



Gestion des données numériques en archéologie : vers un modèle commun de plan de gestion de données

Emmanuelle Bryas (Inrap)
Camille Colin (Inrap)

Nathalie Letellier-Becquart (MAE)
Emmanuelle Morlock (MOM, HiSoMA)

Journées MASA 2018 : traitement et la diffusion des données de l'archéologie - Maison méditerranéenne des sciences de l'homme (MMSH) , 14, 15, 16 novembre 2018.

Contexte réglementaire

- **Engagement 14 : construire un éco système de la science ouverte** (Deuxième plan d'action national pour la France : pour une action publique transparente et collaborative 2018-2020)

« Généraliser progressivement via un accompagnement la mise en place de **plans de gestion des données** dans les appels à projets de recherche, et inciter à une ouverture des données produites par les programmes financés (2019 et en continu). »

- **Plan national pour la science ouverte 4 juillet 2018**

« Généraliser la mise en place de **plans de gestion des données** dans les appels à projets de recherche »

- **Plan d'action 2019 de l'ANR**

« l'ANR appelle les coordinateurs à considérer la question des données de recherche au moment du montage et tout au long du projet ... et imposera de plus un **plan de gestion des données (DMP)** pour les projets financés à partir de 2019. »

Qu'appelle t-on « données de la recherche »

« les données de la recherche sont définies comme **des enregistrements factuels** (chiffres, textes, images et sons), qui sont utilisés comme **sources principales** pour la recherche scientifique et sont **généralement reconnus par la communauté scientifique** comme nécessaires pour **valider les résultats de la recherche.** »

OCDE, *Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics*, 2007

Plus concrètement, ce sont des données qui :

- doivent être **réexploitables** par de nouveaux programmes de recherche ;
- s'insèrent dans un **ensemble documentaire** (archives, publications) pour **témoigner de la manière dont s'est opéré le travail scientifique**, et pour en témoigner à l'ensemble de la société et non uniquement à la communauté scientifique.

(extrait de Fayet (Sylvie), « Données » de la recherche, les mal-nommées, UrfistInfo <https://urfistinfo.hypotheses.org/2581>, consultation du 12/11/2018)

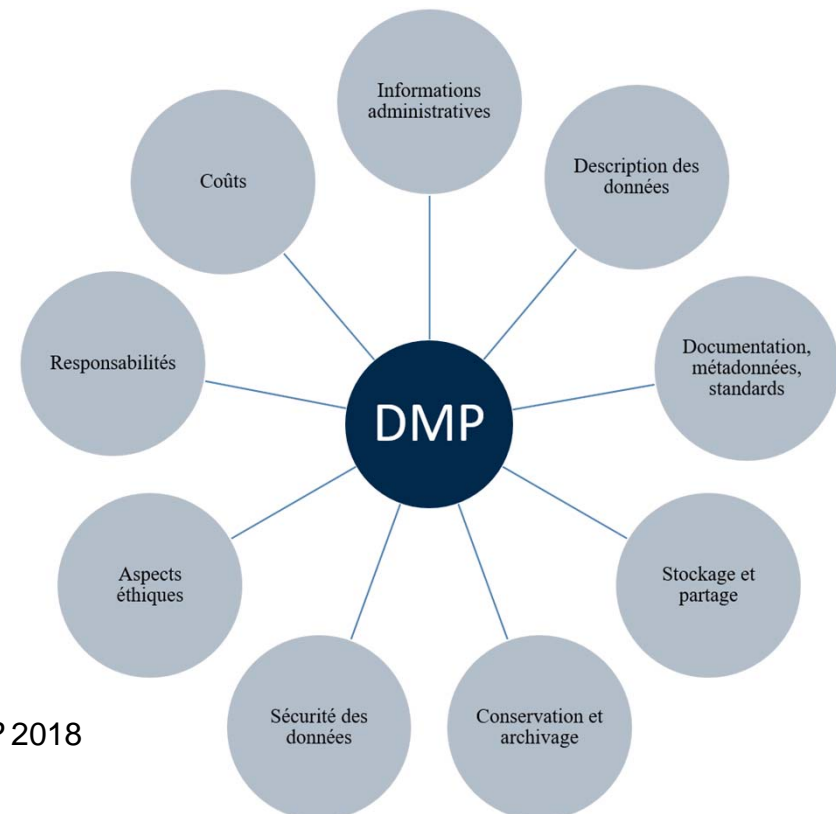
Quest-ce qu'un plan de gestion de données ?

Document **formalisé et évolutif** qui explicite la manière dont les données produites au cours et à l'issue d'un processus ou d'un projet seront :

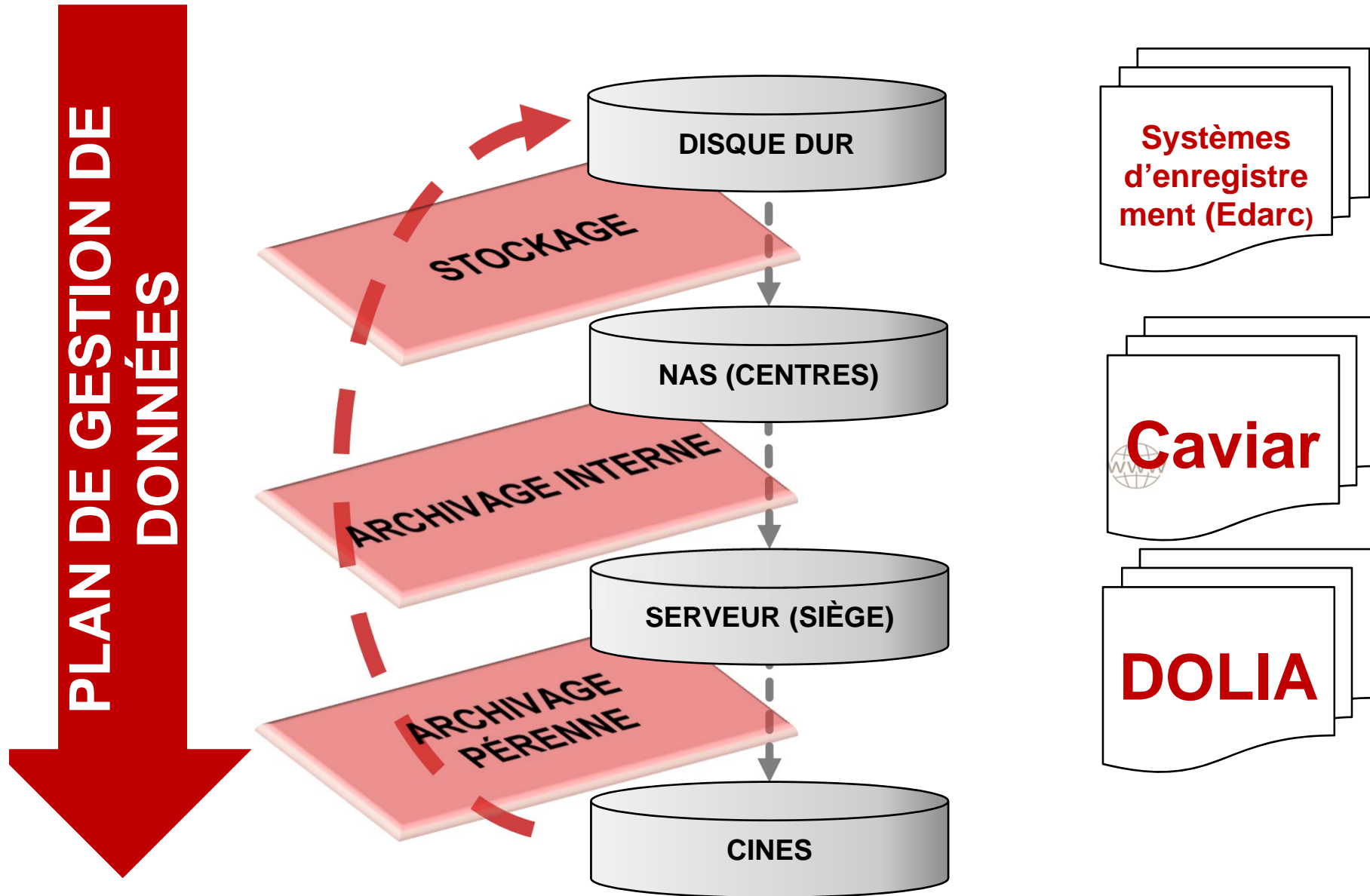
- **obtenues** ;
- **documentées** ;
- **analysées** ;
- **disséminées** ;
- **utilisées.**

C'est un outil de gestion qui s'appuie sur le cycle de vie des données :

- conditions de **production** ;
- conditions de **stockage** ;
- modalité de **conservation** ;
- Modalités de **partage et de diffusion.**



Genèse du projet : un PGD à l'Inrap



Un plan de gestion de données appliqué aux opérations d'archéologie

- Fournir une « **carte d'identité** » de l'opération alimentée par les différents systèmes d'information de l'institut ;
- Établir clairement **la responsabilité de chacun**, producteur(s) et utilisateur(s), sur les documents produits au cours de l'opération ;
- Permettre la **localisation du document** tout au long de son cycle de vie.

⇒ **1^{ère} pierre de l'Open Data**



Un groupe de travail PGD au sein de MASA

Pourquoi rédiger un plan de gestion de données :

- par obligation ;
- pour faciliter le travail collaboratif ;
- pour produire des données FAIR ;
- pour faciliter l'élaboration d'une politique de dissémination.

Au sein du consortium MASA :

- établir un modèle de plan de gestion de données qui puisse être utilisé tel quel ou adapté par toute équipe menant un projet de recherche dans le domaine de l'archéologie ;
- bénéficier d'une masse critique de projets pour tester le plan élaboré à l'INRAP et en tirer une version plus générique.

Un modèle de plan de gestion de données (1/2)

Une première partie consacrée à la **description du projet**

➤ **INFORMATIONS GÉNÉRALES SUR LE PGD**

➤ **INFORMATIONS GÉNÉRALES SUR LES DONNÉES**

Origine des données, volume, contrôle qualité.

INFORMATIONS GÉNÉRALES SUR L'OPÉRATION (à renseigner si nécessaire) *Répéter le bloc autant de fois qu'il y a d'opération*

Ce sous bloc est à utiliser lorsque le projet repose sur l'exploitation de données issues d'une ou plusieurs opérations d'archéologie clairement identifiables.

➤ **INFORMATIONS SUR LE PROJET**

Description du projet : nature, titre, porteur, financement, contenu.

➤ **RESPONSABILITÉ DE LA GESTION DES DONNÉES**

Régime de propriété des données.

➤ **RESSOURCES NÉCESSAIRES À LA MISE EN ŒUVRE DU PGD**

➤ **JEUX DE DONNÉES**

Présentation globale des jeux de données décrits dans le plan de gestion.

Un modèle de plan de gestion de données pour l'archéologie (2/2)

Une deuxième partie consacrée à la **description des jeux de données**, à répéter autant de fois qu'il y a de jeux de données.

➤ JEU DE DONNÉES

Nom, producteur.

➤ **DESCRIPTION DES DONNÉES** (zone à répéter si jeu de données techniquement hétérogène)

Nature, méthode de production, format.

➤ **MÉTADONNÉES : DOCUMENTATION DES DONNÉES** (zone à répéter si jeu de données techniquement hétérogène)

Format, mode de production

➤ **STOCKAGE, ACCÈS ET SÉCURITÉ DES DONNÉES**

Lieu de stockage, volumétrie, sécurité, gestion et garantie des accès.

➤ **À L'ISSUE DU PROJET - DISSÉMINATION**

Règles de diffusion, licence, plateforme/base de données d'enregistrement des données produites.

Identification des données sensibles, mesures de protection.

➤ SÉLECTION ET ARCHIVAGE DES DONNÉES

Tri, volume, durée et lieu de conservation.

Actions 2018

Organisation de deux ateliers en 2018 :

- le 9 février à l'Inrap
- le 21 juin à la MOM

Objectifs : tester, discuter le modèle établi à l'Inrap à partir du PGD proposé par les universités Paris Descartes et Paris Diderot*, afin d'établir un **modèle de plan de gestion de données minimal** appliqué à l'archéologie.

Aujourd'hui, un premier modèle est proposé, qu'il faut maintenant tester en conditions réelles.

* Accessible sur OPIDoR à l'adresse suivante https://dmp.opidor.fr/template_export/322740362.pdf

Perspectives 2019

Au premier semestre :

La validité du modèle établi sera **testée dans le cadre d'un service d'archives**, sur un corpus de données clos, pour un projet déjà achevé.

Deux réunions :

- en début d'année pour tirer les conclusions des premiers tests effectués et ajuster le modèle type de plan de gestion de données en conséquence ;
- une deuxième, à la fin du premier semestre, sous forme d'un **atelier pratique** pour la « prise en main » du modèle établi.

Au second semestre :

- Soumission du modèle élaboré à **OPIdoR** ;
- Début d'une réflexion sur la possibilité de produire des **Data Papers** à partir du plan de gestion de données (chantier 2020 ?).

