

# Les enjeux du « Data Management Plan » pour les chercheurs

Emmanuelle Morlock IE CNRS, Laboratoire HiSoMA (UMR 5189)  
Geneviève Romier, IR CNRS, Institut des grilles et du Cloud



[www.cnrs.fr](http://www.cnrs.fr)





# Objectifs

---

1. Définition, enjeux et utilité du DMP
2. Outils disponibles
3. Exercice



[www.cnrs.fr](http://www.cnrs.fr)

Partie I

# Définition, enjeux et utilité du DMP





# Définition, enjeux & utilité

« **Data Management Plan** » (DMP)  
=  
« **Plan de gestion de données** » (PGD)

⊙ Une définition synthétique :

○ « *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

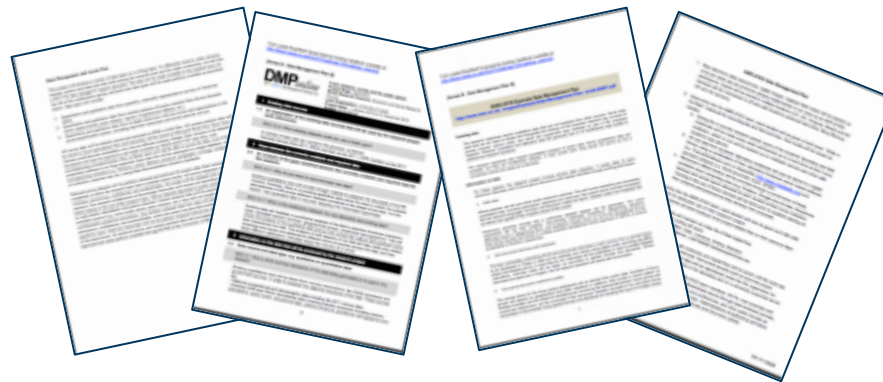
- Le DMP est un **document formalisé** ;
- Son contenu peut varier selon les besoins des projets et les demandes des agences de financement ;
- C'est avant tout une **feuille de route**, en ajustement constant ;
- Le DMP est centré sur les données et les **jeux de données (datasets)** du projet de recherche.

# Un document de politique scientifique

Le DMP n'est ni « technique », ni « informatique ».

Le DMP d'un projet de recherche, c'est :

- ⦿ la traduction concrète de la politique des données d'un projet ;
- ⦿ une récapitulation synthétique des objectifs et des actions qui vont être mises en œuvre ;
- ⦿ en regard d'une demande de financement.

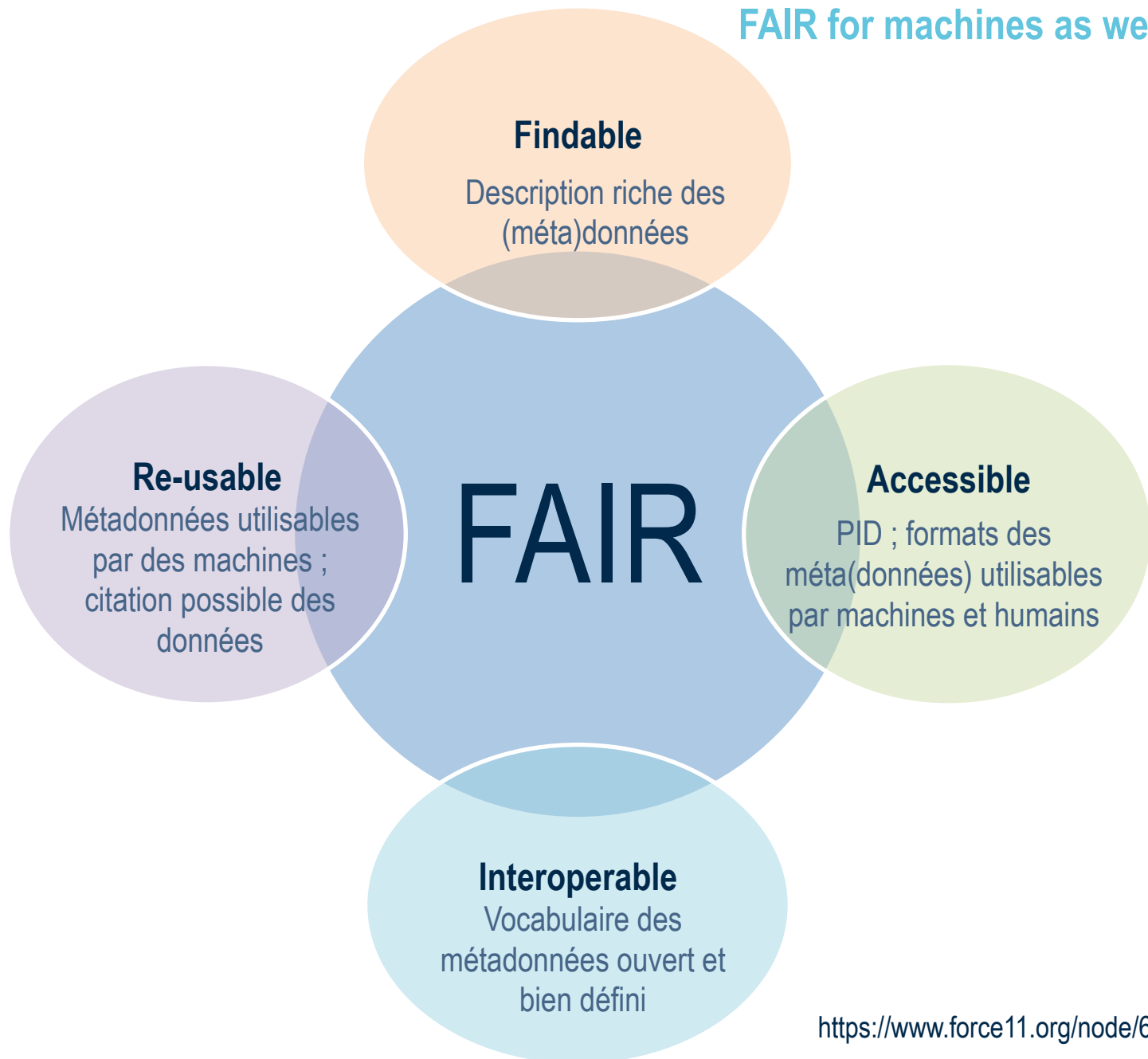


➤ Le DMP est un **livrable** du projet.

# Pourquoi rédiger un DMP ?

- Les données numériques requièrent une « gestion active » :
  - Entretien continu (sauvegarde, migration, conversions...) tout au long du cycle de vie ;
  - Planification des actions (enjeux de qualité, de faisabilité technique, de viabilité économique...)
  - Notion de « [digital curation](#) » (au sens anglo-saxon).
- Objectif : [FAIR](#) (janvier 2014, Lorentz workshop "Jointly design a Data FAIRport")
  - F**indable : Métadonnées descriptives riches - Identifiants pérennes
  - A**ccessible : Utilisation de protocoles libres, ouverts et standardisés pour communiquer les données – les métadonnées toujours accessibles
  - I**nteroperable : formats ouverts, vocabulaires contrôlés & standards de métadonnées
  - R**e-usable : Droits explicités et licences d'utilisation claires





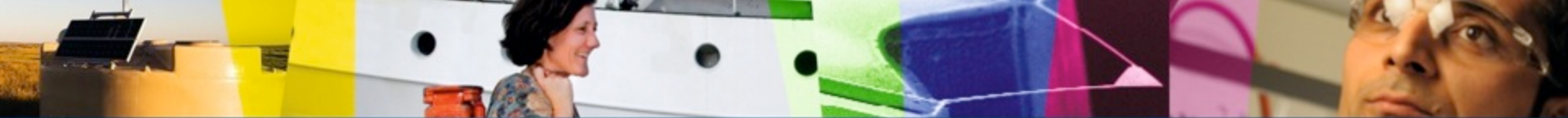
<https://www.force11.org/node/6062/#Annex1>

# Rôle clé du « M » dans DMP

- ⊙ Dans la perspective du Data Management Plan, la gestion des données, ce n'est pas la simple « intendance » des données (*Datastewardship*), mais plutôt un véritable « management », c'est-à-dire l'optimisation de ressources en vue d'un objectif.
- ⊙ Il requiert :
  - Beaucoup de « bon sens »
    - *Par ex. choix du nommage, sauvegarde...*
  - Des choix stratégiques forts
    - *Par ex. les durées de conservation*
  - La mise en cohérence avec les enjeux scientifiques et l'impact du projet
- ⊙ Il permet :
  - d'identifier et de rendre visibles les actions à mener ;
  - de planifier (voir les étapes clés, les échéances, les périodes critiques... opération par opération).







## « Data storage » vs « Data curation »

---



*Data storage is confined to simply keeping data in existence and ensuring that it can be accessed when needed. It does not necessarily entail practices of refreshment or format migration (essential to maintaining the data in a usable form) nor does it entail higher-level curatorial practices such as enhancement of the data through added metadata, or migration from one representational standard to another. **Data curation** thus goes far beyond the scope of data storage.*

<https://guide.dhcurator.org/faq/>



# Exemples

---

1. Responsabilité
2. Gestion des résultats
3. Description structurée
4. Plan de sauvegarde
5. Propriété intellectuelle
6. Devenir des données dans *l'après-projet*

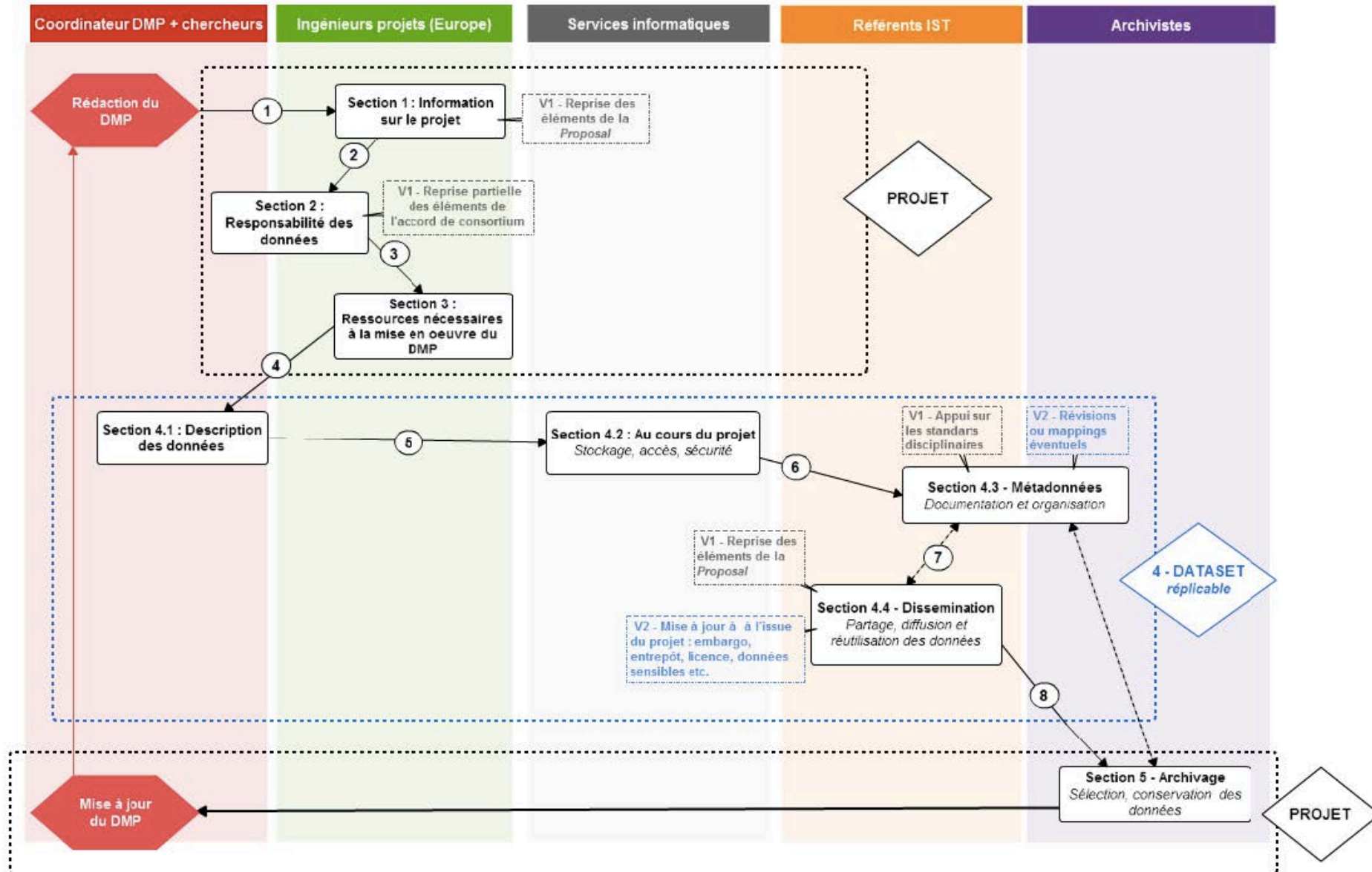
# Exemple 1 : Responsabilité

- ⊙ Qui fait quoi ?
- ⊙ À quel moment ?
- ⊙ Sous la responsabilité de qui ?
  
- ⊙ **Chercheur** :
  - *Responsable des données, coordinateur du DMP*
  - *Description des données et découpage des jeux de données*
- ⊙ **Ingénieur-projet** : coordonne les actions autour du DMP
- ⊙ **Informaticien** : interlocuteur pour le stockage et sécurisation des données, les aspects infrastructure 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



Source du schéma :

A.CARTIER, M.MOYSAN, N.REYMONET. Réaliser un plan de gestion de données : guide de rédaction (V1, 09/01/2015) p. 11



# Exemple 2 : Gestion des jeux de données

- ① Dans quel état et sous quelle forme de classement va-t-on gérer les jeux de données produits par la collecte, la production ou les analyses ?
- ① Exemples d'états problématiques à éviter :
  - **Noms de fichiers**
    - Incompréhensibles
    - Suivent des règles de nommage variables ou inexistantes
    - Contiennent des caractères qui ne sont pas acceptés par tous les systèmes informatiques
  - **Fichiers**
    - Stockés à divers endroits
    - En plusieurs formats et versions
    - Dans des formats propriétaires...





# Exemple 3 : Descriptions et métadonnées

- ① Les données sont-elles compréhensibles pour des chercheurs ne faisant pas partie de l'équipe qui les a produites ?
- ① Quels vocabulaires sont utilisés pour décrire les données ? (métadonnées)
  - Y a-t-il utilisation de standards partagés ? Ou s'agit-il de vocabulaires « maison » ?
  - Les termes sont-ils définis ? Les conventions internes au projet pour la description et d'indexation des données sont-elles documentées elles aussi ?
- ① Les informations qui permettront à un utilisateur potentiel de découvrir le jeu de données sont-elles exprimées explicitement ?
- ① Un dictionnaire des données est-il fourni ?



# Exemple 4 : Plan de sauvegarde



- ⦿ À quel endroit ?
- ⦿ À quelle fréquence ?
- ⦿ Combien de copies ?
- ⦿ Quel espace serveur ?
- ⦿ Quels modalités de cryptage des copies (si besoin)?
- ⦿ Coût(s) ?
- ⦿ Comment restaurer les données en cas de besoin ?

# Exemple 5 : Propriété intellectuelle

- ⦿ Qui sont les détenteurs des droits de la PI ? Et pour quelles données ?
- ⦿ Comment prendre en compte la politique du financeur ?
- ⦿ Quelles obligations sur les durées de conservation découlent des finalités des données ?
  
- ⦿ Exemples :
  - **Données de brevets** = à conserver indéfiniment
  - **Données sous-jacentes** à un article scientifique
    - « Une mention particulière doit être apportée aux données venant en appui des publications et qui doivent à ce titre rester disponibles et accessibles le plus largement possible. » cf. le schéma d'orientation stratégique de l'IST du CNRS (2013 p. 37)



# Exemple 6 : Devenir dans l'après-projet

- ① Qu'arrivera-t-il aux données quand le projet sera terminé ?
  - Quelles données doivent être détruites ? (contraintes légales...)
  - Quelles données doit-on conserver ? (publication associée...)
  - Quelles données est-il souhaitable de conserver ?
  - Quelles données doit-on partager ? (contraintes du financeur...)
  - Quelles données peut-on partager et quand ? (embargo ?...)
- ① Quelle valeur ont mes données ?
  - Pour qui ?
  - Comment l'évaluer ?
- ① Quelles sont les options pour la publication des données ?
- ① Quelles sont les options pour l'archivage ?
- ① Quels sont les coûts associés ?



# Le contenu du DMP en bref

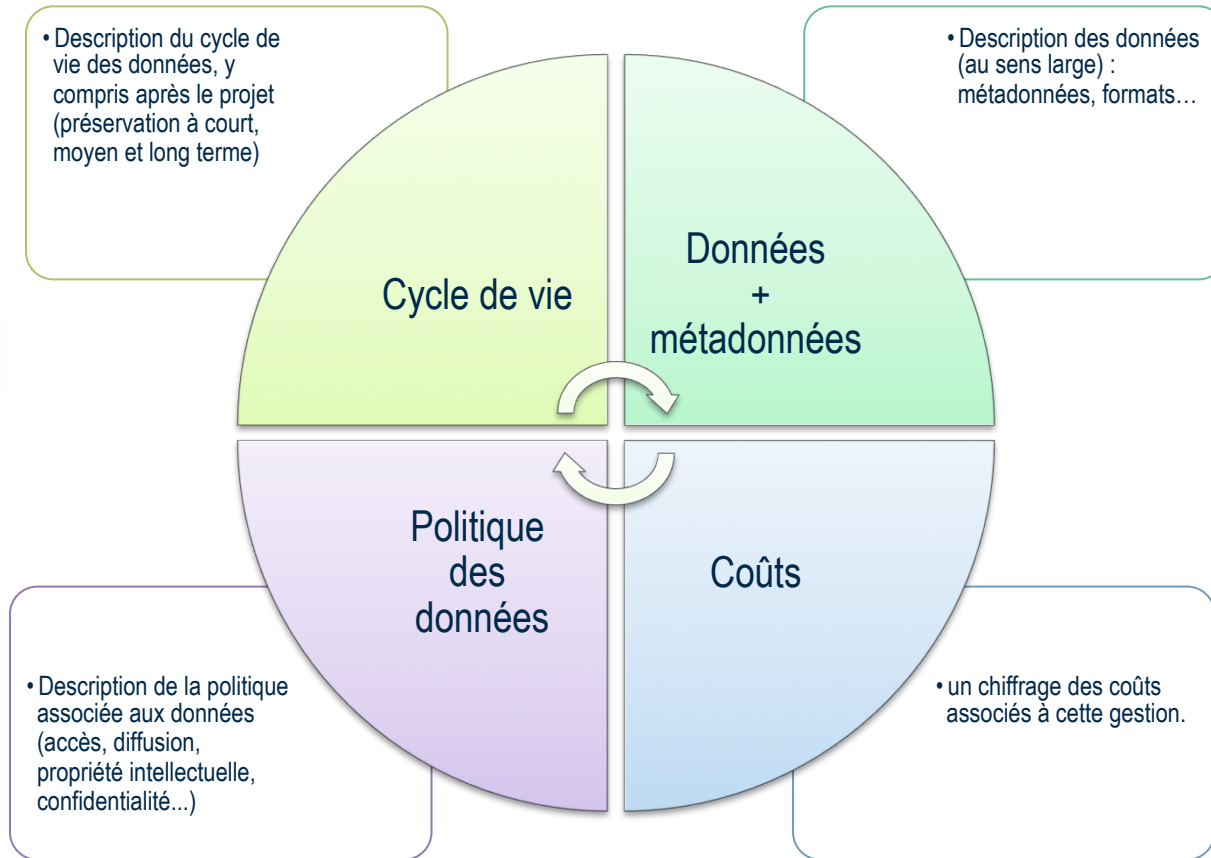
- ⦿ Faire un DMP, c'est définir comment les données du projet seront :



- ⦿ Un DMP contient :
  - une description du **cycle de vie** des données, y compris après le projet (*préservation long terme*) ;
  - une description des **données** (au sens large) : *métadonnées, formats...* ;
  - une description de la **politique associée aux données** (accès, diffusion, propriété intellectuelle, confidentialité...);
  - un **chiffrage des coûts** associés à cette gestion.
- *Le DMP participe à la **sécurisation** et à la **pérennisation** des données.*



# Le contenu du DMP en bref



# Quand et pourquoi ?

## ⦿ Quand ?

- Le DMP doit être préparé avant que les premières données ne soient collectées, produites et mises à jour (souvent un "draft" dans la proposition).
- Le DMP est mis à jour régulièrement.

## ⦿ Pourquoi ?

- Le DMP est plus en plus nécessaire :
  - demandé par les agences de financement de la recherche dès la proposition et dans le suivi et l'évaluation des projets (cf. le **critère d'impact global** des projets ANR).
- C'est un livrable du projet pour de nombreux financements.

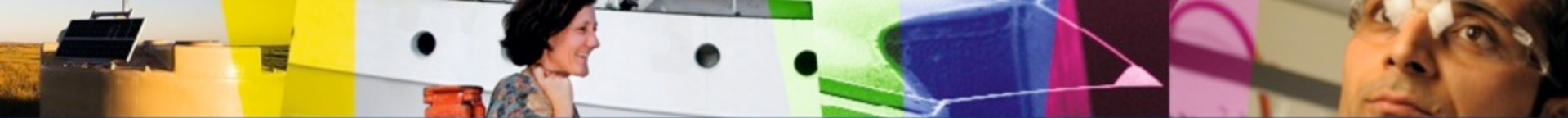
## ⦿ Qui ?

- Les responsables de projets scientifiques

### ➤ Remarque

*Les grandes communautés internationales ont ouvert la voie : Infrastructures de recherche, projets internationaux structurés autour d'instruments ou communautés produisant et utilisant de grandes quantités de données.*





## 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](http://www.univ-paris-diderot.fr/DocumentsFCK/recherche/Realiser_un_DMP_V1.pdf)



# Exemples de cadres de DMP



## Appels à projets internationaux

- ✓ Horizon 2020
- ✓ T-AP « Digging into Data Challenge »
- ✓ Démarche similaire aux Etats-Unis (NSF) et dans différents pays d'Europe (Pays-Bas, Royaume-Uni...).

# L'exemple de H2020 : un peu d'histoire



COMMISSION EUROPÉENNE

Bruxelles, le 17.7.2012  
COM(2012) 401 final

**COMMUNICATION DE LA COMMISSION AU PARLEMENT EUROPÉEN, AU  
CONSEIL, AU COMITÉ ÉCONOMIQUE ET SOCIAL EUROPÉEN ET AU COMITÉ  
DES RÉGIONS**

**Pour un meilleur accès aux informations scientifiques:  
dynamiser les avantages des investissements publics dans le domaine de la recherche**

➤ Annonce les mesures prises dans H2020



# Les objectifs de l'Union européenne

- ⦿ Accélérer l'innovation, permettre et accélérer la mise sur le marché des résultats de la recherche ;
- ⦿ Permettre de nouvelles formes de recherche sur les données (*data-intensive research*) ;
- ⦿ Inciter aux collaborations, éviter les doublons, améliorer l'efficacité ;
- ⦿ Améliorer la qualité et la transparence ;
- ⦿ Améliorer le retour sur investissement de la R&D.



# Horizon 2020

① <http://www.horizon2020.gouv.fr/>

① Programme-cadre européen actuel de financement de la recherche et de l'innovation

① 79 milliards d'euros sur la période 2014-2020.

① Appels à projets (*Calls*) à destination des scientifiques et industriels de l'Union européenne.

① Contient des dispositifs pour favoriser :

- l'accès aux résultats des projets, aux publications (**Open Access**) ;
- la ré-utilisation des données de recherche (**Open Data**) ;
- la publication des codes permettant la production et/ou le traitement de ces données (**Open Source**).



# Horizon 2020, de l'action pilote « Open Research Data » à la systématisation

- ⊙ 2014-2015 : « Open Research Data » (**ORD Pilot**) systématique pour certains programmes et projets volontaires.
- ⊙ 2016-2017 : idem
  - liste des programmes concernés (cf. transparent suivant)
- ⊙ On peut se soustraire à l'ORD sous conditions (liste précise, ex. données personnelles)
- ⊙ Le DMP est annoncé comme obligatoire dans H2020 par la suite !
- ⊙ Participer à l'Open Research Data Pilot n'implique pas d'ouvrir toutes les données. L'objectif du pilot est d'encourager le bon management des données comme une bonne pratique de recherche.
- L'approche de la Commission peut être résumée ainsi :  
« *as open as possible, as closed as necessary* ».

## Pour en savoir plus :

[http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm)



# Liste des programmes 2016-2017 concernés

- **Future & Emerging Technologies**
- **Research infrastructures**
- **Leadership in enabling & industrial technologies** – Information & Communication Technologies
- **Nanotechnologies, Advanced Materials, Advanced Manufacturing & Processing, & Biotechnology** – 'nanosafety' & 'modelling' topics
- **Societal Challenge – Food security**, sustainable agriculture & forestry, marine & maritime & inland water research & the bioeconomy - selected topics as specified in the work programme
- **Societal Challenge – Climate Action**, Environment, Resource Efficiency & Raw Materials – except raw materials
- **Societal Challenge – Europe in a changing world** – inclusive, innovative & reflective societies
- **Science with & for Society**
- **Cross-cutting activities** – focus areas – part Smart & Sustainable Cities

cf.

[http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm)



# Horizon 2020 : les plus

- ⦿ Les coûts associés au DMP sont éligibles dans le cadre et pendant la durée du projet.
- ⦿ Des Guidelines pour vous aider !
  - « *These guidelines are intended to help applicants and beneficiaries of projects under Horizon 2020 meet their responsibilities as regards research data quality, sharing and security.* »
- ⦿ Modèle proposé :  
[http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm#A1-template](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template)
  - cf. transparents suivants
- ⦿ Outil proposé : <https://dmponline.dcc.ac.uk/>
  - cf. 3<sup>ème</sup> partie de l'atelier





# Annex 1: Data Management Plan template 1/2

The purpose of the Data Management Plan (DMP) is to provide an analysis of the main elements of the data management policy that applicants will apply to all datasets that will be generated by the project.

The DMP is not a fixed document, but evolves during the lifespan of the project.

Address the points below on a dataset-by-dataset basis in your DMP and reflect the current status of reflection within the consortium about the data that will be produced.



- ① **Dataset reference & name**

- Identifier for the dataset to be produced.

- ② **Dataset description**

- Description of the data that will be generated or collected, its origin (if collected), nature & scale and to whom it could be useful, and whether it underpins a scientific publication. Information on the existence (or not) of similar data and the possibilities for integration & reuse.

- ③ **Standards & metadata**

- Reference to existing suitable standards of the discipline. If these do not exist, an outline on how and what metadata will be created.

Source : [http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm#A1-template](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template)

# Annex 1: Data Management Plan template 2/2

## ⦿ Data sharing

- Description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination & necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).
- If the dataset cannot be shared, give the reasons why (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related)

## ⦿ Archiving & preservation (including storage & backup)

- Description of the procedures that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered.



# Annex 2: Additional guidance for Data Management Plans 1/2

This can be applied to any project that produces, collects or processes research data, and is included here as reference for elaborating DMPs in Horizon 2020 projects. This guide is structured as a series of questions that should be ideally clarified for all datasets produced in the project.

Scientific research data should be easily:

## 🕒 Discoverable

- **DMP question:** are the data and associated software produced and/or used in the project discoverable (and readily located), identifiable by means of a standard identification mechanism (e.g. Digital Object Identifier)?

## 🕒 Accessible

- **DMP question:** are the data and associated software produced and/or used in the project accessible and in what modalities, scope, licenses (e.g. licensing framework for research and education, embargo periods, commercial exploitation, etc.)?

## 🕒 Assessable & intelligible

- **DMP question:** are the data and associated software produced and/or used in the project assessable for and intelligible to third parties in contexts such as scientific scrutiny and peer review (e.g. are the minimal datasets handled together with scientific papers for the purpose of peer review, are data provided in a way that judgments can be made about their reliability and the competence of those who created them)?



# Annex 2: Additional guidance for Data Management Plans 2/2

- ① **Useable beyond the original purpose for which it was collected**
  - **DMP question:** are the data and associated software produced and/or used in the project useable by third parties even long time after the collection of the data (e.g. is the data safely stored in certified repositories for long term preservation and curation; is it stored together with the minimum software, metadata and documentation to make it useful; is the data useful for the wider public needs and usable for the likely purposes of non-specialists)?
- ② **Interoperable to specific quality standards**
  - **DMP question:** are the data and associated software produced and/or used in the project interoperable allowing data exchange between researchers, institutions, organisations, countries, etc. (e.g. adhering to standards for data annotation, data exchange, compliant with available software applications, and allowing re-combinations with different datasets from different origins)?



# Références et liens proposés

## 🎯 Références

- [Guidelines on Data Management in Horizon 2020](#)
- [Annotated Model Grant Agreement - Open access to research data](#)

## 🎯 Liens

- [OpenAIRE - Open Access Infrastructure for Research in Europe](#)
- [Digital Curation Centre's DMP Online tool](#)



# Les projets H2020 SHS sont concernés quelques exemples

⦿ [http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/ssh\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/ssh_en.htm)

⦿ (Societal Challenge – Europe in a changing world – inclusive, innovative & reflective societies)

⦿ <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/3083-cult-coop-07-2017.html>

ou

⦿ <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/3087-cult-coop-09-2017.html>







# RESEARCH & INNOVATION

## Participant Portal

European Commission > Research & Innovation > Participant Portal > Opportunities

HOME FUNDING OPPORTUNITIES HOW TO PARTICIPATE EXPERTS SUPPORT Search PP LOGIN REGISTER

**EU Programmes 2014-2020**

Search Topics

Updates  

Calls  

H2020

Research Fund for Coal & Steel

3rd Health Programme

Promotion of Agricultural Products

Consumer Programme

COSME


Justice Programme

FP7 & CIP Programmes 2007-2013

### TOPIC : Cultural heritage of European coastal and maritime regions

**Topic identifier:** CULT-COOP-07-2017  
**Publication date:** 14 October 2015

**Types of action:** RIA Research and Innovation action  
**DeadlineModel:** single-stage  
**Planned opening date:** 04 October 2016  
**Deadline:** 02 February 2017 17:00:00  
 Time Zone : (Brussels time)

 **Horizon 2020**  
 Pillar: Societal Challenges  
 Work Programme Year: 2016-2017  
 Work Programme Part: Europe in a changing world - inclusive, innovative and reflective Societies  
 Call : H2020-SC6-CULT-COOP-2016-2017

#### Topic Description

##### Specific Challenge:

This RIA complements previous RIA in rural, coastal and urban settings. European

### TOPIC : European cultural heritage, access and analysis for a richer interpretation of the past.

**Topic identifier:** CULT-COOP-09-2017  
**Publication date:** 14 October 2015

**Types of action:** RIA Research and Innovation action  
**DeadlineModel:** single-stage  
**Planned opening date:** 04 October 2016  
**Deadline:** 02 February 2017 17:00:00  
 Time Zone : (Brussels time)



**Horizon 2020**  
 Pillar: Societal Challenges  
 Work Programme Year: H2020-2016-2017  
 Work Programme Part: Europe in a changing world - inclusive, innovative and reflective Societies  
 Call : H2020-SC6-CULT-COOP-2016-2017

[H2020 website](#)

[Call budget overview](#)

#### Topic Description

+ More

##### Specific Challenge:

Collections in archives, museums, and at cultural heritage sites contain a wealth of digital texts, images, audio-visual content and 3D representations of objects or scenes as well as other

# DMP & H2020 en résumé

- ① Le DMP systématique pour les programmes du Pilot « Open Research Data » (ORD Pilot) et les projets volontaires.
- ① Les coûts associés au DMP sont éligibles dans le cadre et pendant la durée du projet.
- ① Un modèle de DMP et des guidelines sont proposés.
- ① Un outil est proposé (avec modèle intégré).
- ① Le DMP est un livrable du projet : le document de soumission du projet de recherche ne contient qu'un *draft*, qui sera ensuite développé et mis à jour en continu.



# Autre contexte, autre exemple : T-AP Digging into Data Challenge



<http://diggingintodata.org/>

- Dans le cadre de cet appel est requis un plan de gestion des données INCLUANT les codes (informatiques) développés qui **doivent** être publiés :
  - "Plan de gestion des données [de deux pages au maximum] : les candidats doivent préparer conjointement un plan de gestion des données pour leur projet. Ce plan doit décrire comment l'équipe du projet gèrera et diffusera les données recueillies ou produites dans le cadre du projet. Par exemple, les projets dans cette catégorie peuvent produire des codes de logiciel, des algorithmes, des outils numériques, des ensembles de données, des rapports, des articles, des notes de recherche et des sites Web."
  - "Dépôt de copies de tous les codes mis au point grâce à la subvention dans un dépôt approprié – Vous devez décrire votre méthode de conservation du code dans le plan de gestion des données."
  - source "[RFP-French](#)"

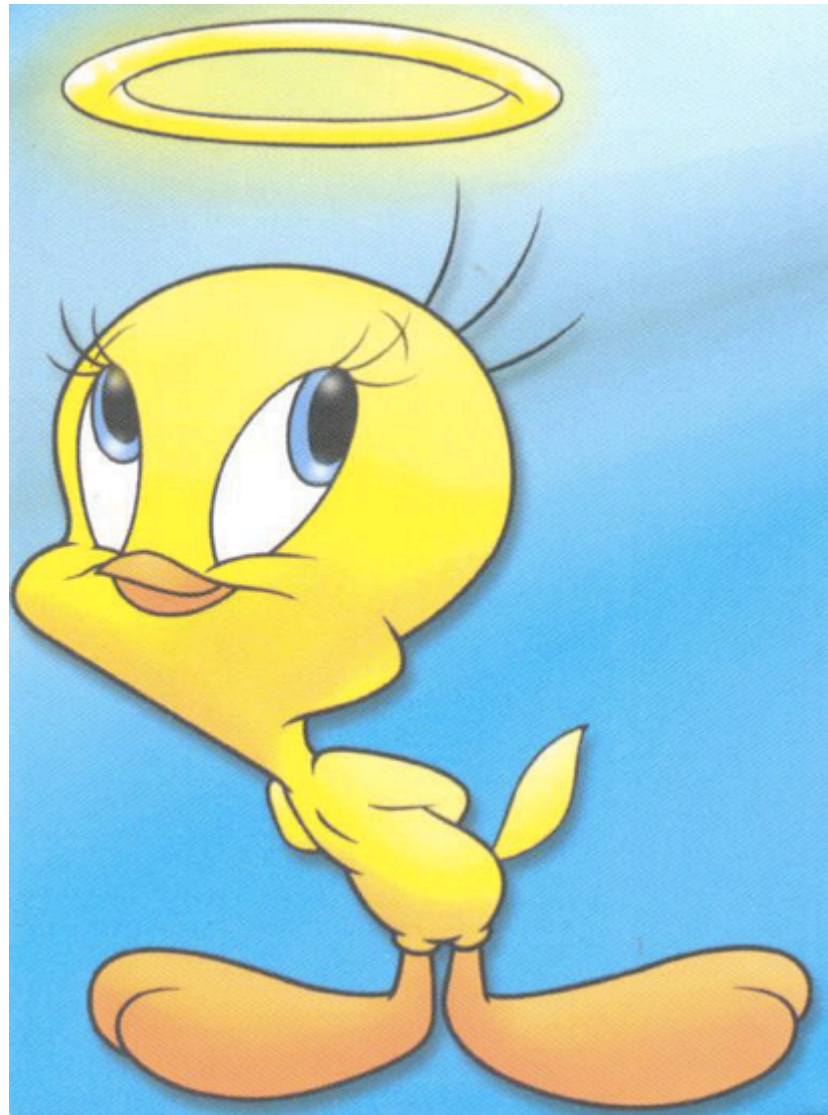




# Quelques questions

- ⦿ Et si le financeur ne l'exige pas, pourquoi un DMP ?
- ⦿ N'est-ce pas une perte de temps ?
- ⦿ Et si on ne souhaite pas partager/publier les données ?
- ⦿ Pour les projets existants, n'est-il pas déjà trop tard ?

# Et si le financeur ne l'exige pas, pourquoi un DMP ?





# Et si le financeur ne l'exige pas, pourquoi un DMP ?

- ⊙ C'est une **bonne pratique** de recherche qui permet de :
  - formaliser la gestion de son travail sur les données
  - gérer son investissement - au bon moment au bon endroit
  - montrer que les données pourront être réutilisées (en dehors de l'équipe, dans le futur...) - ouvrir les données leur donne de la visibilité
- ⊙ Il peut être la base de **datapapers** offrant :
  - peer review ;
  - visibilité ;
  - possibilité d'être cité.



# Dataper, un exemple

- Une publication scientifique dont le but principal est de décrire un ensemble de données ou un groupe d'ensembles de données, plutôt que de rendre compte d'analyses et de résultats de recherche.
- <http://coop-ist.cirad.fr/aide-a-la-publication/rediger/data-paper/1-qu-est-ce-qu-un-data-paper>
- Un exemple issu de la revue *Open Archeology Data*



Journal of open archaeology data

Home About Contact Content Research Integrity

Start Submission Become a Reviewer

LATEST ARTICLES POPULAR ARTICLES

Intensive Survey Data from Antikythera, Greece  
Bevan & Conolly — 15 Feb 2012

Chemical Composition by Neutron Activation Analysis (NAA) of Neo-Assyrian Palace Ware from Iraq, Syria and Israel  
Hart & Neftci — 20 May 2012

Archaeobotanical Data from Two Middle and Later Bronze Age Round House Sites in Cork, Ireland  
Johnson — 08 May 2012

Diachronic and Spatial Distribution of Khabor Ware in the Early Second Millennium BC  
Patterson — 01 May 2012

CONVERSATIONS

Intensive Survey Data from Antikythera, Greece  
Bevan & Conolly — 15 Feb 2012

<http://openarchaeologydata.metajnl.com/articles/10.5334/4f3bcb3f7f21d/>



# Dataper, un exemple



Journal of open archaeology data

Intensive Survey Data from Antikythera, Greece

Authors: Andrew Bevan, James Conolly

Abstract

The Antikythera Survey Project was an interdisciplinary programme of fieldwork, artefact study and laboratory analysis that considered the long-term history and human ecology of the small island of Antikythera, Greece. Directed by Andrew Bevan (UCL) and James Conolly (York) in collaboration with Aris Tsaravopoulos (Greek Archaeological Service), the project was supported by the Canadian Institute in Greece and the Hellenic Ministry of Culture. The project generated various primary datasets are unusual, both in the Mediterranean and beyond, for providing intensive survey coverage of an entire island's surface archaeology.

Keywords: archaeology, GIS, landscape, lithics, pottery

How to Cite: Bevan, A. & Conolly, J., (2012). Intensive Survey Data from Antikythera, Greece. Journal of Open Archaeology Data, 1, p.e3. DOI: <http://doi.org/10.5334/4f3bcb3f7f21d>

1469	179	1	167	2
Views	Downloads	Twitter	Facebook	Wikipedia

Published on 15 Feb 2012 | Peer Reviewed | CC BY 3.0

Intensive Survey Data from Antikythera, Greece

Andrew Bevan and James Conolly

The Antikythera Survey Project was an interdisciplinary programme of fieldwork, artefact study and laboratory analysis that considered the long-term history and human ecology of the small island of Antikythera, in the southwest of the Peloponnese (Greece). Directed by Andrew Bevan (UCL) and James Conolly (York) in collaboration with Aris Tsaravopoulos (Greek Archaeological Service), the project was supported by the Canadian Institute in Greece and the Hellenic Ministry of Culture. The project generated various primary datasets are unusual, both in the Mediterranean and beyond, for providing intensive survey coverage of an entire island's surface archaeology.

Keywords: archaeology, GIS, landscape, lithics, pottery

How to Cite: Bevan, A. & Conolly, J., (2012). Intensive Survey Data from Antikythera, Greece. Journal of Open Archaeology Data, 1, p.e3. DOI: <http://doi.org/10.5334/4f3bcb3f7f21d>

```
<?xml version="1.0" encoding="UTF-8"?>
<JOURNAL_ARTICLE xmlns="http://www.metajnl.com/2012-02-01" >
  <JOURNAL_ARTICLE_TITLE>Intensive Survey Data from Antikythera, Greece</JOURNAL_ARTICLE_TITLE>
  <JOURNAL_ARTICLE_AUTHOR>Andrew Bevan; James Conolly</JOURNAL_ARTICLE_AUTHOR>
  <JOURNAL_ARTICLE_ABSTRACT>The Antikythera Survey Project was an interdisciplinary programme of fieldwork, artefact study and laboratory analysis that considered the long-term history and human ecology of the small island of Antikythera, Greece. Directed by Andrew Bevan (UCL) and James Conolly (York) in collaboration with Aris Tsaravopoulos (Greek Archaeological Service), the project was supported by the Canadian Institute in Greece and the Hellenic Ministry of Culture. The project generated various primary datasets are unusual, both in the Mediterranean and beyond, for providing intensive survey coverage of an entire island's surface archaeology.</JOURNAL_ARTICLE_ABSTRACT>
  <JOURNAL_ARTICLE_KEYWORDS>archaeology, GIS, landscape, lithics, pottery</JOURNAL_ARTICLE_KEYWORDS>
  <JOURNAL_ARTICLE_HOW_TO_CITE>Bevan, A. & Conolly, J., (2012). Intensive Survey Data from Antikythera, Greece. Journal of Open Archaeology Data, 1, p.e3. DOI: http://doi.org/10.5334/4f3bcb3f7f21d</JOURNAL_ARTICLE_HOW_TO_CITE>
  <JOURNAL_ARTICLE_VIEWS>1469</JOURNAL_ARTICLE_VIEWS>
  <JOURNAL_ARTICLE_DOWNLOADS>179</JOURNAL_ARTICLE_DOWNLOADS>
  <JOURNAL_ARTICLE_TWITTER>1</JOURNAL_ARTICLE_TWITTER>
  <JOURNAL_ARTICLE_FACEBOOK>167</JOURNAL_ARTICLE_FACEBOOK>
  <JOURNAL_ARTICLE_WIKIPEDIA>2</JOURNAL_ARTICLE_WIKIPEDIA>
  <JOURNAL_ARTICLE_PUBLISHED_DATE>2012-02-15</JOURNAL_ARTICLE_PUBLISHED_DATE>
  <JOURNAL_ARTICLE_PEER_REVIEWED>true</JOURNAL_ARTICLE_PEER_REVIEWED>
  <JOURNAL_ARTICLE_LICENSE>CC BY 3.0</JOURNAL_ARTICLE_LICENSE>
</JOURNAL_ARTICLE>
```

<http://openarchaeologydata.metajnl.com/articles/10.5334/4f3bcb3f7f21d/>

# Dataper, un autre exemple

Journal of  
**open archaeology data**

Moshenska, G 2012 Selected Correspondence from the Papers of Thomas Pettigrew (1791-1865), Surgeon and Antiquary. *Journal of Open Archaeology Data* 1:e9, DOI: <http://dx.doi.org/10.5334/4f913ca0cbb89>

## DATA PAPER

### Selected Correspondence from the Papers of Thomas Pettigrew (1791-1865), Surgeon and Antiquary

Gabriel Moshenska<sup>1</sup>

<sup>1</sup> UCL Institute of Archaeology, 31-34 Gordon Square, London WC1H 0PY, United Kingdom

The papers of the nineteenth century antiquary and surgeon Thomas Pettigrew are of considerable interest in the histories of science, medicine and archaeology. The letters presented here were received by Pettigrew between 1817 and 1859, and form part of the Pettigrew Papers held at Yale University. They were transcribed during research into Pettigrew's life and work, focusing in particular on his contributions to archaeology and Egyptology, and are presented in the form of texts. They are likely to be of interest and use to scholars of early nineteenth century intellectual history, and historians of archaeology in particular.

**Keywords:** British Archaeological Association, History of Archaeology, History of Medicine, Thomas Pettigrew

#### Funding statement

Leverhulme Trust, Early Career Research Fellowship, 6/SRF/2009/0290.

#### Context

##### Spatial Coverage

UK (54 00 N, 2 00 W) and Holland (52 30 N, 5 45 E)

##### Temporal Coverage

1817AD-1859AD

#### Methods

##### Steps

The dataset was collected by transcribing original letters from the period 1817-59 into an MS Word document. The letters form part of the Pettigrew Papers held by the Beinecke Rare Book and Manuscript Library at Yale University,

through to make sure that details such as dates and names had been correctly copied.

##### Constraints

The entire collection of Pettigrew papers extends to thousands of letters, many of them of little interest, and many that are difficult or impossible to read. For this reason the dataset that has been selected constitutes a small sub-section of the whole collection.

#### Dataset Description

##### Object Name

pettigrew\_letters\_gm.docx / .txt



[www.cnrs.fr](http://www.cnrs.fr)

# N'est-ce pas une perte de temps ?



« [Time management](#) » by Ryan Hyde... via Flickr (CC BY-SA 2.0)

# N'est-ce pas une perte de temps ?

- ① Un DMP est un outil pour améliorer la gestion de ses données,
- ① moins de temps perdu par la suite,
- ① on peut commencer avec une version simple et incomplète et l'améliorer au fil du temps.





# et si on ne souhaite pas partager / publier les données ?



« [Freedom 2](#) » by Ian... via Flickr (CC BY-NC 2.0)

# et si on ne souhaite pas partager/ publier les données ?

- ⊙ Embargo possible (2 ans)
- ⊙ le DMP peut rester interne à l'équipe/au labo
- ⊙ les données pourront être réutilisées facilement par l'équipe dans le futur
- ⊙ permet de repérer des unités négligées mais valorisables (index par exemple)
- ⊙ pour une publication complémentaire
- ⊙ en open data ?



# Et pour les projets existants ?

- ⦿ Tant que les données et leur signification ne sont pas perdues, ce n'est pas trop tard !
  - exemple malheureux “Les données de l'Ecothèque Méditerranéenne”, MC Quidoz, [présentation](#) et [vidéo](#)
- ⦿ On peut prendre le train en marche pour au moins garder une trace de l'existant au niveau de l'équipe, du laboratoire...
- ⦿ Le DMP devient un outil essentiel dans le cas de plusieurs financements successifs dans un même projet scientifique au long cours (photos de l'existant, actions de valorisation, gestion de la PI, pérennisation, travail stagiaires...). C'est un outil essentiel pour l'optimisation, la rentabilisation, la pérennisation. Il permet d'affecter des ressources de la manière la plus optimale et la projection vers de nouveaux projets !







[www.cnrs.fr](http://www.cnrs.fr)

Partie II

# Outils du DMP



# Comment choisir un outil ?

- ⦿ Faut-il utiliser un outil (informatique) ?
- ⦿ Comment organiser le travail collectif autour du DMP ?
- ⦿ Comment gérer l'ajustement constant ?
- ⦿ Quel outil ?
  - Comment choisir selon le projet (confidentialité du DMP) ?
- ⦿ Qui peut aider le chercheur et à quoi ? archiviste, documentaliste, informaticien...
  - cf. slide schéma de Paris 8 Diderot



# DMPonline



- ① <https://dmponline.dcc.ac.uk/>
- ① Outil conçu par le Digital Curation Centre (DCC) et basé à l'origine sur les besoins des chercheurs au Royaume-Uni.
- ① Proposé par la Commission européenne dans le cadre H2020
- ① Peut être utilisé dans d'autres contextes :
  - Modèles liés à d'autres agences de financement
  - Modèle générique
- ① Version en ligne vs version installée localement
- Convergence DMPonline (UK) et DMPTool (US) en cours

# copies d'écran 1/2

**DMP ONLINE** Home About Roadmap Help

Welcome.  
DMPOnline has been developed by the Digital Curation Centre to help you write data management plans.

Screencast on how to use DMPOnline

Sign in

Email address\*

\*\*\*\*\*

[Forgot your password?](#)

Remember me

Sign in

[Or, sign in with your institutional credentials \(UK users only\)](#)

Sign up

New to DMPOnline? Sign up today.

Contact us | Terms of use | DMPOnline previous version  
© 2004 - 2018 Digital Curation Centre (DCC)

**DCC** **Jisc**

**DMP ONLINE** View plans Create plan About Roadmap Help

My plans

The table below lists the plans that you have created, and any that have been shared with you by others. These can be added, shared, exported or deleted at anytime.

Filter plans

Name	Owner	Shared?	Last edited	Select an action
My plan (Horizon 2020 DMP)	Me	No	14-04-2016	Edit Share Export Delete
My plan (DCC Template)	Me	No	23-05-2016	Edit Share Export Delete

Create plan

Contact us | Terms of use | DMPOnline previous version  
© 2004 - 2018 Digital Curation Centre (DCC)

**DCC** **Jisc**

# copies d'écran 2/2

Shared as **Genevieve Remier**

**DMP ONLINE** View plans Create plan About Roadmap Help

## My plan (Horizon 2020 DMP)

Plan details Initial DMP Mid-term Review DMP Final review 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 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
Genevieve Remier	Owner

### Add collaborator

Email

Permissions:

Shared as **Genevieve Remier**

**DMP ONLINE** View plans Create plan About Roadmap Help

## My plan (Horizon 2020 DMP)

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

Please fill in the basic project details below and click 'Update' to save

Plan name	<input type="text" value="My plan (Horizon 2020 DMP)"/>
ID	<input type="text"/>
Grant number	<input type="text"/>
Principal Investigator/Researcher	<input type="text" value="Genevieve Remier"/>
Principal Investigator/Researcher ID	<input type="text"/>
Plan data contact	<input type="text"/>
Description	<input type="text"/>

**Questions to consider:**

- What is the nature of your research project?
- What research questions are you addressing?
- For what purpose are the data being collected or created?

**Guidance:**

Clearly summarise the type of study for students to help others understand the background for which the data are being collected or created.

**This plan is based on:**

Funder	European Commission (Horizon 2020)
Institution	Centre National de la Recherche Scientifique (France)

Horizon 2020 includes a limited and flexible pilot action on open access to research data. Participating projects must develop a Data Management Plan (DMP) specifying which data will be openly accessible.

Projects taking part in the Open Research Data Pilot and those intending to develop a DMP voluntarily must provide a first version of the Data

# Documentation

## 🕒 Tutoriels de l'INIST :

### [Le libre accès aux résultats de la recherche dans le cadre d'Horizon 2020](#)

- Un tutoriel sur les principes généraux de dépôt des publications et données de la recherche dans le cadre d'Horizon 2020
- **5 tutoriels** sur la rédaction d'un DMP (Data Management Plan), c'est-à-dire le plan de gestion des données
- **1 tutoriel** sur le dépôt des données de recherche dans un entrepôt

## 🕒 <https://dmponline.dcc.ac.uk/help#PlanningHelp>

- En particulier :  
[How to develop a Data Management and Sharing](#)







[www.cnrs.fr](http://www.cnrs.fr)

Partie III (Exercice)

# Créer son premier DMP





# Exercices

utilisation en binôme de l'outil en ligne <http://dmponline.dcc.ac.uk>



- ⦿ Création d'un compte chacun (a priori déjà créé)
- ⦿ Choix d'un modèle (générique ?)
- ⦿ Création d'un plan et partage
- ⦿ Identification d'un jeu de données et saisie pour l'exercice
  - Critères de définition d'un jeu de données (dataset)
  - Mettre en avant le raisonnement suivi pour la décision (reproductibilité, coût...)
- ⦿ Quels autres jeux de données seraient à renseigner aussi ?
  - Bonne granularité – stratégie et pratique concrète – impact
- ⦿ Partage du plan avec le collègue du binôme : utilisation de la fonction commentaire
- ⦿ Export du plan dans le format de son choix

# Aide-mémoire DMP



- ⦿ Y a t'il un modèle requis (financeur, tutelle..) de DMP pour le projet ?
- ⦿ Qui va contribuer au DMP ? (dans le laboratoire, chez les partenaires)
- ⦿ Qui peut aider ?
- ⦿ Qui va utiliser le DMP ?
- ⦿ Outil en ligne ?
- ⦿ Première version (livrable)
- ⦿ Quand remettre à jour le DMP ? => « On doit » (livrable) vs « On peut » (suivi des évolution du projet) ?
- ⦿ Version finale
- ⦿ Quelles sont les données du projets ? identifier les datasets

# Liens utiles : DMP dans les projets H2020



- Recommandations de la Commission européenne :
  - [http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm)
  - [http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm](http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm)
- Traduction française : INIST-CNRS, France
  - [http://www.donneesdelarecherche.fr/IMG/pdf/lignes\\_directrices\\_gestion\\_des\\_donnees\\_horizon\\_2020\\_version2\\_1\\_tr\\_fr.pdf](http://www.donneesdelarecherche.fr/IMG/pdf/lignes_directrices_gestion_des_donnees_horizon_2020_version2_1_tr_fr.pdf)
- Réaliser un plan de gestion de données de la recherche - Data management Plan (DMP), guide de rédaction de l'université Diderot :
  - [http://www.univ-paris-diderot.fr/sc/site.php?bc=recherche&np=Data\\_DMP&g=m](http://www.univ-paris-diderot.fr/sc/site.php?bc=recherche&np=Data_DMP&g=m)

# Pour aller plus loin - 1

## 🕒 Thématique « données de la recherche » traitée par les organismes et réseaux métier en IST/données de la recherche

- DIST/CNRS <http://www.cnrs.fr/dist/>
- INIST/CNRS <http://www.inist.fr/?Donnees-de-la-recherche&lang=fr>
- TGIR Huma-Num <http://www.huma-num.fr/>, Progedo <http://www.progedo.fr/>
- URFIST <http://urfistinfo.hypotheses.org/>
- BSN 10  
<http://www.bibliothequescientifiquenumerique.fr/bsn-10-donnees-de-la-recherche/>
- Réseau Base de données <http://rbdd.cnrs.fr/>
- Réseau Renatis <http://renatis.cnrs.fr/>
- Réseau Mate-SHS <http://mate-shs.cnrs.fr/>
- A suivre : action DoRANum <http://urfistinfo.hypotheses.org/2980>

# Pour aller plus loin - 2



## Le site du [DCC](#)

A screenshot of the Digital Curation Centre (DCC) website. The header features the DCC logo and the tagline 'because good research needs good data'. A navigation menu includes links for Home, Digital curation, About us, News, Events, Resources, Training, Projects, Community, and Tailored support. The main content area is titled 'Data Management Plans' and includes a list of resources such as Briefing Papers, How-to Guides &amp; Checklists, and a 'Data Management Plans' section with sub-links like Checklist, DMPonline, and FAQ. A 'Useful links' section provides links to 'Funders DMP requirements' and 'Checklist for a DMP'. A video player is embedded, titled 'Managing Research Data (video)', with a thumbnail showing a globe and a person.

## Workshop on "Active Data Management Plans"

- 28-30 juin 2016, CERN, Suisse, organisé par le CERN et le Research Data Alliance (RDA) [Active Data Management Plans Interest Group](#)
- Inscriptions en cours...

# Repères biblio

- ① [Gestion des données de la recherche : quelques pistes pour démarrer, comprendre et se former](#), billet très riche en références de Odile Contat (site *Corist-SHS*)
- ① [Introduction à la gestion des données de recherche](#), cours en ligne de Natalie Clairoux
- ① [Policy, Practice, and Law](#), billet de Melissa Levine







[www.cnrs.fr](http://www.cnrs.fr)

# A vous de jouer !

