



HAL
open science

A la découverte du plan de gestion des données : présentation et utilisation

Léa Mosnier

► **To cite this version:**

Léa Mosnier. A la découverte du plan de gestion des données : présentation et utilisation. Semaine Data-SHS, Dec 2022, Aix en Provence, France. hal-03940445

HAL Id: hal-03940445

<https://amu.hal.science/hal-03940445>

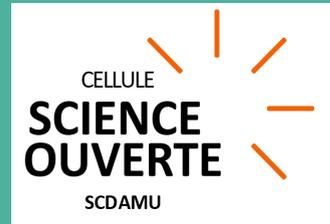
Submitted on 16 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A la découverte du plan de gestion des données présentation et utilisation

Léa Mosnier - Chargée de mission données de la recherche
Cellule Science Ouverte - **Service Commun de Documentation AMU**



CONTEXTE & POLITIQUES ACTUELLES



LE PGD/DMP, QU'EST CE QUE C'EST ?

Plan de gestion des données (PGD) ou Data Management Plan (DMP)

- Document qui permet de **mettre en place les actions indispensables** pour gérer les données de recherche
- **Décrit** la façon dont les données sont gérées depuis leur création, collecte jusqu'au partage et archivage
- **Evolutif** : la rédaction commence dès le début du projet et peut/doit être mis à jour tout au long du projet
- **Aide à organiser et anticiper** les étapes du cycle de vie des données

Un outil au service du projet de recherche

UN DOCUMENT INCONTOURNABLE



A l'échelle mondiale

De plus en plus recommandé

A l'échelle européenne



- Exigence de la Commission européenne
- Outils et infrastructures déployés spécialement

A l'échelle nationale



- PNSO 1 et 2
- Incité ou obligatoire selon les organismes de recherche, universités, agences de financements

A QUOI SERT-IL ?

Montrer **comment les données sont gérées** tout au long du projet et au-delà

- Selon de bonnes pratiques
- Dans le respect d'un cadre éthique et juridique

Respecter les **principes FAIR** :
« aussi ouvert que possible, aussi fermé que nécessaire »

- Faciliter les échanges entre partenaires
- Anticiper les besoins et gagner du temps

SES BÉNÉFICES



Réfléchir et évaluer les
ressources nécessaires



Rendre les données
action-machinable



Définir les rôles et
responsabilités de chacun



Opportunité de dialogue
entre les différents
acteurs



Rendre les données FAIR



Faciles à trouver

F



Intéropérables

I

A

Accessibles



R

Réutilisables



FAIR data VS Open data

- **Les données ouvertes (open data)** : contenus **librement accessibles** qui peuvent être **utilisés, modifiés et partagés** par n'importe qui pour n'importe quel but
- **Les données FAIR (FAIR data)** : données qui suivent un certain nombre de **principes** pour être **lisibles** et **utilisables** par des **humains** comme par des **machines**

"Aussi ouvert que possible, aussi fermé que nécessaire"

QUAND FAIRE UN PGD/DMP ?



QUEL CONTENU DANS UN PGD/DMP ?



Description des données



Documentation et qualité des données



Stockage et sauvegarde des données en cours



Exigences légales et éthiques



Stratégie de partage et ouverture des données



Préservation à long terme



Responsabilité et ressources



DESCRIPTION DES DONNÉES

« Les données produites seront les suivantes :

- **Documents textuels** des textes antiques non lemmatisés et lemmatisés en langue grec (**alphabet grec ancien avec accent, format Unicode**) en **.txt regroupés en dossiers par auteurs**, c'est-à-dire **6 dossiers de maximum 30 Mo** [...] nous disposons de **119 fichiers .txt** (16+103) des lemmatisations [...] »

Exemple extrait du DMP public du projet « Hospitam », modèle ANR, de L. Locatelli





DOCUMENTATION ET QUALITÉ DES DONNÉES

« [...] Nous avons déployé l'**outil de stockage de la TGIR Huma-Num ShareDocs** en formant les membres du projet. [...] Nous avons **1 To de disponible sur cette plateforme**. L'organisation par dossier a été pensée en équipe et prend la forme suivante : **AAP_ANR ; AXE1** [...].

Cette représentation numérique est stockée sur ShareDocs, par terrain et par habitant en utilisant le **codage suivant** :

Ex : **E2GF_B1= Enquête 2 Grenoble Flaubert Baluchon 1** »

Exemple extrait du DMP public du projet « PROSECO », modèle Science Europe, de T. Manola





STOCKAGE ET SAUVEGARDE DES DONNÉES EN COURS

« [...] Les données biologiques seront **stockées dans trois disques durs externes conservés dans des endroits distincts à accès contrôlé et dans un serveur dédié à la sauvegarde** [...].

Les données génétiques seront stockées sous forme de **fichier .txt** [...].

L'**hébergement** [...] avec une **localisation des données assurées en Union Européenne**, à travers un **entrepôt de données de type SQL**. »

Exemple extrait du DMP public du projet « PrediMAP », modèle ANR, de K. MESBAHI-IHADJADENE





EXIGENCES LÉGALES ET ÉTHIQUES

« Le fichier informatique utilisé [...] est mis en œuvre conformément à la **réglementation française (loi Informatique et Libertés modifiée) et européenne (Règlement Général sur la Protection des Données–RGPD)**. [...] aucune Investigation Clinique ne peut être pratiquée sur une personne **sans son consentement libre et éclairé, recueilli par écrit** après que lui a été délivrée l'information [...]. »

Exemple extrait du DMP public du projet « PrediMAP », modèle ANR, de K. MESBAHI-IHADJADENE





STRATÉGIE DE PARTAGE ET OUVERTURE DES DONNÉES

« Comme il s'agit de **données personnelles et potentiellement sensibles**, un **partage sur un entrepôt de données n'est pas envisagé de façon ouverte**. Les données récoltées seront **conservées en France pendant 3 ans après la fin du projet et anonymisées**, [...]. **Seuls les membres du projet scientifique ANR PROSECO** auront accès aux baluchons et **à leur version numérisée de manière sécurisée**. »

Exemple extrait du DMP public du projet « PROSECO », modèle Science Europe, de T. MANOLA





PRÉSERVATION À LONG TERME

« Les données cliniques seront **conservées en base active jusqu'à 2 ans** après la dernière publication de la recherche. Les données biologiques **seront conservées sur le CLE jusqu'à 10 ans après la dernière publication de recherche**. Les métadonnées [...] être **archivées sur le site d'archivage EGA**. Un DAC contrôlera l'**accès aux données anonymisées à des tiers pour des recherches scientifiques dans l'intérêt public**. »

Exemple extrait du DMP public du projet « PrediMAP », modèle ANR, de K. MESBAHI-IHADJADENE





RESPONSABILITÉS ET RESSOURCES

« Le **responsable de la gestion des données** est le **Chef de projet RHU PrediMAP- APHP**. Les personnes impliquées dans le **traitement des données cliniques** :

- **Technicien de recherche clinique** pour la saisie des données clinique
- **Attaché de Recherche Clinique** : gestion et traitement des queries
- **Datamanagement programmation** de la base de données / eCRF

[...]

- 120 000 euros pour la **supervision des aspects techniques de la gestion des données** de la recherche clinique [...] »

Exemple extrait du DMP public du projet « PrediMAP », modèle ANR, de K. MESBAHI-IHADJADENE



Les outils pour faire un plan de gestion de données



DMP Tool

- Créer par l'Université de Californie
- Modèles simplifiés de DMP



DMP Online

- Créé par le Digital Curation Center (DCC)
- Prend en compte les recommandations de la Commission Européenne (H2020)



DMP Opidor

- Instance française de DMP Online du DCC qui est mis à disposition de l'ESR
- Hébergé et géré par l'Inist-CNRS

QUEL OUTIL UTILISER POUR RÉDIGER ?



- Gratuit
- Accessible à l'ensemble de l'ESR et partenaires français ou étrangers
- Sécurisé
- Evolutif

- **Modèles financeurs** et d'**établissements** disponibles
- **Référentiels pour certains modèles**, par ex Science Europe, qui permet de fournir des listes standardisés de termes
- **Espace institutionnel AMU**
- Téléchargeable sous différents formats
- **Assistance personnalisé de rédaction**

DMP du projet "TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles "

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "ANR - Modèle de PGD (français)" fourni par Agence nationale de la recherche (ANR).

Renseignements sur le plan

Titre du plan	DMP du projet "TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles "
Langue	fra
Date de création	2020-06-26
Date de dernière modification	2020-07-22
Identifiant	

Renseignements sur le projet

Titre du projet	TRANSFUNERAIRE: approche comparative des rituels collectifs de ré-inhumation en Europe et en Amérique-Latine XXe-XXIe siècles
Résumé	<p>Notre programme est réalisé par deux équipes d'anthropologues sociaux travaillant sur l'Europe (sous la direction d'Elisabeth Anstett) et sur l'Amérique-Latine (sous la direction de Valérie Robin Azevedo) ; il propose d'effectuer une analyse comparative des pratiques funéraires qui ont actuellement lieu en Europe et en Amérique-Latine, dans des contextes marqués par la violence extrême et la mort de masse.</p> <p>Il vise à éclairer leurs transformations les plus contemporaines mais aussi les bricolages rituels, religieux et politiques mis en œuvre pour faire avec des corps fragmentés ou une absence inéluctable de corps. Au-delà d'une interrogation sur l'adaptabilité des grammaires et des lexiques funéraires, et d'un questionnement sur l'impact des pratiques médico-légales, notre objectif est de saisir les enjeux de ces transformations et les modalités de reconfiguration de la relation aux défunts ré-inhumés. Dans ces situations mortuaires inédites, une compréhension empirique et précise des procédures matérielles et symboliques entourant le destin posthume des restes humains ou des corps absents, ambitionne ainsi de mieux cerner les mécanismes socioculturels de la résilience en temps de crise.</p>

Sources de financement

- Agence nationale de la recherche (ANR) : ANR-19-CE27-0022

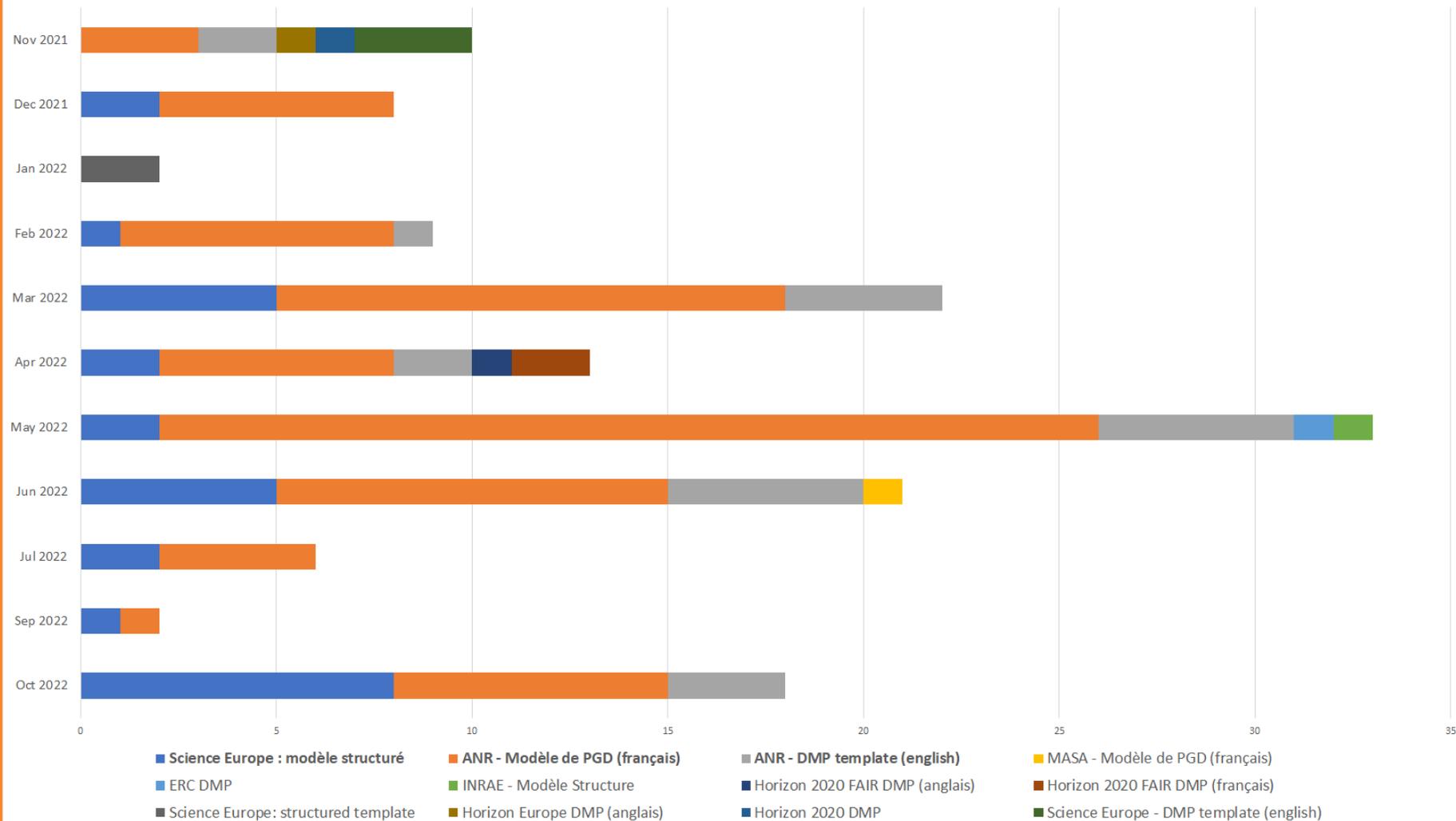
Produits de recherche :

1. Données collectées par les membres du programme Transfunéraire (Collection)
2. Données produites par les membres, dans le cadre du programme Transfunéraire (Collection)
3. Fichier croisant les informations relatives aux données collectées et produites (Texte)

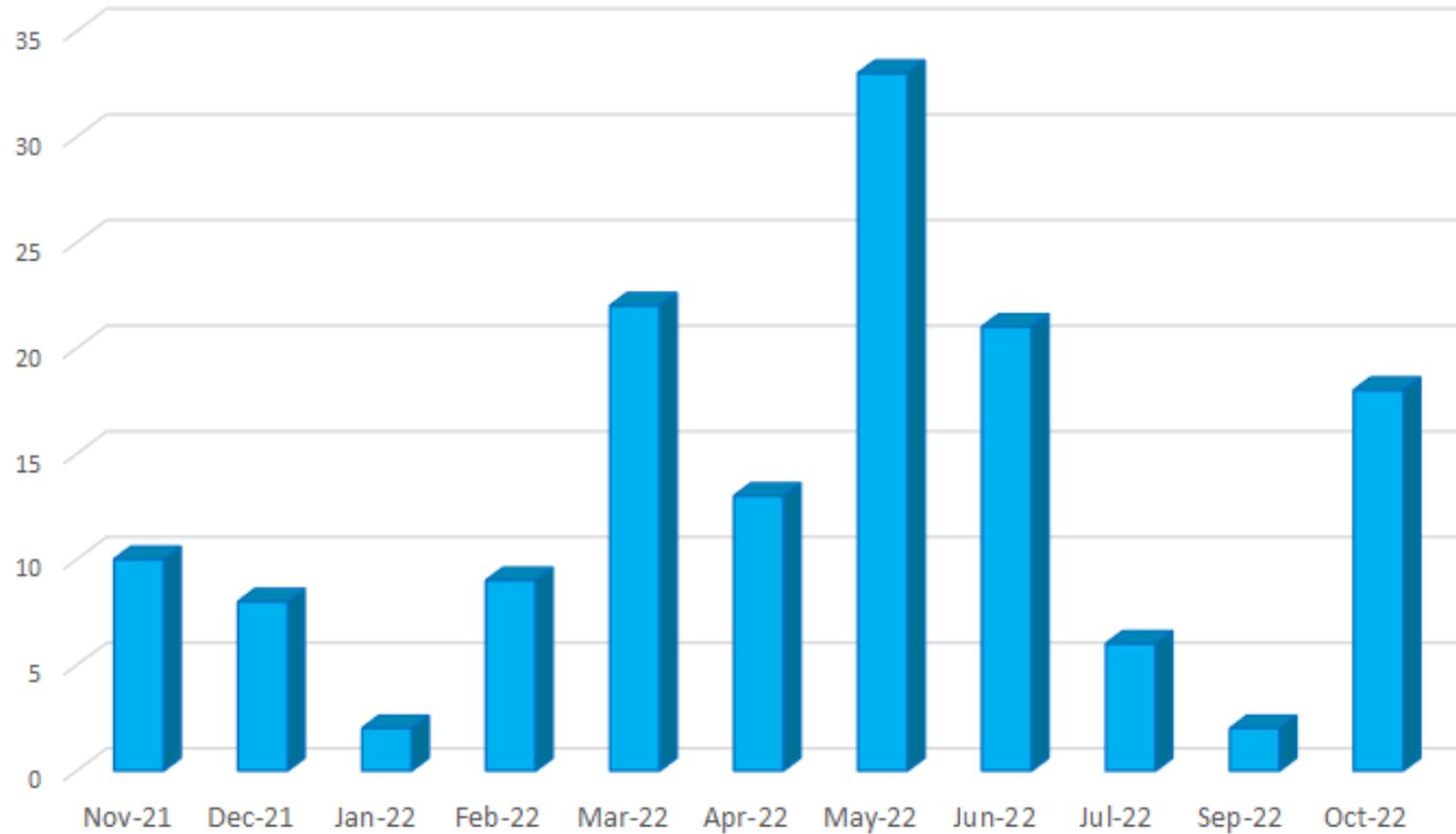
LE PGD/DMP CÔTÉ AMU

- 233 utilisateurs actifs
- 331 plans au total
- 144 PGD fait en un an (2021-2022)
- 33 plans réalisés au mois de mai
- Majorité de plan sur le modèle ANR

Nombre de plans par modèle des 12 derniers mois



Nombre de plans créés au cours des 12 derniers mois





POUR CONCLURE ... 10 ASTUCES POUR RÉDIGER UN PGD/DMP

- 1 Commencer tôt
- 2 Réfléchir à la réutilisation des données
- 3 Vérifier les politiques en vigueur
- 4 Se faire aider par les services d'appui
- 5 Penser large
- 6 S'inspirer du travail d'autrui
- 7 Etre précis là où c'est nécessaire
- 8 Etre concret
- 9 Oser avouer qu'on ne sait pas (encore)
- 10 Mettre à jour



MERCI POUR VOTRE ATTENTION! DES QUESTIONS ?

ASSISTANCE À LA RÉDACTION ET RELECTURE DE PGD/DMP :
SCD-DMP-ASSISTANCE@UNIV-AMU.FR