
Plan Overview

A Data Management Plan created using DMPonline

Title: CLSR ICONOGRAPHIE

Creator: Nam Pham

Principal Investigator: Nam Pham

Data Manager: Nam Pham

Project Administrator: Nam Pham

Contributor: Stéphane Pétermann

Affiliation: Université de Lausanne

Template: DMP UNIL en français

ORCID iD: 0009-0007-9635-7491

Project abstract:

Ce projet est dédié à la conservation et à la gestion desarchives numérisées du CLSR (photographies, correspondances, manuscrits et documents iconographiques). Chaque unité est identifiée par une nomenclature normalisée commençant par le préfixe "**CRLR**" suivie de **six chiffres**. Ces fichiers, conservés au format**TIFF**, ont pour vocation d'être diffusés auprès de la **communauté scientifique** et d'un plus large public. Afin de garantir l'intégrité des données sur le long terme, les archives sont encapsulées dans des structures **BagIt**, permettant un contrôle technique rigoureux de chaque fichier.

ID: 193405

Start date: 05-01-2026

End date: 01-01-2100

Last modified: 06-01-2026

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

CLSR ICONOGRAPHIE

1. DESCRIPTION DES DONNEES [cette section se rapporte à la question 1.1 du DMP mySNF]

Des données seront collectées, étudiées, générées ou réutilisées dans le cadre de ce projet :

- Oui

Votre projet implique-t-il la réutilisation de données existantes (produites par vous ou des tiers) ?

- Oui

D'où proviennent les données existantes réutilisées, que contiennent-elles et quel est leur format ?

Les fonds proviennent des **producteur·rices** (auteur·rices, collectivités ou institutions) qui confient leurs archives au CLSR par voie de don ou de dépôt. Le répertoire "**CLSR ICONOGRAPHIE**" est spécifiquement dédié aux reproductions numériques haute définition (format **TIFF**). Ce corpus comprend des numérisations de photographies, de correspondances, de manuscrits et de divers documents iconographiques. Ces ressources sont destinées à être valorisées auprès de la **communauté scientifique**, des professionnel·les de l'édition et des médias, ainsi que du grand public.

Indiquez le type des données produites dans le cadre du projet (nouvelles données) :

- Images

Décrivez brièvement le contenu des données produites ainsi que leur format :

Les images numérisées sont en format TIFF.

2. METHODOLOGIE ADOPTEE ET TRAITEMENT DES DONNEES [cette section se rapporte à la question 1.2 du DMP mySNF]

Décrivez comment les données seront collectées et/ou produites dans le cadre de votre projet :

La production d'images haute définition répond aux besoins des usagers, aux travaux des **chercheur·euses** du centre et aux campagnes de numérisation institutionnelles. Ces travaux sont généralement confiés aux photographes de la BCU (Laurent Dubois ou Claudia Renna). En cas de contraintes techniques particulières, le CLSR peut recourir à une expertise externe.

Décrivez la manière dont les données seront traitées, étudiées et analysées dans le cadre de votre projet :

Les images au format **TIFF** sont renommées selon une nomenclature unique ("**CRLR**" suivi de **quatre chiffres**) et organisées dans une structure de répertoires hiérarchisée. Chaque unité documentaire est encapsulée dans un "**Bag**" respectant les spécifications du protocole **BagIt**. La préservation est assurée par un **script de contrôle d'intégrité** exécuté annuellement. Par ailleurs, un second script automatise le traitement des demandes : il extrait les fichiers requis, y injecte les **métadonnées embarquées** nécessaires et génère un dossier local destiné à l'envoi aux **demandeur·euses**.

3. ORGANISATION ET REGLES DE NOMMAGE [cette section se rapporte à la question 1.2 du DMP mySNF]

Décrivez les règles de gestion et l'organisation des dossiers et des fichiers de données adoptée (arborescence classificatoire) :

Les fichiers **TIFF** sont organisés au sein d'une structure de dossiers hiérarchisée et nommés selon un système de **numerus currens** (le préfixe "**CRLR**" suivi de quatre chiffres). Cette cote unique assure l'interopérabilité avec la base de données **CATIMA**, où sont centralisées les descriptions détaillées de chaque image. Les répertoires regroupant ces fichiers sont identifiés par la mention "**BAG_HD**" complétée par les cotes extrêmes de la série (ex. : BAG_HD501-600).

Décrivez les règles de nommage des fichiers adoptées (nomenclature) :

La nomenclature repose sur une règle de nommage stricte : le préfixe "**CRLR**" suivi d'un identifiant numérique à quatre chiffres. Ce système garantit à chaque unité documentaire une **cote unique et pérenne**.

4. DOCUMENTATION ET METADONNEES [cette section se rapporte à la question 1.3 du DMP mySNF]

Allez-vous rédiger de la documentation qui accompagne les données ?

- Oui

Sous quelle forme envisagez-vous de rédiger la documentation qui accompagne les données ?

- Un fichier texte (.docx, .odt, .pdf)

La documentation n'est pas prioritaire par rapport aux autres tâches. Le DMP servira de documentation.

Quelles informations cette documentation fournira-t-elle ?

- Les standards et les variables utilisés (noms, questions, descriptions, algorithmes, syntaxes, etc.)
- La licence de réutilisation des données
- Les formats et les types de données

Indiquez quelles métadonnées ou standards de métadonnées accompagneront les données :

- Je ne sais pas encore

Dans la perspective du déploiement d'un **SAE (Système d'Archivage Électronique)**, nous envisageons d'adopter des standards de métadonnées reconnus tels que le **Dublin Core** ou l'**EAD**.

5. STOCKAGE [cette section se rapporte à la question 3.1 du DMP mySNF]

Veuillez fournir une estimation du volume nécessaire pour stocker les données :

- De 501Go à 1023Go

Quelles infrastructures de stockage utiliserez-vous ?

- Infrastructure de stockage de la Division calcul et soutien à la recherche (DCSR)

6. SAUVEGARDE [cette section se rapporte à la question 3.1 du DMP mySNF]

Par quels moyens les données seront-elles sauvegardées ?

- NAS DCSR (double copie)

7. QUESTIONS ETHIQUES ET LEGALES [cette section se rapporte à la question 2.1 du DMP mySNF]

Veuillez cocher les questions d'ordre éthique soulevées par votre projet :

- Mon projet ne soulève aucune question d'ordre éthique

Bien que la majorité des fonds soit d'un usage courant, une **infime minorité** des données peut être classée comme "**sensible**". Ces cas exceptionnels font l'objet d'un protocole de protection spécifique.

Votre projet implique-t-il la récolte ou le traitement de données personnelles et/ou sensibles ?

- Oui

8. SECURITE DES DONNEES [cette section se rapporte à la question 2.2 du DMP mySNF]

Décrivez les mesures mises en place pour assurer la sécurité informatique des données tout au long du projet :

La gestion des données repose sur des droits d'accès différenciés : seul l'archiviste du CLSR est habilité à les **modifier** ou à les **supprimer**. Les autres collaborateurs disposent exclusivement de droits de **consultation** et de **téléchargement**. Par ailleurs, la sécurité informatique du système est **déléguee** à la DCSR.

Décrivez les mesures de sécurité supplémentaires mises en place pour assurer les impératifs de protection des données personnelles et/ou sensibles :

Afin de garantir la protection de la vie privée, seules les **métadonnées** relatives aux documents contenant des données personnelles ou sensibles sont publiées en ligne. L'**accès à l'image source** (l'objet numérique) demeure restreint. En présence d'informations sensibles dans les descriptions, les ressources visuelles ne sont en aucun cas mises à disposition du public.

9. PROPRIETE INTELLECTUELLE [cette section se rapporte à la question 2.3 du DMP mySNF]

Sous quelles conditions la réutilisation des données créées par des tiers est possible ?

Les fonds conservés sont généralement protégés par le **droit d'auteur** ; toute exploitation est ainsi soumise à l'autorisation préalable des **ayants droit**. Dans ce cadre, le CLSR agit en tant qu'**intermédiaire** entre les demandeur·euses et les titulaires de droits. Concernant la base de données **CATIMA**, conformément à l'**art. 24e de la LDA**, la consultation publique est limitée à un "aperçu global" ou à de "courts extraits de faible résolution", garantissant le respect de la propriété intellectuelle tout en permettant la recherche scientifique.

Les données créées dans le cadre de votre projet seront-elles soumises à des restrictions en lien avec les brevets ou les inventions, ou encore à un contrat ?

- Non

10. CONSERVATION A LONG TERME [cette section se rapporte à la question 3.2 du DMP mySNF]

Selon votre estimation, quelle proportion des données de recherche issues de votre projet est destinée à être conservée à long terme (plus de 10 ans) :

- Toutes les données

Sur quels critères de sélection se base cette estimation ?

Ces ressources sont des **substituts numériques** susceptibles d'évoluer au gré des avancées technologiques (optimisation de la résolution ou de la compression). Compte tenu de l'investissement substantiel, tant **financier qu'humain**, requis pour leur production, leur **préservation à long terme** constitue un impératif stratégique pour le CLSR.

Quels formats d'archivage seront utilisés ?

- Images: TIFF, JPEG 2000, PDF, PNG, GIF, BMP

La stratégie de conservation du CLSR repose sur une gestion dynamique des formats. Elle inclut la possibilité de **migrer les données** vers des standards plus adaptés à la **préservation numérique**, tels que le **JPEG 2000**, afin de garantir une pérennité optimale face à l'évolution des technologies.

11. PARTAGE ET DEPOT DES DONNEES [cette section se rapporte aux questions 4.1 et 4.2 du DMP mySNF]

Le partage des données de votre projet de recherche est-il soumis à des restrictions d'accès (embargo, partage sur demande, données qui ne peuvent pas être partagées en raison de clauses légales, éthiques, etc.) ?

- Oui

Précisez ces restrictions et justifiez-les :

La majeure partie du corpus iconographique demeure soumise au **droit d'auteur**. À ce stade, le protocole de traitement technique est **indifférencié**, que les œuvres soient sous droits ou qu'elles relèvent du domaine public. En revanche, le CLSR encourage activement la découverbarilité et la consultation de ces ressources par le biais de sa base de données **CATIMA**, qui sert de levier à leur mise à disposition. »

Sous quelle licence les données seront-elles mises à disposition ?

- Autre (à préciser)

Le CLSR ne détenant pas les droits de la majorité des ressources diffusées, l'usage de licences **Creative Commons** est proscrit, à l'exception des œuvres appartenant au domaine public (**PDM**). En lieu et place, le Centre adopte le standard international **RightsStatements**. Ces déclarations permettent d'informer précisément le public sur le statut juridique des objets numériques et leurs conditions de réutilisation.

Les déclarations suivantes ont été retenues :

- **InC** (*In Copyright*) : Protégé par le droit d'auteur.
- **InC-RUU** (*Unknown Rightsholder*) : Protégé par le droit d'auteur, mais titulaire(s) non identifié(s) ou non localisé(s).
- **PDM** (*Public Domain Mark*) : Domaine public.
- **CNE** (*Copyright Not Evaluated*) : Statut juridique non encore évalué.
- **CC-BY 4.0** (*Attribution*) : Utilisé exclusivement pour les ressources produites par le Centre ou libérées de droits.

Conformément à l'avis de droit régissant notre base de données, ces informations sont actuellement

masquées pour le public et servent exclusivement à la gestion interne des droits.

Dans quel dépôt (*repository*) les données seront-elles déposées ?

- Je ne vais pas déposer mes données dans un dépôt (expliquez pourquoi)

Contrairement aux projets de recherche clôturés, ces reproductions ne constituent pas des jeux de données figés. Elles doivent être perçues comme des "**données de recherche activables**", dont le contenu a vocation à être **enrichi et valorisé** de manière continue au fil des années.