



HAL
open science

Charte de réutilisation des données-Projet Anarmed

Leila Galai

► **To cite this version:**

| Leila Galai. Charte de réutilisation des données-Projet Anarmed. CNRS. 2024. hal-04666019

HAL Id: hal-04666019

<https://hal.science/hal-04666019v1>

Submitted on 1 Aug 2024

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.

DIM PAMIR

Patrimoines matériels – innovation, expérimentation et résilience

Charte de réutilisation des données

Version 1.1 du 05/06/2023

Auteur·e·s : Leila Galai

1. Objectifs de la charte de réutilisation des données

La Charte de réutilisation des données a pour but de permettre à un projet scientifique – qu'il s'agisse d'un projet de recherche doctorale, post-doctorale, d'un équipement, etc. – d'exprimer de manière simple et pratique quels sont ses engagements et ses attentes quant à l'usage de données entrantes ou la réutilisation de données sortantes, et de l'afficher vis-à-vis des parties prenantes du projet.

Les projets financés par le DIM PAMIR associent des scientifiques venant d'établissements de recherche et des partenaires socio-économiques, dont des établissements culturels. Les données sortantes issues du projet sont destinées à la communauté scientifique, mais peuvent aussi être destinées au grand public, par exemple lors d'actions de médiation ou de science participative.

Le questionnaire ci-dessous est un guide pas à pas, qui va vous permettre d'élaborer une charte de réutilisation des données adaptée à vos besoins¹. Vous allez vous positionner à partir des six principes rassemblés dans l'*Heritage Data Reuse Charter*² – qui s'appuie notamment sur les principes FAIR³.

En partant de votre pratique, vous allez expliciter les mesures mises en place pour appliquer ces six principes, les contraintes relatives aux données entrantes et sortantes de votre projet (par exemple les formats, la documentation, la licence) et vos attentes vis-à-vis d'autres parties prenantes en matière de réutilisation des données.

Ce document peut évoluer au cours de votre projet ; c'est pourquoi nous vous invitons à le mettre à jour régulièrement.

Un conseil : répondez aux questions ci-dessous le plus précisément possible. Si vous rencontrez des difficultés pour répondre à certaines questions, vous pouvez bien sûr [contacter l'équipe de coordination du DIM PAMIR](#). Enfin, si certaines questions ne concernent pas votre projet de recherche ou votre équipement, signalez-le

¹ A titre d'exemple, DARIAH-Campus a créé sa propre charte de réutilisation des données en s'inspirant des principes de l'*Heritage Data Reuse Charter*. Voir « DARIAH-Campus Reuse Charter », 19 octobre 2022, <https://campus.dariah.eu/docs/reuse-charter>

² Le contenu de l'*Heritage Data Reuse Charter* est explicité dans le carnet Hypothèses qui y est dédié. Voir Laurent Romary, Erzsébet Tóth-Czifra et Dorian Seilier, *Heritage Data Reuse Charter*, s. d., <https://datacharter.hypotheses.org/>

³ Voir GO FAIR, « FAIR Principles », s. d., <https://www.go-fair.org/fair-principles/>

simplement. Une fois votre charte remplie, nous vous demandons de la déposer sur le portail ouvert [HAL](#) sous la typologie « Rapport / Rapport technique », pour un partage de bonnes pratiques.

2. Définitions

Se reporter au [glossaire](#) du DIM PAMIR pour les termes suivants :

- Données entrantes
- Données sortantes
- Enrichissement
- Équipe projet
- Format ouvert
- Licence
- Métadonnée
- Modèle de citation
- Parties prenantes

Charte de réutilisation des données du projet Anarmed

* = le champ est requis

Voir les options des listes déroulantes = faites un clic droit sur la petite flèche ou sur la zone grisée

Identification des parties prenantes*

Nom du projet* : Documenter les procédés de fabrication des armures anciennes avec le support des sciences archéologiques : nouvelles approches expérimentales

Acronyme du projet* : Anarmed

Nom du ou de la responsable* : Philippe Dillmann

Identifiez les autres parties prenantes*

Nom de l'entité	Nom du ou de la responsable de l'entité
ICMMO	Emilie Bérard
LAPA	Philippe Dillmann
Musée de l'Armée	Olivier Renaudeau
C2RMF	Benoît Mille
Synchrotron Soleil	Solenn Reguer
Université Paris 8-Arscan	Catherine Verna
Université de Cergy-Heritages	Valérie Toureille
Grand public ⁴	

Ajoutez des lignes si besoin

Principes de la charte de réutilisation des données

Réciprocité

Le principe de **réciprocité** porte sur le fait que l'équipe projet et les autres parties prenantes échangent sur les savoirs et connaissances partagés, notamment en matière d'enrichissement des données.

- Avez-vous prévu d'enrichir les données entrantes de votre projet ? Si oui, prévoyez-vous d'en informer les autres parties prenantes qui les ont fournies ?

Oui, nous avons prévu d'enrichir les données entrantes par de nouvelles données. Oui, les parties prenantes qui ont fournis les données entrantes sont informées.

- Pour les données sortantes de votre projet, souhaitez-vous que les autres parties prenantes les enrichissent ? Si oui, quels apports attendez-vous ?

Le projet actuel porte sur le développement méthodologique de techniques de caractérisation non invasives pour l'analyse d'armures dans le but de documenter les procédés de fabrication. Les données sortantes pourront de ce fait être enrichies par d'autres études qui auront lieu à la suite de ce projet post-doctoral par les mêmes parties prenantes. Ces apports peuvent se présenter sous forme de données d'analyses par DRX, LIBS ou d'autres techniques de caractérisations.

- Quand vos données sortantes sont enrichies par d'autres parties prenantes, quel mécanisme avez-vous mis en place pour recueillir ces enrichissements ?

⁴ Précisez si le grand public est impliqué oui ou non dans votre projet.

Partage des données par échange de fichiers (excel ou fichiers brutes)

Interopérabilité

Le principe d'**interopérabilité** porte sur le fait que les données sont mises à disposition dans un format ouvert et réutilisable par les parties prenantes. Toutes les parties prenantes s'engagent à privilégier des standards internationaux et des protocoles ouverts⁵.

- Si votre équipe ou votre équipement héberge des données, quels formats de données entrantes acceptez-vous ?

Selon les techniques de caractérisation utilisés, les données seront au format d'images de microscopie (TIF, Jpeg..), de spectres liés aux techniques d'analyses chimiques (ASCII,CSV) et de diffractogrammes. (ASCII). L'ensemble de ces formats sont des formats ouverts.

- Quels sont les formats de données sortantes de votre projet ?

Même réponse que la question précédente.

- Si les formats des données sortantes ne sont pas ouverts, qu'allez-vous mettre en place pour en proposer et selon quel calendrier ?

Les formats cités sont ouverts.

Citabilité

Le principe de **citabilité** porte sur le fait que le projet (avec ses données et ses résultats) doit être cité selon un modèle décidé par les parties prenantes.

- Quel est votre modèle de citation pour le projet ?

Projet Anarmed, IDF- IDF-DIM-PAMIR-2022-2-001

- Quel est votre modèle de citation pour les données sortantes de votre projet ?

Citation de l'article publiant les données.

Ouverture

Le principe d'**ouverture** porte sur le fait d'associer une licence à des données. Il est recommandé de partager ces dernières autant qu'il est possible, sous des licences internationales telles que Creative Commons⁶.

- Comment mettez-vous à disposition aux différentes parties prenantes les données sortantes de votre projet ?

A terme, dépôt et partage sur Huma-num

- Quelle licence associerez-vous aux données sortantes de votre projet ?

Cela n'a pas encore été décidé.

Maintenance des données

Le principe de **maintenance des données** porte sur le fait de préserver les données à court, moyen et long terme.

- Qui est en charge de la maintenance des données sortantes de votre projet ? Indiquez un contact mail.

Les données sortantes seront conservées sur le serveur NAS du LAPA (philippe.dillmann@cea.fr) et copie sur Human-num.

⁵ Se référer aux principes FAIR A1.1 et I1. Voir GO FAIR, « A1.1: The protocol is open, free and universally implementable », s. d., <https://www.go-fair.org/fair-principles/a1-1-protocol-open-free-universally-implementable/> et GO FAIR, « I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation », s. d., <https://www.go-fair.org/fair-principles/i1-metadata-use-formal-accessible-shared-broadly-applicable-language-knowledge-representation/>

⁶ Voir Wikipédia, « Licence Creative Commons », 19 octobre 2022, https://fr.wikipedia.org/wiki/Licence_Creative_Commons

- Est-ce que votre équipe propose à la communauté une solution d'hébergement pour les données du projet ? Si oui, quelle est cette solution ?
Pour les données acquises pendant le projet, notre équipe propose d'héberger les données sur le serveur NAS du LAPA. Les données de diffraction acquises au Synchrotron Soleil sont également conservées dans un serveur dédié.
- Avez-vous une politique de préservation des données à long terme ? Si oui, quelle est-elle ?
Les données sont conservées à long terme sur la plateforme Huma-Num.

Fiabilité

Le principe de **fiabilité** porte sur le fait de documenter de manière lisible et accessible la provenance et assurer la traçabilité des-données du projet.

- Pour documenter la provenance des données du projet, quelles métadonnées (par exemple : auteur-e-s et date du jeu de données, institution de production et/ou de conservation des données, etc.) allez-vous fournir ?

Les métadonnées associées aux données seront principalement :

- une référence unique associée à chaque échantillon pour chaque mesure réalisée; - l'identité et les contacts des chercheurs impliqués (personnes en charge des techniques de caractérisations), le statut des données (publiées, non publiées), le programme de recherche associé, la date d'analyse ;
- le protocole d'acquisition : identification du montage utilisé pour acquérir les données chimiques et structurales, mode opératoire, caractéristiques des équipements utilisés et incertitudes de mesure associées. La documentation associée aux données sera :
 - Une note technique qui décrit les métadonnées et la façon dont elles ont été collectées et enregistrées, ainsi que les processus de traitement associés.

- Pour assurer la traçabilité des changements apportés aux données, quelles métadonnées (par exemple version, date de mise à jour, personnes contribuant à la mise à jour du jeu de données, etc.) prévoyez-vous de fournir ?

Les numéros de versions seront incrémentés à chaque changement apporté.

Pour ce faire, il importe de se référer à des normes, des standards, des vocabulaires contrôlés. Vous pouvez par exemple consulter le schéma de métadonnées de DataCite⁷ qui s'appuie sur le Dublin Core⁸ qui est utilisé pour décrire des ressources de manière simple et standardisée.

⁷ DataCite est un consortium international regroupant des bibliothèques et institutions spécialisées en sciences de l'information, qui propose des services autour du partage et de la citation des données, dont l'attribution d'identifiants pérennes (DOI). Voir DataCite Schema, « DataCite Metadata Schema 4.4 », 30 mars 2021, <http://schema.datacite.org/meta/kernel-4.4/>

⁸ Le Dublin Core est un standard de métadonnées. Voir DublinCore, « DCMI Metadata Terms », 20 janvier 2020, <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>