

La donnée au cœur de la recherche

Open (research) data, Data Management Plan

Philippe Saby
10 juillet 2017



Sommaire

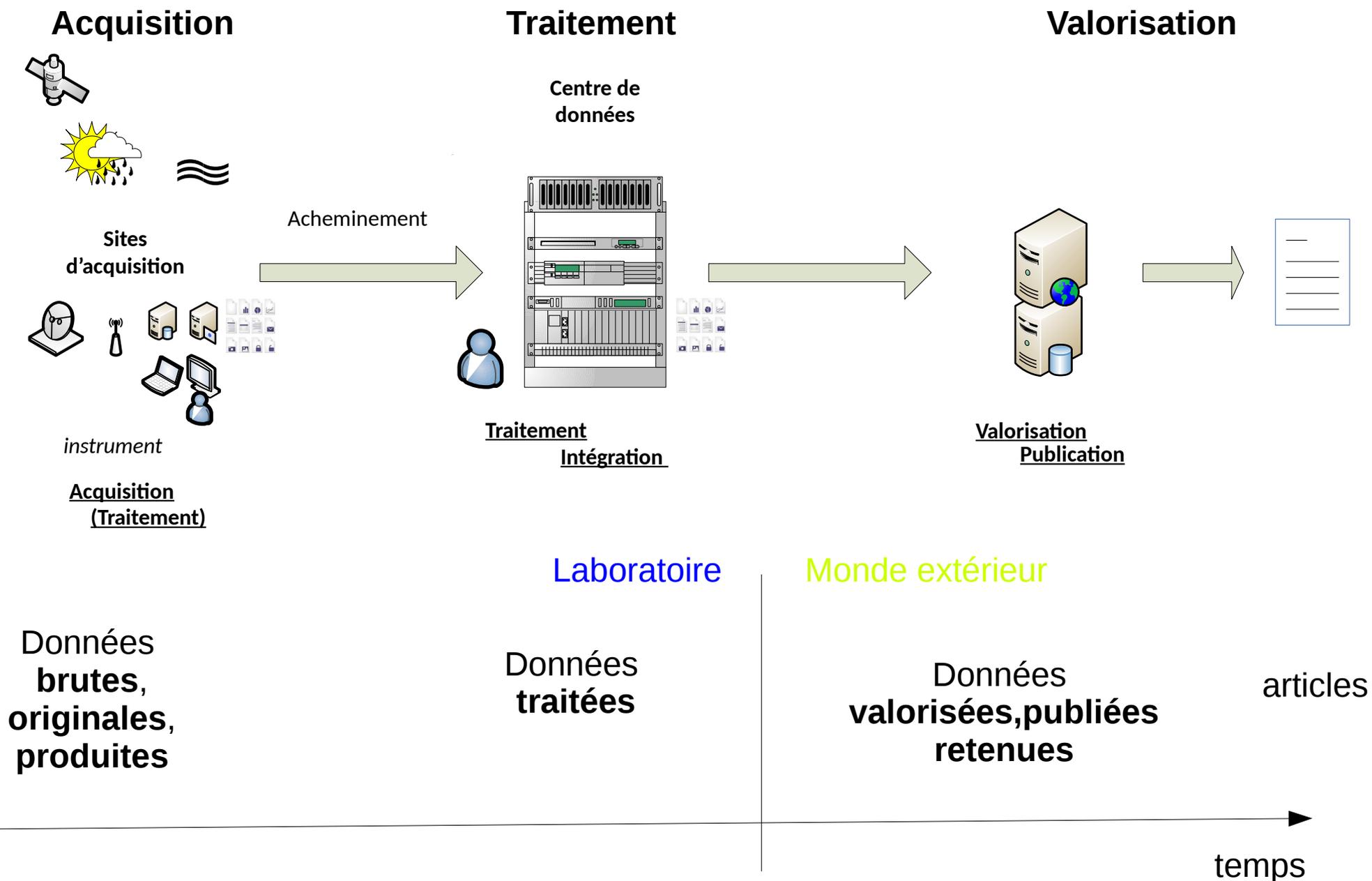
- Quoi ?
 - Exemple d'un dataflow
- Pourquoi ?
 - Notion de « databerg »
- Comment ?
 - H2020 et data management plan
- Conclusion

Quoi ?

Exemple d'un dataflow



Exemple d'un dataflow

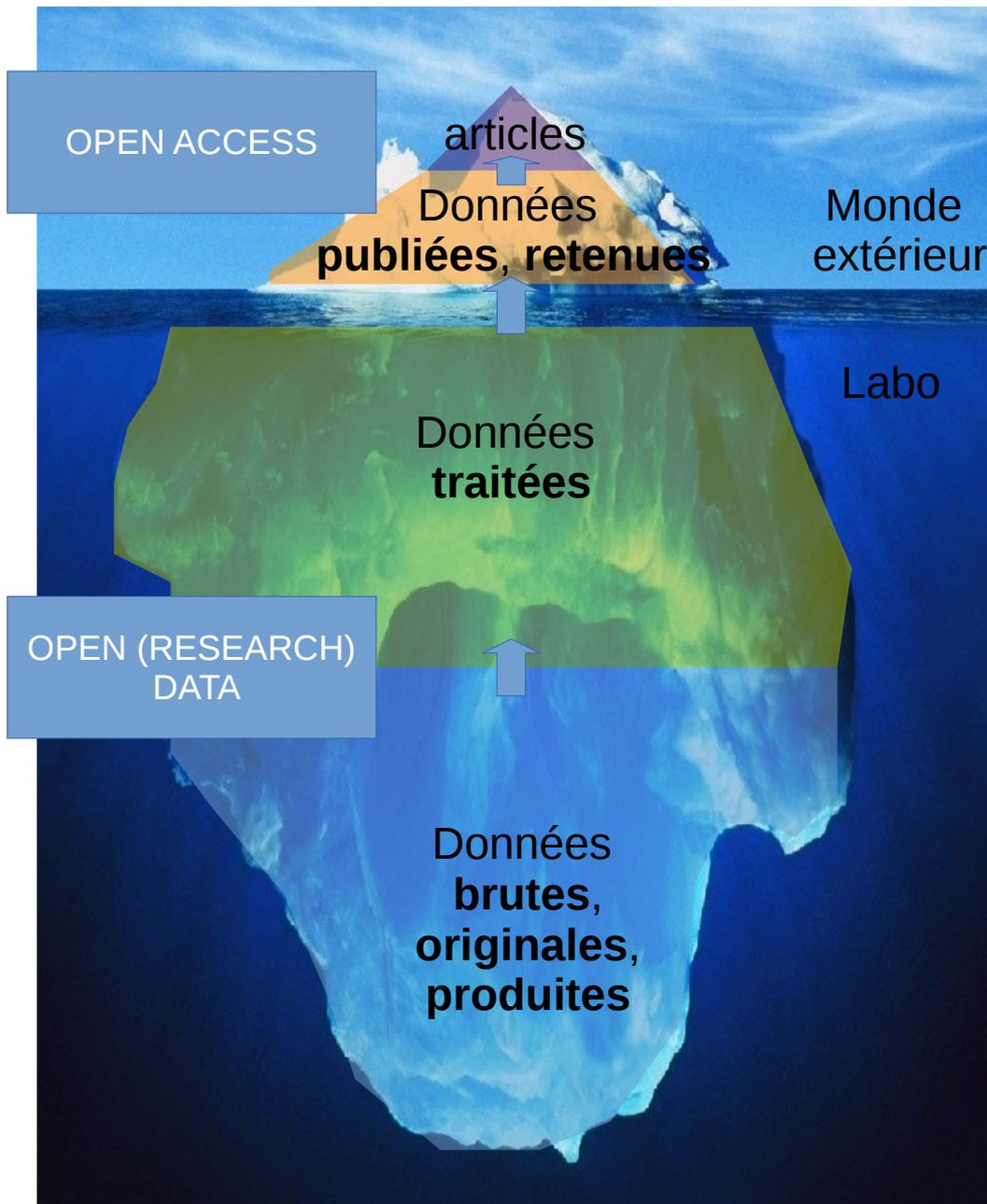


Pourquoi ?

Notion de « databerg »



Notion de « databerg »



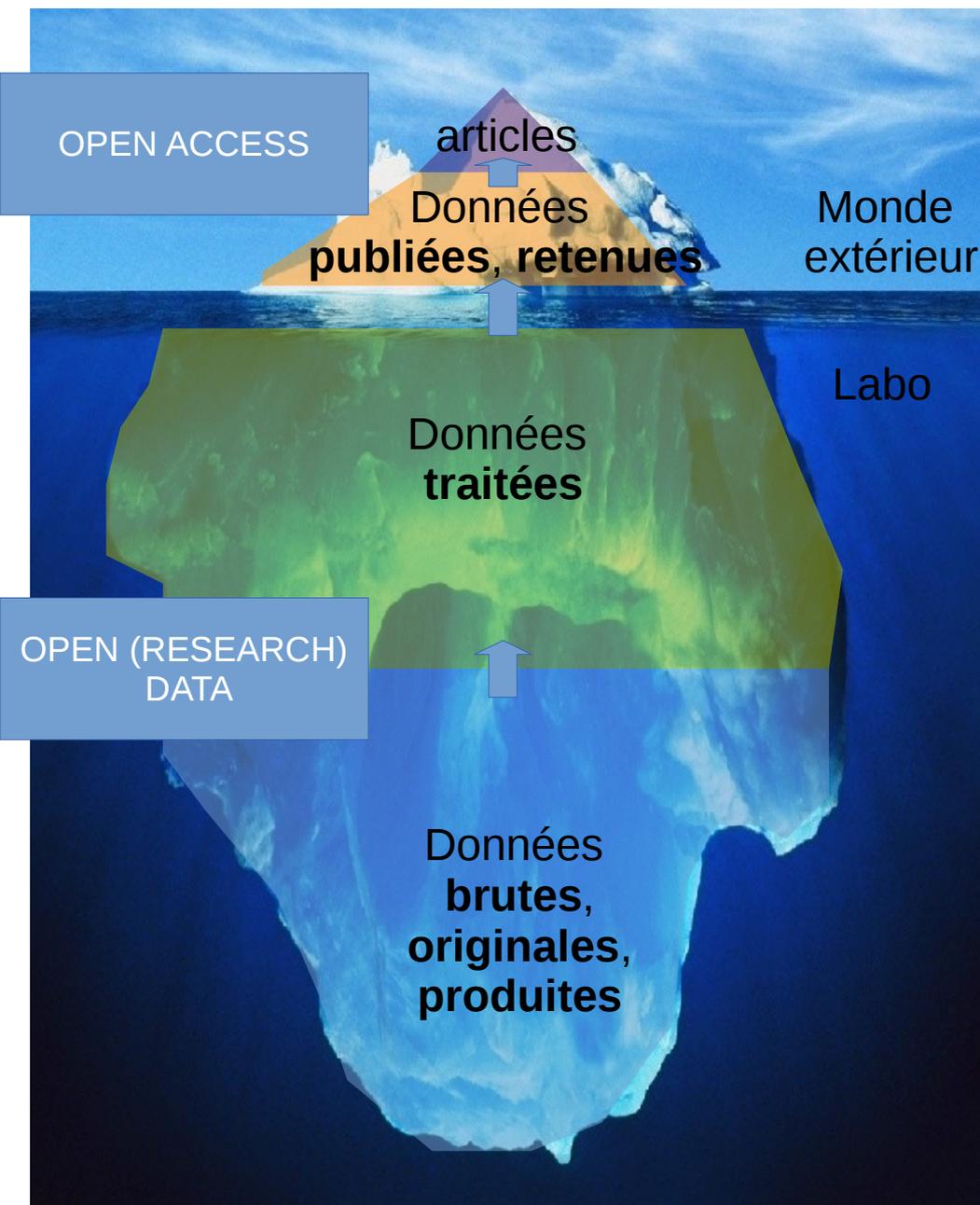
Données intégrées dans les publications

Données exploitées pour l'élaboration des résultats

Données analysées, agrégées, compilées ou transformées

Données primaires produites ou collectées au cours de la recherche

Constat



Du fait de cette phase « sous-marine »,

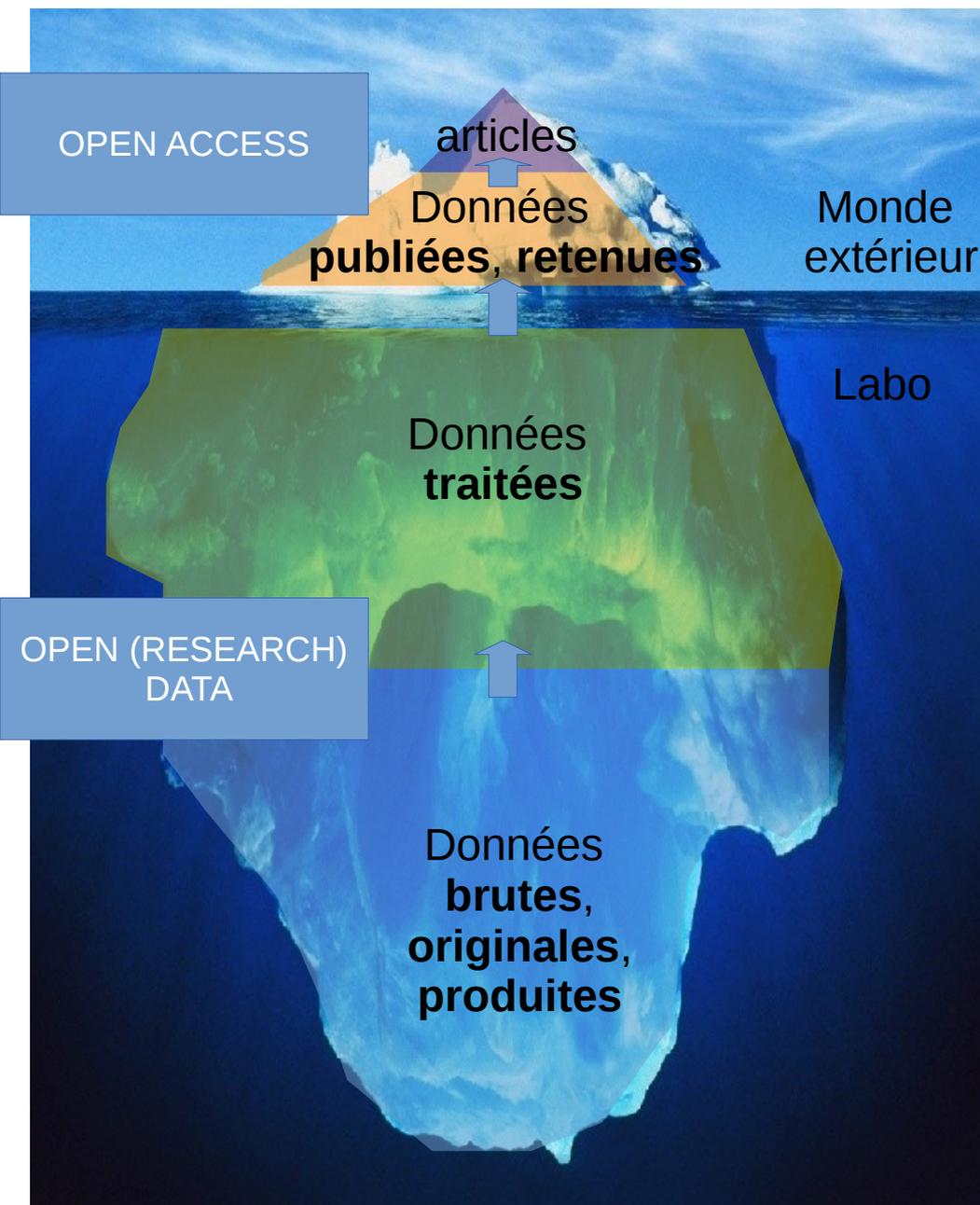
- Une part importante des données « financées » reste invisible et donc inexploitée,
- Les données ne sont pas accessibles,
- Les formats ne sont pas ouverts,
- Utilisation uniquement verticale (silot) pas d'exploitation horizontale
- Pas de droits associés

Comment ?

H2020 et Data Management Plan



Soyons « FAIR »



FAIR (janvier 2014, Lorentz workshop "Jointly design a DataFAIRport »)

•Findable

- Métadonnées descriptives riches -
- Identifiants pérennes

•Accessible

- Utilisation de protocoles libres, ouverts et standardisés pour communiquer les données - les métadonnées toujours accessibles

• Interoperable

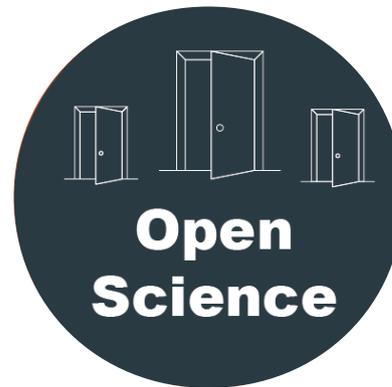
- formats ouverts, vocabulaires contrôlés & standards de métadonnées

•Re-usable

- Droits explicités et licences d'utilisation claires

Open Science Monitor

- <http://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>

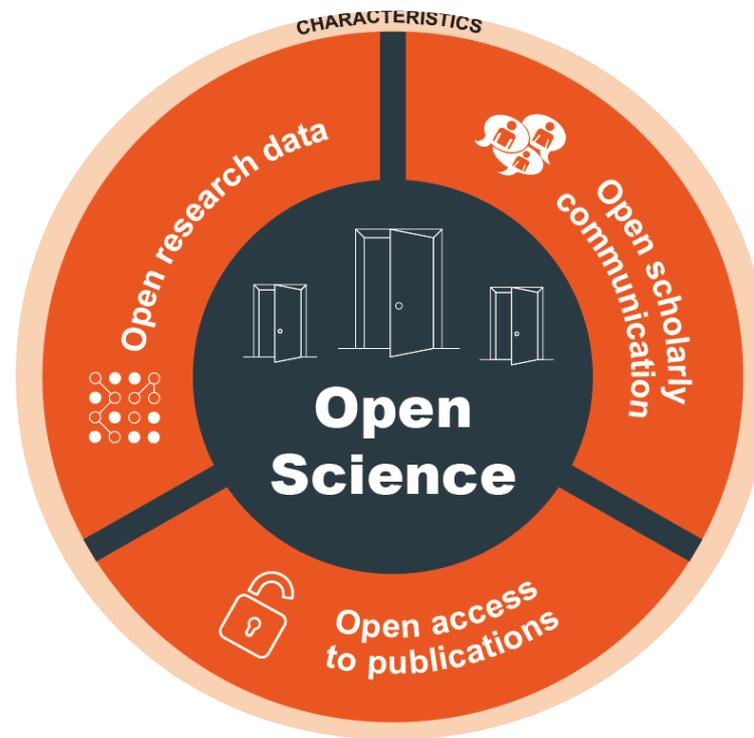


Open Science Monitor

- <http://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>

Enseignement

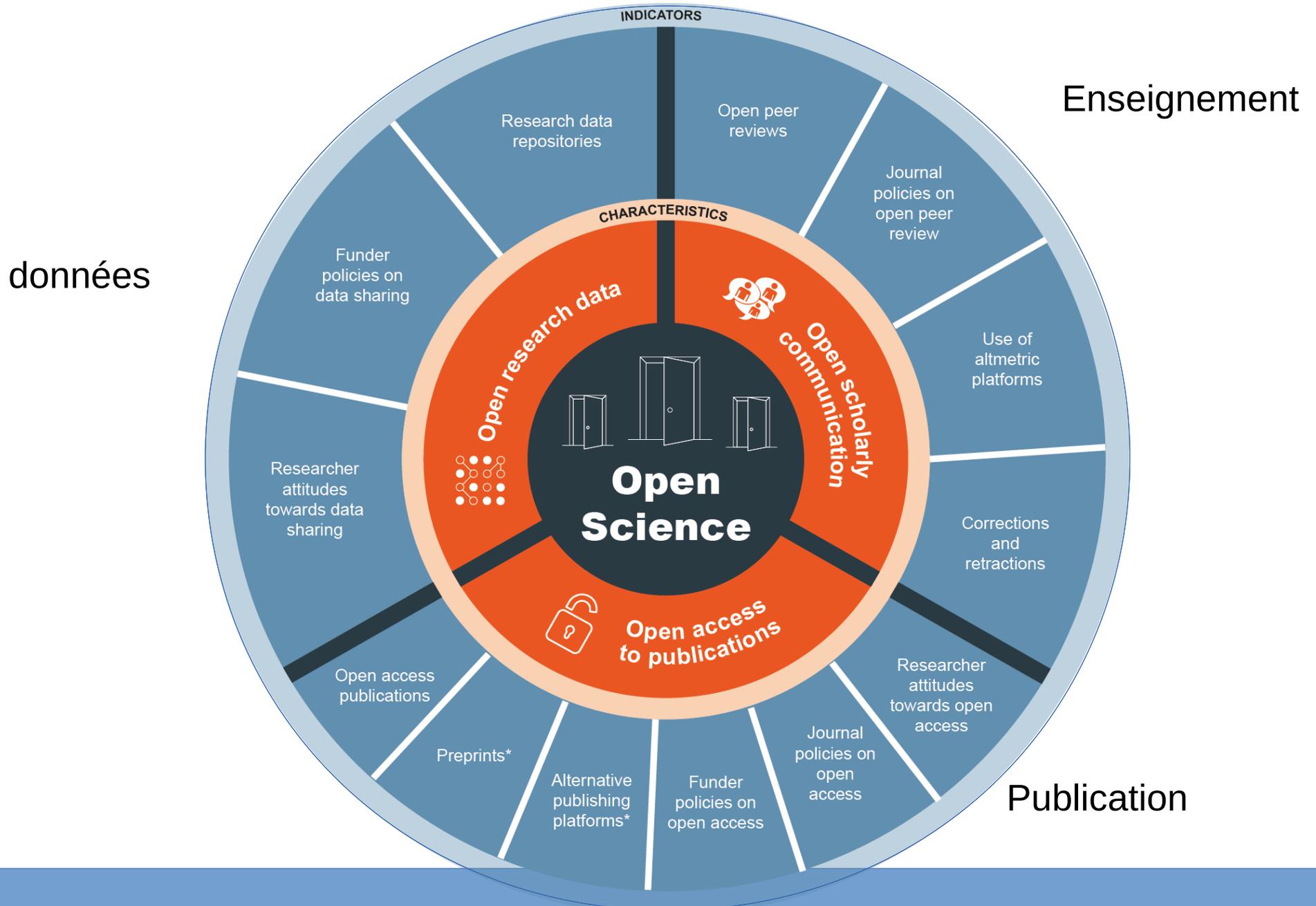
données



Publication

Open Science Monitor

- <http://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>



Data Management Plan DMP

«A 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, 15 février 2016

Commission européenne

- Document formel précisant la manière dont les données seront produites, traitées, décrites, partagées, protégées et conservées au cours et à l'issue du projet
- feuille de route, centré sur les jeux de données (datasets) du projet
- Traduction concrète de la politique des données d'un projet
- ➔ C'est un livrable du projet
- Pourquoi rédiger un DMP ?
 - Pour obtenir un financement
 - Planifier les actions et rendre mes données « FAIR »

Data Management Plan DMP

Le Data Management Plan ou Plan de gestion de données est un **document synthétique** qui **aide à organiser et anticiper toutes les étapes du cycle de vie de la donnée**. Il explique pour chaque jeu de données comment seront gérées les données d'un projet, depuis leur création ou collecte jusqu'à leur partage et leur archivage.

Une bonne pratique de recherche



De plus en plus utilisée

Le **contexte politique va en faveur de l'Open Data**. Recommandé par les financeurs, le DMP devient l'outil de gestion incontournable des projets de recherche.

Pour optimiser la gestion de vos données

Exemples

L'Europe, les Etats-Unis ou encore l'Australie ont adopté le DMP. Les projets rentrant dans le pilote du programme Horizon 2020 doivent obligatoirement en établir un.



Orientée libre accès

Le DMP est très lié au principe du libre accès aux données de recherche. En fonction de votre choix et de vos contraintes en matière de partage, des critères sont à définir.

Pour vous poser les bonnes questions

Quelles métadonnées renseigner ? Faut-il des outils spécifiques pour accéder aux données ? Quelle licence correspond à vos souhaits et/ou contraintes de partage ? Etc.

Le saviez-vous ?

Un DMP peut être établi aussi bien dans une optique de **partage** des données que pour des données en **accès restreint ou fermé**, total ou partiel. Le DMP mentionnera dans ce cas les raisons de non partage.

<http://doranum.fr/>

Data Management Plan DMP

La réalisation d'un DMP



Suivant un calendrier

Le DMP est un document évolutif. Des mises à jour et des livrables précis peuvent être définis selon les financeurs et/ou projets.

Pour respecter les échéances



Au moyen d'outils dédiés

Des outils existent pour vous aider à produire votre DMP. Ils sont gratuits après création d'un compte.

Pour vous guider

Exemples

Mise à jour lors de la production d'un nouveau jeu de données, lors du dépôt d'une demande de brevet, etc.

DMP OPIDoR, instance française de DMPonline du Digital Curation Centre (DCC)

Data Management Plan DMP



Data Management Plan DMP

Les principaux champs d'un DMP



Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires

**Propre à chaque jeu de données sauf si
valable à l'échelle du projet**

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data Management Plan DMP

Section 1 : En-tête

Objectif : identifier le document et son contenu

Principaux champs : version ; historique

Section 2 : Informations générales sur le projet

Objectif : identifier le projet

Principaux champs :

- Financier
- Référence de la convention de financement
- Mots-clés

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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



Data management plan

Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires

Data Management Plan DMP

Section 3 : Informations générales sur les données

Objectif : préciser comment les données ont été produites ou collectées, à qui elles appartiennent, comment y accéder

Principaux champs :

- Méthode de production
- Propriété
- Gestion des accès

Data management plan

Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data Management Plan DMP

Section 4 : Sécurité, gestion des risques et éthique

Objectif : Evaluer les risques et menaces qui pèsent sur les données tout au long du projet, prendre les mesures nécessaires pour leur protection.

Principaux champs :

- Risques et menaces
- Mesures de protection
- Ethique

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data management plan

Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires

Data Management Plan DMP

Section 5 : Jeux de données

Objectif : Décrire pour chaque jeu de données sa nature, ses caractéristiques techniques et ses modalités de diffusion.

Principaux champs :

- Identifiant du jeu de données
- Formats
- Standards de métadonnées

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data management plan

Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires



Data Management Plan DMP

Section 6 : Dissémination et protection

Objectif : identifier les données sensibles, leur besoin de protection, le potentiel d'ouverture des données

Principaux champs :

- Exceptions au principe de diffusion
- Embargo
- Licence
- Modalités de diffusion

Ces éléments sont décrits pour le projet dans la première version du DMP, puis par jeu de données dans les versions suivantes.

Data management plan

Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires

”as open as possible, as closed as necessary“

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data Management Plan DMP

Section 7 : Archivage des données

Objectif : Sélectionner et prévoir l'archivage à long terme des données ayant vocations à être conservées

Principaux champs :

- Sort
- Sélection
- Volume
- Durée de conservation préconisée

Data management plan

Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

Section 8 : Ressources nécessaires

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data Management Plan DMP

Section 8 : Ressources nécessaires

Objectif : contribuer au budget prévisionnel du projet en anticipant les coûts, dont ceux éligibles

Principaux champs :

- Matériel
- Personnels
- Formation
- Archivage

Data management plan



Section 1-2 : Informations générales sur le projet

Section 3 : Informations générales sur les données

Section 4 : Sécurité, gestion des risques et éthique

Section 5 : Jeux de données

Section 6 : Principes de dissémination et protection

Section 7 : Archivage des données

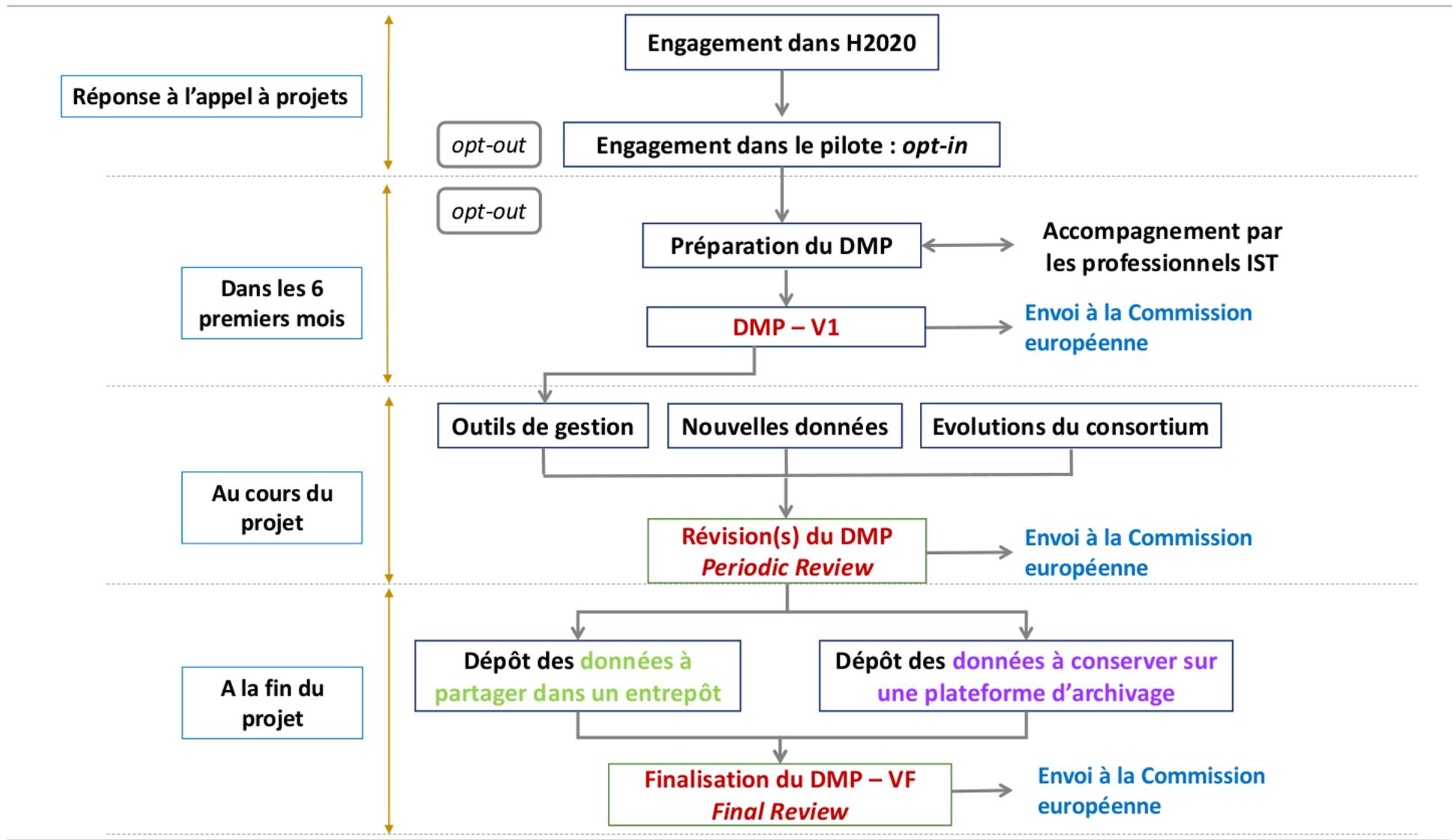
Section 8 : Ressources nécessaires

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data Management Plan DMP



Les outils pour le DMP

- DMPOnline

- Digital Curation Centre (DCC) et basé à l'origine sur les besoins des chercheurs au Royaume-Uni.
- <https://dmponline.dcc.ac.uk/>
- Plusieurs modèles en fonction des agences de financement

- DMPTool (US)

- University of California Curation Center
- <https://dmptool.org/>

Convergence des deux outils en cours

- DMP OPIDOR (INIST)

➔ Rédaction collaborative du DMP (différentes phases), export dans différents formats (en fonction du dossier de demande de financement)

Les outils pour le DMP



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

Save

Not answered yet

Data set description

B / *I* / [List icons] / [Link icon] / [Grid icon]

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 [Dropdown arrow]

Export

Aurore CARTIER - aurore.cartier@parisdescartes.fr

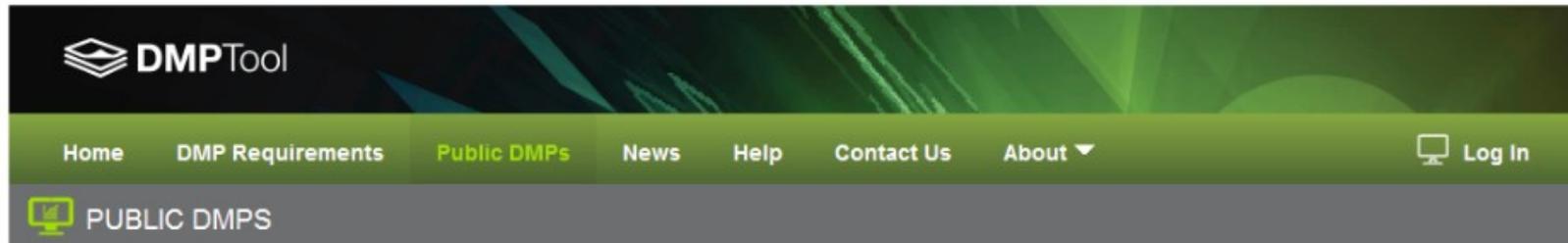
Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Les outils pour le DMP



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 Schlemmer	
Late Season Productivity, Carbon, and Nutrient Dynamics in a Changing Arctic	NSF-EAR: Earth Sciences	Oregon State University	Miguel Goni	
Language Endangerment and Revitalization at the Dialectal Level: A current analysis of Salasaca Kichwa (ISO 639-3 = qxl)	NSF-SBE: Social, Behavioral, Economic Sciences	Temple University	Robin Aronow	

[View All](#)

Aurore CARTIER - aurore.cartier@parisdescartes.fr
Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr
Nathalie REYMONET - nathalie.reymonet@univ-paris-diderot.fr

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



Les outils pour le DMP

L'outil OPIDOR de l'INIST - CNRS



Create a new plan

Please select from the following drop-downs so we can determine what questions and guidance should be displayed in your plan.

If you aren't responding to specific requirements from a funder or an institution, [select here to write a generic DMP](#) based on the Horizon 2020 FAIR DMP

If applying for funding, select your research funder.

Otherwise leave blank.



To see institutional questions and/or guidance, select

You may leave blank or select a different organisation to

- View plans
- Create plan
- About
- Help

Choose a template

There are a number of possible templates you could use:

Tick to select any other sources of guidance you wish

 Digital Curation Centre

Create plan

My plan (Horizon 2020 FAIR DMP)

Plan details	Initial DMP	Share	Export
1. Data summary	(1 question, 0 answered)		+
2. FAIR data	(4 questions, 0 answered)		+
3. Allocation of resources	(1 question, 0 answered)		+
4. Data security	(1 question, 0 answered)		+
5. Ethical aspects	(1 question, 0 answered)		+
6. Other	(1 question, 0 answered)		+

Export

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Les outils pour le DMP

My plan (Horizon 2020 FAIR DMP)



Plan details Initial DMP **Share** Export

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 an account. A notification is also issued when a user's permission level

Logged in as Magalie Moysan ▾

Collaborators

Email address

Magalie Moysan



View plans

Create plan

About

Help

My plan (Horizon 2020 FAIR DMP)

Add collaborator

Email

Permissions:

Read only ▾

Add collaborator

Plan details Initial DMP Share **Export**

From here you can download your plan in various formats. This may be useful if you need to submit your plan as part of a grant application. Select what format you wish to use and click to 'Export'.

Format

- pdf
- csv
- html
- json
- pdf
- text
- xml
- docx

Export

Aurore CARTIER - aurore.cartier@parisdescartes.fr

Magalie MOYSAN - magalie.moysan@univ-paris-diderot.fr

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

Data Management Plan DMP

- Un processus inter disciplinaires
 - Chercheur : initiateur, production, traitement et valorisation
 - IST : métadonnées, open access, DOI
 - Informaticien : infrastructure, entrepôts, archivage, sécurisation
 - Juriste : droits et licences
 - Ingénieur projet européen (IPE) : supervision, pilotage
- ➔ mettre autour de la table tous les acteurs (vision complète processus)
- Un processus qui va nécessiter des ressources
 - Humaines
 - Matérielles et logicielles (entrepôts, archivage, etc.)
 - Financières (maintenance infrastructures, formation, fonctionnement)
- ➔ obtenir des financements pérennes (20 ans) pour un projet, par nature, fini

Conclusion

- Nous entrons dans une période « ouverte »
 - Open science, open access, open data, open research data
 - Pas d'ouverture = moins (pas) de financement
- Aujourd'hui, le DMP est facultatif car réalisé après l'acceptation du projet
- Demain, il sera probablement obligatoire et intégré à la demande initiale
- Nous devons donc :
 - nous préparer pour la rédaction,
 - former une (des) équipe(s) inter disciplinaires (IST, chercheur, informaticien etc.) en charge de cette thématique,
 - identifier (créer, rendre visible) des entrepôts de données et des centre d'archivage reconnus pour la conservation de nos données.

Objectif : anticiper aujourd'hui pour avoir, demain, des structures visibles et pérennes au niveau régional, national mais aussi international

Références

- Open Science Monitor

- <http://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>

- Documentations / formations

- <http://doranum.fr/> apprentissage numérique sur les données

- **Aurore CARTIER** - aurore.cartier@parisdescartes.fr

- Magalie MOYSAN** - magalie.moysan@univ-paris-diderot.fr

- Nathalie REYMONET** - nathalie.reymonet@univ-paris-diderot.fr

- <https://appui-recherche.univ-paris-diderot.fr/data-management-plan-dmp-guide-de-redaction>

- Michel roland (unice)

- http://wiki-urfist.unice.fr/wiki_urfist/index.php/Open_Science_et_donn%C3%A9es_de_la_recherche

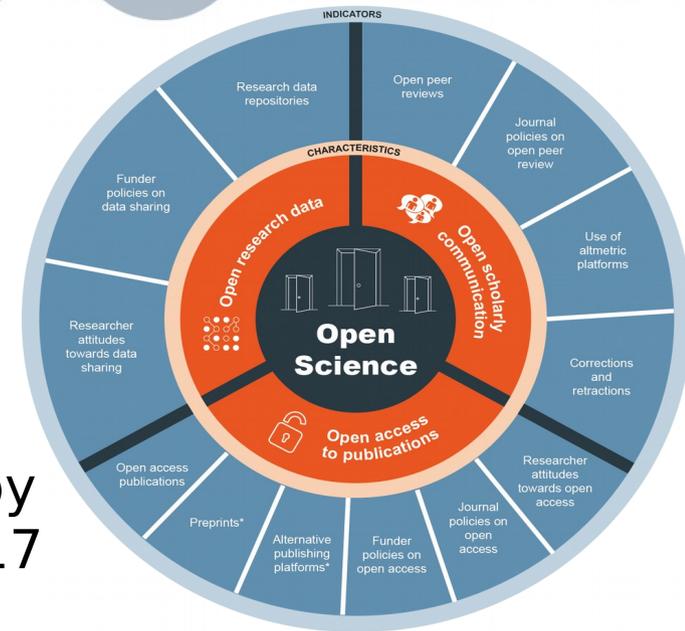
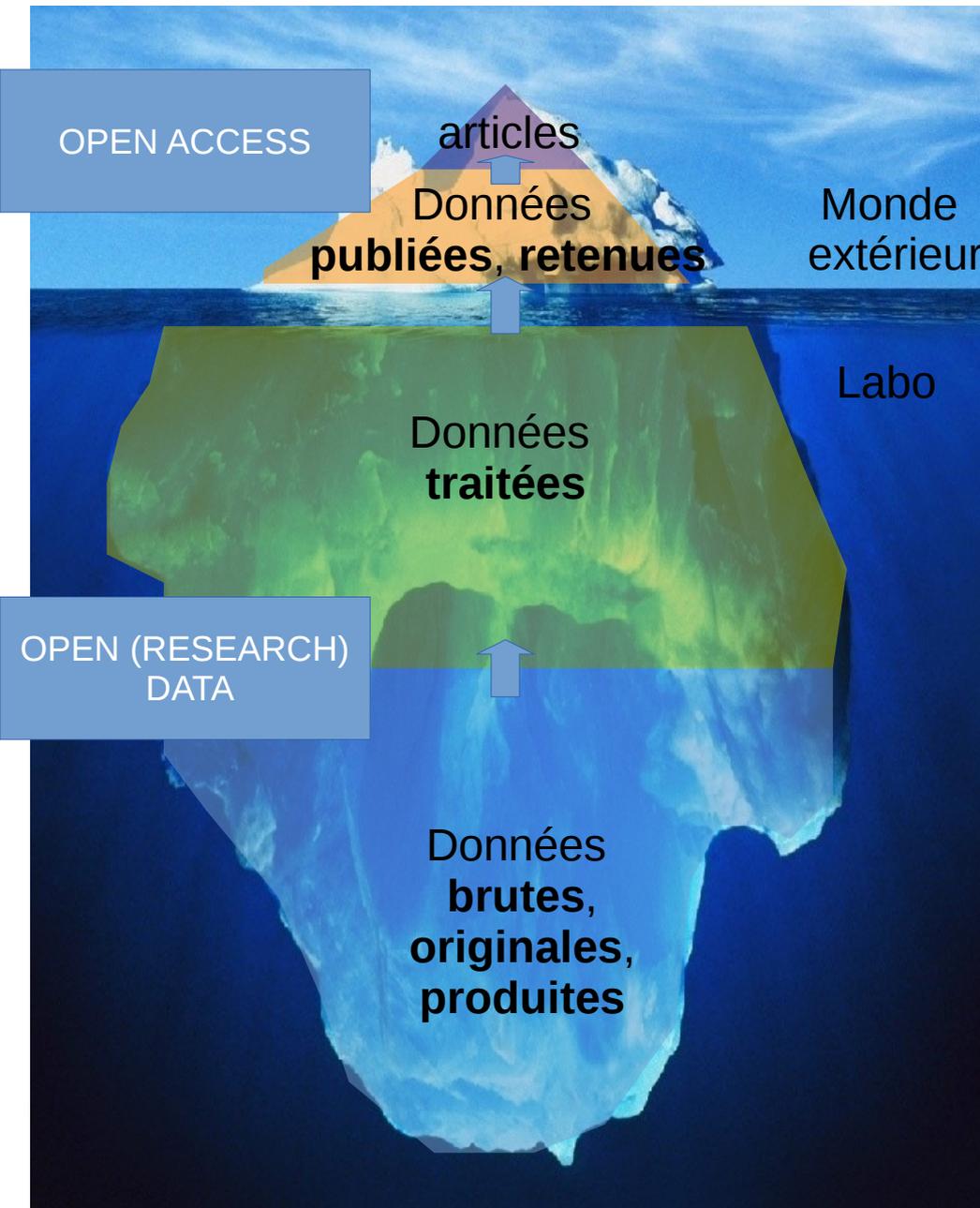
- Outil DMP

- <http://www.cdlib.org/>

- <http://www.dcc.ac.uk>

- <https://dmp.opidor.fr>

Merci de votre attention



Philippe Saby
10 juillet 2017