
Plan Overview

A Data Management Plan created using DMPonline

Title: CLSR ARCHIVES NUMERIQUES

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:

Le "projet" a pour but de stocker les archives nées numériques du CLSR (texte, image, vidéo, son, etc).

Ces archives sont organisées dans un SIP et un AIP. Le premier contient les fichiers au format d'origine et le second les fichiers convertis en format pérenne.

Dans chacun, les dossiers sont conservés dans des Bag It, facilitant la vérification de l'intégrité des fichiers.

ID: 193623

Start date: 05-01-2026

End date: 01-01-2100

Last modified: 17-12-2025

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 ARCHIVES NUMERIQUES

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 données proviennent des "producteurs de fonds", c'est-à-dire des auteur.rices, des collectivités et des institutions qui déposent leurs archives au CLSR.

Dans le dossier CLSR ARCHIVES NUMERIQUES, les archives nées numériques y sont stockées, renommées après un tri et un nettoyage de données.

Elles sont de divers formats, majoritairement du texte et des images.

Le dossier pourra également accueillir les courriels qui composeraient une part importante des versements.

Elles seront utilisées par la communauté des chercheur.euses, les éditeurs, les journalistes et un plus large public.

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

- Autre (à préciser)
- Textuelles
- Vidéo
- Audio
- Tableurs
- Images

Le dossier pourrait accueillir des copies images d'ordinateurs, tablettes ou smartphone.

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

Les données contiendraient des informations relatives à l'activité littéraire du producteur.rice du fonds, à ses engagements professionnels, associatifs ou politiques, à sa vie personnelle. Dans le cas d'un fonds d'archives d'une collectivité (maison d'édition, revues, associations, etc), les données contiendraient des informations sur l'activité et le fonctionnement de celle-ci.

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 :

Les archives font l'objet d'un don ou d'un dépôt au CLSR. Les données sont extraites des supports numériques (clé USB, CD-ROM, disque dur externe, etc)

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

Les doublons seront identifiées à l'aide de logiciel et, éliminé selon leur per

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 stockés dans un système hiérarchique de dossiers et sont renommées "CRLR" suivi de 4 chiffres (numerus currens). L'organisation sui La cote unique renvoie à notre base de données [CATIMA](#) qui décrit ces images.

Le dossier qui contient les TIFF est renommé "BAG_HD" suivi des cotes extrêmes (par exemple, "BAG_HD501-600")

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

Le nommage suit cette règle : "CRLR" suivi de 4 chiffres, fournissant ainsi une cote unique et attribuée.

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.

Quelles informations cette documentation fournira-t-elle ?

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

Indiquez quelles métadonnées ou standards de métadonnées accompagneront les données :

- Je ne sais pas encore

Si nous parvenons à développer un SAE, il est probable que nous nous tournions vers 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

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 :

Seul l'archiviste du CLSR peut supprimer et modifier les données, les autres collaborateurs peuvent uniquement consulter et télécharger.

La sécurité informatique est déléguée à la DCSR.

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

Seules les métadonnées des données personnelles et/ ou sensible sont mises en ligne. Si les descriptions comportent des données personnelles ou sensibles, les images ne sont tout simplement pas 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 données sont généralement soumises au droit d'auteur et ne peuvent être utilisées sans l'autorisation des ayant droits.

Le CLSR fait office d'intermédiaire entre le demandeur.euses et l'ayant-droit.

Pour sa base de données CATIMA, l'art. 24e de la LDA, permet la consultation sous forme d' « aperçu global » ou de « court extrait de faible résolution ou de format réduit »

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) :

- Entre 76% et 99%

Sur quels critères de sélection se base cette estimation ?

Il s'agit techniquement de reproductions numériques qui peuvent, selon les avancées technologiques, être améliorées (meilleure qualité, meilleure compression).

Etant donné le coût de réalisation (ressources financières et humaines) de ces reproductions, il est impératif de les préserver à long terme.

Quels formats d'archivage seront utilisés ?

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

Avec la possibilité de migrer vers un format plus adapté pour la préservation numérique (par exemple vers le JPG2000).

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 majorité de ces images sont soumise au droit d'auteur.

A ce stade, il n'est pas prévu de différencier les images libérées de leur droit et celles soumises au droit d'auteur. Par contre, leur mise à disposition est encouragée par notre base de données CATIMA.

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

Question not answered.

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

- Je ne sais pas encore