumé, mots-clés...)
 - Contexte de production (date, producteur...)
 - Caractéristiques des fichiers et des données (formats, taille...)
 - Propriété intellectuelle et droits d'utilisation
- Elles facilitent
 - l'accessibilité des données,
 - leur interopérabilité,
 - leur pérennité.
- Elles sont indispensables à la rédaction d'un DMP

Créer un plan de gestion des données

3. Le partage des données

Les modalités de dépôt

- Où déposer?
 - Dans un entrepôt de données de la recherche
 - A préciser dans le DMP : quel entrepôt vous allez utiliser?
 - Pour trouver un entrepôt de données, des répertoires et des catalogues sont à votre disposition:
 - La page wiki “Data repositories” fait partie du projet [Open Access Directory](#) . Elle fournit une liste d’entrepôts pour les données ouvertes. Classé par discipline, chaque entrepôt est aussi décrit:
http://oad.simmons.edu/oadwiki/Data_repositories
 - [re3data](#) ou “Registry of Research data Repositories” est un catalogue d’entrepôts de données.

Créer un plan de gestion des données

3. Le partage des données

Quelques exemples d'entrepôts

Nom	URL	Type	Description
Zenodo	https://zenodo.org/	Institution publique	Fruit de la coopération entre le CERN et OpenAire. il accueille des jeux de données très différentes et fournit un DOI (Digital Object Identifier).
Dryad	http://datadryad.org/	Organisation à but non lucratif	Entrepôt de données à vocation pluridisciplinaire, mais dans les faits, il contient surtout des données issues du monde médical.

Créer un plan de gestion des données

3. Le partage des données

Quelques exemples d'entrepôts

Nom	URL	Type	Description
Datahub	https://datahub.io/fr/	Organisation à but non lucratif	Entrepôt interdisciplinaire. Il est le fruit du travail de l'Open Knowledge Foundation
Figshare	https://figshare.com/	Organisation à but non lucratif	Les chercheurs peuvent déposer gratuitement leurs données dans l'entrepôt. Figshare propose aussi une solution payante aux institutions pour les aider à gérer leurs données.

Créer un plan de gestion des données

3. Le partage des données

Les modalités de dépôt

- Quand déposer?
 - Jeux de données générés ou recueillis dans le cadre du projet : dès que possible!
 - Autres données? Possible, à préciser dans le DMP avec la date limite de dépôt

Créer un plan de gestion des données

3. Le partage des données

Les modalités de partage

- Procédures d'accès aux données
 - Accès immédiat ou embargo
 - Accès ouvert à tous ou restreint à des groupes spécifiques
- Comment peut-on y accéder?
 - Outils et instruments nécessaires pour vérifier et réutiliser les données (logiciels, protocoles d'analyse...)
 - Associer un identifiant unique et pérenne à vos données. Exemple : [DOI](#) pour les objets, [ORCID](#) pour les contributeurs/producteurs

Créer un plan de gestion des données

3. Le partage des données

Les licences

- Pour que les données soient réutilisables, il est nécessaire de les associer à des licences.
- Ajouter la déclaration de ce type aux données : « Ces données/ Cette base de données est mise à disposition sous la [licence choisie]. Le texte complet de [nom de la licence choisie et version si nécessaire] est disponible ici : [url correspondante] »

Créer un plan de gestion des données

3. Le partage des données

Les licences

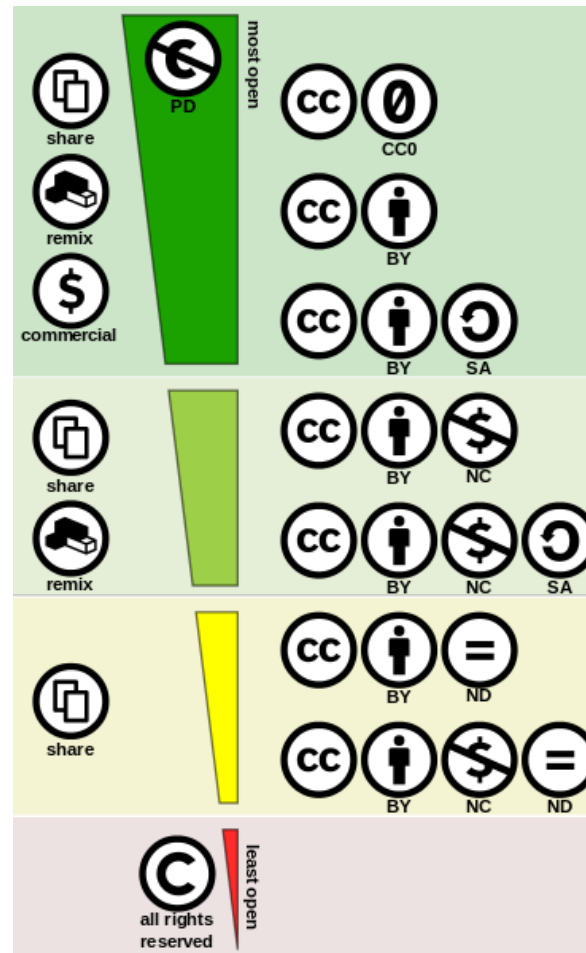
- Les licences [Creative Commons](#) créées en 2002 pour la diffusion de contenus numériques.
 - Combinent 4 clauses afin de préciser les droits auxquels l'utilisateur renonce ou non
 - Attribution (sigle BY) : paternité.
 - Non commercial (sigle NC)
 - Pas de modification (sigle ND - No Derivative Works)
 - Partage dans les même conditions (sigle SA - Share Alike)

Recommandation : utiliser une licence CC-BY (respect de l'attribution).

Créer un plan de gestion des données

3. Le partage des données

Les licences
Creative Commons



Créer un plan de gestion des données

3. Le partage des données

Les licences

- La licence [CC-by 4.0](#) (Creative Commons Attribution) impose seulement de créditer l'auteur(s) des données.
- La licence [CCo](#) (Creative Commons Public Domain Dedication) de type domaine public. Imposée par l'entrepôt de données pluridisciplinaire Dryad.

Créer un plan de gestion des données

3. Le partage des données

Les licences

- La licence [ODC-by](#) (Open Database Commons) est très proche de la licence CC-by.
- La licence [ODC-ODBL](#) (Open database License) est très proche de la licence Creative Commons by-sa.

Créer un plan de gestion des données

3. Le partage des données

Les licences

- La licence [PDDL](#) (Public domain dedication and license) de type domaine public
- La [Licence ouverte](#) (LO) créée en 2011 par la mission Etalab pour faciliter la diffusion des données publiques françaises sur la plateforme data.gouv.fr

Créer un plan de gestion des données

3. Le partage des données

Les exceptions

- Du point de vue légal, certains jeux de données ne sont pas librement communicables:
 - les données à caractère personnel => anonymisation (si possible)
 - les données relatives à l'exécution d'un contrat de prestation de services exécuté pour le compte d'une ou de plusieurs personnes déterminées (non publiques)
 - les données relatives à la sécurité publique ou au secret
 - les données relatives au secret professionnel.
- Du point de vue éthique, il peut aussi être nécessaire de ne pas divulguer certaines données.
- Dans le DMP, il faut expliquer les raisons qui vous empêchent de partager vos données.

Créer un plan de gestion des données

4. L'archivage des données

Archiver ≠ diffuser

- Diffuser:
 - Accès facile aux données
 - Explorer les données
 - Navigation riche
- Archiver:
 - Conserver sur le long terme
 - Accès pérenne dans le temps
 - Navigation simple (en général)

Créer un plan de gestion des données

4. L'archivage des données

Des compétences multiples

- Techniques et informatiques : veille sur les supports, les formats, les technologies
- Archivistiques : politique d'archivage
- Documentaires : métadonnées et catalogage
- Juridiques
- Scientifiques

Créer un plan de gestion des données

4. L'archivage des données

Métadonnées de préservation

- Cela requiert aussi de conserver toutes les **données d'accompagnement en plus** des données brutes.
- Avec les métadonnées de préservation, on assure:
 - La compréhension des données
 - Le contexte de leur collecte
 - Les conditions de restitution

Créer un GDlan de gestion des données

4. L'archivage des données

Où archiver ses données?

- Evaluation et labellisation des sites d'archivage pérenne:
 - [Data Seal of Approval](#)
 - [Drambora](#) (Digital Repository Audit Method Based on Risk Assessment) : méthodologie pour l'auto-évaluation (DCC)
 - [TRAC](#) (Trustworthy Repositories and Certification Checklist) : outils pour l'audit de sites d'archivage (DCC)
- En France:
 - CINES (Centre Informatique National de l'Enseignement supérieur) : [plateforme d'archivage](#)
 - Huma-Num : [service d'archivage](#) pour les données en SHS

Créer un plan de gestion des données

Les outils disponibles

- [DMPOnline](#) :
 - créé par le Digital Curation Center (DCC)
 - En accord avec les recommandations de la Commission Européenne (intègre H2020)
 - Propose modèles et conseils
- [DMP OPIDoR](#) : adaptation de DMPOnline par le CNRS
- [DMPTool](#):
 - Créé par l'Université de Californie
 - Modèle simplifié de DMP

Bibliographie (1)

Définir les données de la recherche

- Fayett, Sylvie, « [Données](#) » de la recherche, les mal-nommées » dans *UrfistInfo*, 15 novembre 2013.
- Fournier Thierry, « [Les données de la recherche : définitions et enjeux](#) », in *Arabesques*. Janvier-mars 2014, n°73, pp. 4-6.
- Rosemberg Noëmie, « [De la définition des données de la recherche](#) » dans *En quête des données*, le 30/09/2015,

Pourquoi partager les données de la recherche?

- LeClere Felicia, “[Too Many Researchers Are Reluctant to Share Their Data](#)”, *The Chronicle of Higher Education*, August 3, 2010
- Mulvany Ian, “[Open Data, Challenges towards implementation and possible solutions](#)”, *mulvany.net*, September 2016
- Pienta Amy M., Alter George, Lyle Jared, [The Enduring Value of Social Science Research: The Use and Reuse of Primary Research](#), IS CPR (Inter-university Consortium for Political and Social Research), Institute for Research, University of Michigan

Bibliographie (2)

La gestion des données de la recherche

- ANDS (Australian National Data Service), [Capability maturity:](#)
- ANDS, [Research Data Management in Practice](#)
- Clairoux Nathalie, [Introduction à la gestion des données de la recherche](#), Université de Montréal, Hiver 2016.
- Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, National Academy of Sciences, [Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age](#), Washington, DC: National Academy Press, 2009
- EUDAT (EUropean DATa), [Research data Management: An introductory Webinar from OpenAIRE and EUDAT](#), 26 May 2016 :
- Gaillard Rémi, [De l'Open data à l'Open research data : quelle\(s\) politique\(s\) pour les données de recherche ?](#), mémoire d'études sous la direction de Monique Joly pour l'obtention du diplôme de conservateur de bibliothèque, Lyon, Enssib, janvier 2014:
- Contat Odile, [Gestion des données de la recherche, quelques pistes pour démarrer, comprendre et se former](#), série de 9 billets
- RDNL (Research Data Netherlands), [Essentials 4 Data Support](#), RDNL
- University of Bristol, [Research data bootcamp](#)
- University of Oregon Libraries, "Best practices", [Research Data Management](#)

Bibliographie (3)

Guides européens et internationaux sur la gestion des données de la recherche

- European Commission, "[Horizon 2020: Work Programme update supports competitiveness through open science](#)", July 2016
- European Commission, "[Open access & Data management](#)", *Participant Portal H2020 Online Manual*
- European Commission, "[Open research data in Horizon 2020](#)"
- European Commission (Directorate-General for Research & Innovation), [H2020 Programme Guidelines on FAIR Data Management in Horizon 2020](#), Version 3.0, 26 July 2016
- European Commission (Directorate-General for Research & Innovation), [H2020 Programme Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#), Version 3.0, 26 July 2016
- OCDE, [Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics](#), Paris, 2007, 28 pages

Bibliographie (4)

Créer un plan de gestion des données

- ANDS, "[Creating a data management framework](#)", ANDS Guides:
- ANDS, "[Data Management Plans](#)", ANDS Guides
- Australian National University, "[Data Management: Data Management Plans](#)", LibGuides
- 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
- Chaudrion Stéphane, Maignant Catherine, Schöpfel Joachime, Westeel Isabelle, Les données de la recherche dans les thèses de doctorat - Livre blanc, [Rapport de recherche] Université de Lille 3, 2015. <[hal-01192930](#)>
- Inist, Tutoriels "[Le libre accès aux résultats de la recherche dans le cadre d'Horizon 2020](#)", Inist-CNRS, juin 2016
- DCC (Digital Curation Center), [Checklist for a Data Management Plan](#), version 4.0, 2013, Edinburgh: Digital Curation Centre:
- ICPSR (Interuniversity Consortium for Political and Social Research), "[Guidelines for Effective Data Management Plans](#)", Data management and Curation
- Jones Sarah, "[How to Develop a Data Management and Sharing Plan](#)", DCC How-to Guides, Edinburgh: Digital Curation Centre, 2011
- Morlock Emmanuelle, Romier Geneviève, "[Les enjeux du « Data Management Plan » pour les chercheurs](#)", Humanités numériques et données ouvertes, Université de Lyon, Lyon, 25 May 2016

Bibliographie (5)

Exemples de DMP

Deux DMP pour le programme Horizon 2020

- ESiWACE (Centre of Excellence in Simulation of Weather and Climate in Europe),
[*Data Management Plan. Deliverable 5.4*](#), 28 February 2016
- FREME (Open Framework of E-Services Multilingual and Semantic Enrichment of Digital Content),
[*Data management plan*](#), version 1, 24 July 2015

Un DMP soumis au AHRC (Arts and Humanities Research Council)

- University of Bristol Research Data Service
[*Sample AHRC Technical Plan*](#), version 2, 1 August 2013

Deux DMP en sciences sociales

- [Example plans](#) from researchers at the University of Leeds

Bibliographie (6)

Exemples de DMP

Un DMP en psychologie

- DMP “[Interference in serial verbal reactions](#)”

Des DMP provenant de l'University of California San Diego (UCSD)

- UCSD Sample NSF (National Science Foundation) [Data Management Plans](#)

Des DMP en biologie et chimie

- [Trois exemples de DMP](#)