



HAL
open science

Charte de réutilisation des données du projet PaléoLASER

Thibault Clauzel

► **To cite this version:**

Thibault Clauzel. Charte de réutilisation des données du projet PaléoLASER. Centre national de la recherche scientifique. 2024. hal-04684898

HAL Id: hal-04684898

<https://hal.science/hal-04684898v1>

Submitted on 3 Sep 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 du projet PaleoLASER

Version 2 du 02/09/2024

Auteur : Thibault CLAUZEL

Identification des parties prenantes*

Nom du projet* : Application de la VCOF-CRDS à l'étude des environnements du passé

Acronyme du projet* : PaléoLASER

Nom du ou de la responsable* : Mathieu Daëron

Identifiez les autres parties prenantes*

Nom de l'entité	Nom du ou de la responsable de l'entité
ISTeP	Olivier LACOMBE
GEOPS	Christophe COLIN
LiPhy	Bahram HOUCHMANDZADEH
Grand public ¹	non

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 ?

Le projet n'accueille pas de données entrantes extérieures, elles sont fournies directement par l'équipe.

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

Non, il n'est pas attendu des autres parties prenantes qu'elles enrichissent les données sortantes du projet.

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

Dans le cas d'enrichissements, leurs définitions sont réalisées de manière collaborative lors des publications scientifiques.

¹ Précisez si le grand public est impliqué oui ou non dans votre projet.

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 ouverts².

- Si votre équipe ou votre équipement héberge des données, quels formats de données entrantes acceptez-vous ?
Notre équipement n'héberge pas de données.
- Quels sont les formats de données sortantes de votre projet ?
Les données sortantes sont produites sous format numérique type .txt ou .csv. Ces formats sont ouverts, utilisables et interopérables par tous.
- 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 des données sortantes 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 ?
Le projet en lui-même n'est pas citable. Les résultats associés sont citables selon le modèle académique avec des articles scientifiques associé à un (des) Digital Object Identifier (DOI).
- Quel est votre modèle de citation pour les données sortantes de votre projet ?
Les données sortantes seront citées selon les normes de chaque article scientifique (avec DOI associé) où elles seront présentées. Si les données sont hébergées sur une plateforme (Zenodo), elles seront également associées à des DOI qui permettront leur citation.

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 Commons³.

- Comment mettez-vous à disposition aux différentes parties prenantes les données sortantes de votre projet ?
Les données sortantes, ainsi que les protocoles utilisés pour les calibrer, sont archivés sur Zenodo et GitHub. Leurs interprétations seront mises à disposition dans des articles scientifiques qui seront archivés sur HAL.
- Quelle licence associez-vous aux données sortantes de votre projet ?
Les données sortantes seront associées à une licence Creative Commons CC-BY 4.0, tandis que les codes et programmes seront associés à une Licence MIT, en accord avec les principes FAIR.

Maintenance des données

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

² 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/>

³ Voir Wikipédia, « Licence Creative Commons », 19 octobre 2022, https://fr.wikipedia.org/wiki/Licence_Creative_Commons

- Qui est en charge de la maintenance des données sortantes de votre projet ? Indiquez un contact mail. La personne en charge de la gestion des données est Mathieu Daëron (mathieu.daeron@lsce.ipsl.fr)
- Est-ce que votre équipe propose à la communauté une solution d'hébergement pour les données du projet ? Si oui, quelle est cette solution ?
Non, l'équipe n'est pas à l'origine d'une solution d'hébergement pour les données, elle exploite d'autres solutions comme GitHub et Zenodo.
- Avez-vous une politique de préservation des données à long terme ? Si oui, quelle est-elle ?
La politique de préservation des données est celle du CNRS, telle que décrite dans les "plans données de la recherche", et celles des plateformes d'hébergement des données (Zenodo, HAL).

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 ?

L'ensemble des métadonnées suivantes seront fournies avec les données : date et heure de l'acquisition, opérateur, quel type d'échantillon a été analysé, la nature des données (brute ou corrigée).

- 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 ?

La traçabilité sera apportée par les métadonnées suivante : date d'acquisition des données, type de logiciel de correction utilisé. Dans le cadre de PaléoLASER, les données brutes et corrigées seront fournies. Or, seules les données corrigées sont susceptibles d'être modifiées, contrairement aux données brutes qui ne subiront aucun changement une fois archivées. Les changements sont tracés dans les plateformes d'archivages (comme Zenodo) qui peuvent fournir l'historique des versions existantes.

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 DataCite⁴ qui s'appuie sur le Dublin Core⁵ qui est utilisé pour décrire des ressources de manière simple et standardisée.

⁴ 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/>

⁵ 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/>