

---

# DMP de la plateforme "Estrades"

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

<b>Titre du plan</b>	DMP de la plateforme "Estrades"
<b>Version</b>	Version initiale
<b>Objet/périmètre du plan</b>	<p>Ce PGD est un descriptif général de la gestion par défaut des données au sein de la plateforme Estrades. La plateforme assure un soutien auprès de projets qui sont invités à élaborer leur propre PGD à partir de celui-ci.</p> <ul style="list-style-type: none"><li>• plateforme : désigne ici Estrades en tant qu'équipe de soutien à la recherche adossée à la MISHA ;</li><li>• projet : désigne le projet de recherche accompagné par la plateforme, pour lequel la rédaction d'un DMP est recommandée par la plateforme à défaut d'être demandé par les instances de financement.</li></ul>
<b>Domaines de recherche (selon classification de l'OCDE)</b>	History and archaeology, Computer and information sciences, Languages and literature
<b>Langue</b>	fra
<b>Date de création</b>	2021-10-05
<b>Date de dernière modification</b>	2023-06-09
<b>Licence</b>	Creative Commons Attribution 4.0 International
<b>Documents (publications, rapports, brevets, plan expérimental....), sites web associés</b>	<ul style="list-style-type: none"><li>• Page web : <a href="https://estrades.huma-num.fr/">https://estrades.huma-num.fr/</a></li><li>• Carnet hypothèses : <a href="https://estrades.hypotheses.org/a-propos">https://estrades.hypotheses.org/a-propos</a></li><li>• Gitlab : <a href="https://git.unistra.fr/estrades">https://git.unistra.fr/estrades</a></li></ul>

## Renseignements sur le projet

**Titre du projet** Estrades : Services de TRanscription, d'Analyse et de Diffusion pour l'Édition Structurée

**Acronyme** ESTRADES

**Résumé** ESTRADES est une plateforme proposant des solutions techniques, des services et de la formation visant à soutenir le développements de projets en humanités numériques autour de l'édition structurée. Par son nom, le projet sous entend la création d'une infrastructure *relativement simple mais suffisamment stable*, permettant une meilleure mise en avant des projets qu'elle soutient. Élaboré et porté au sein de l'unité de recherche ARCHE UR3400 depuis 2019, Estrades a également fait l'objet d'un financement de la fondation partenariale de l'UHA en 2022 (via le laboratoire ILLE) et réunit des EC et BIATSS de différentes composantes et laboratoires au sein d'un comité d'orientation depuis la même année. Depuis 2023, Estrades fait partie des plateformes soutenues par la MISHA.

**Date de début** 2022-05-19

**Partenaires**

- MISHA (200919794F)
- Université de Strasbourg ()
- université de Haute-Alsace ()
- Arts, civilisation et histoire de l'Europe (EA 3400 - UR 3400 depuis 01.1.2020) (200114208U)
- INSTITUT DE RECHERCHE EN LANGUES ET LITTÉRATURES EUROPÉENNE (ILLE) - UR 4363 (200918521X)

**Produits de recherche :**

1. Default research output (Jeu de données)

**Contributeurs**

Nom	Affiliation	Rôles
PHuN (Plateforme Humanités Numériques)	Maison Interuniversitaire des Sciences de l'Homme - Alsace (MISHA)	
Porte Guillaume	ARCHE UR3400	<ul style="list-style-type: none"> <li>• Coordinateur du projet</li> <li>• Personne contact pour les données</li> <li>• Responsable du plan de gestion de données</li> </ul>

**Droits d'auteur :**

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre

soumission.

# DMP de la plateforme "Estrades"

---

## 1. Description des données et collecte ou réutilisation de données existantes

### 1a. Comment de nouvelles données seront-elles recueillies ou produites et/ou comment des données préexistantes seront-elles réutilisées ?

La plateforme accompagne des projets de recherche en Sciences Humaines qui travaillent, notamment, sur l'édition de textes d'après des ouvrages ou des archives. La question de l'acquisition de ces données doit être prise en charge par les projets.

En ce qui concerne la réutilisation de codes sources ou l'utilisation de logiciels, la plateforme s'assure du respect des licences associées.

---

### 1b. Quelles données (types, formats et volumes par ex.) seront collectées ou produites ?

- Reproductions de documents (.jpg, .png, .pdf)
  - Transcriptions, commentaires (.xml, .txt)
- 

## 2. Documentation et qualité des données

### 2a. Quelles métadonnées et quelle documentation (par exemple méthodologie de collecte et mode d'organisation des données) accompagneront les données ?

- Utilisation du vocabulaire TEI (Text Encoding Initiative)
  - Métadonnées de description (teiHeader) basées sur les recommandations de la MRSH de Caen : <https://hal-rnmsh.archives-ouvertes.fr/hal-03830903v1>
  - Utilisation du protocole IIIF pour l'affichage des images
- 

### 2b. Quelles mesures de contrôle de la qualité des données seront mises en œuvre ?

La plateforme est compétente en matière d'utilisation du langage XML dans le cadre de l'édition structurée. Elle ne propose pas de validation concernant le contenu de la production, dont la validation reste à la charge de l'équipe du projet.

---

## 3. Stockage et sauvegarde pendant le processus de recherche

### 3a. Comment les données et les métadonnées seront-elles stockées et sauvegardées tout au long du processus de recherche ?

- GitLab Unistra : <https://git.unistra.fr>

- Seafire Unistra : <https://seafire.unistra.fr>

---

### 3b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long du processus de recherche ?

- Les documents traités sont essentiellement anciens et ne comportent pas de données sensibles.
- Le cas échéant, les projets sont invités à prendre contact avec le service DPO de l'Université de Strasbourg lors de l'élaboration du partenariat avec la plateforme.
- Par ailleurs, la plateforme ne stocke pas d'informations personnelles concernant ses utilisateurs. Cependant, les outils utilisés peuvent stocker certaines de ces informations (GitLab, Seafire, etc.) ; nous invitons à consulter leurs closes d'utilisation.

---

## 4. Exigences légales et éthiques, codes de conduite

### 4a. Si des données à caractère personnel sont traitées, comment le respect des dispositions de la législation sur les données à caractère personnel et sur la sécurité des données sera-t-il assuré ?

La plateforme ne traite pas directement de données à caractère personnel entrant dans le cadre de la législation (données historiques). Toutefois, le cas échéant, les porteur.eu.ses de projets devront contacter le ou la DPO de l'Université de Strasbourg afin de permettre à la plateforme la mise en place des protocoles permettant le respect des dispositions légales.

---

### 4b. Comment les autres questions juridiques, comme la titularité ou les droits de propriété intellectuelle sur les données, seront-elles abordées ? Quelle est la législation applicable en la matière ?

- Par défaut, toutes les données de recherche hébergées sont au format [CC BY-NC-SA 4.0](#).
- Les autres productions (code source) sont diffusées sous licence [CeCILL 2.1](#).

---

### 4c. Comment les éventuelles questions éthiques seront-elles prises en compte, les codes déontologiques respectés ?

Ces questions incombent au projet de recherche qui doit s'assurer de leur prise en compte en amont du travail avec la plateforme.

---

## 5. Partage des données et conservation à long terme

### 5a. Comment et quand les données seront-elles partagées ? Y-a-t-il des restrictions au partage des données ou des raisons de définir un embargo ?

Codes sources et contenus sont rendus disponibles à l'issue du projet via le GitLab de l'Unistra et l'entrepôt de données Nakala (voir ci-dessous)

---

### 5b. Comment les données à conserver seront-elles sélectionnées et où seront-elles préservées sur le long terme (par ex. un

entrepôt de données ou une archive) ?

Les contenus (fichiers XML TEI, images, photographie, transcriptions brutes en txt, etc.) sont déposées sur l'entrepôt [Nakala](#) et disponible de manière publique. Le DOI du dépôt est ensuite indiqué sur la page d'accueil du projet.

---

**5c. Quelles méthodes ou quels outils logiciels seront nécessaires pour accéder et utiliser les données ?**

- Les fichiers déposés sont enregistrés dans des formats textes ouverts (xml, csv, txt, json...) qui peuvent être ouverts par de simples éditeurs de texte.
- La reproduction du fonctionnement global des éditions est décrite dans la documentation de la plateforme sur le GitLab.

---

**5d. Comment l'attribution d'un identifiant unique et pérenne (comme le DOI) sera-t-elle assurée pour chaque jeu de données ?**

- DOI attribué pour chaque dépôt par Nakala
- DOI attribué la page d'accueil de chaque édition (contenant les métadonnées générales, les liens vers les fichiers XML-TEI, et vers les dépôts)

---

## **6. Responsabilités et ressources en matière de gestion des données**

**6a. Qui (par exemple rôle, position et institution de rattachement) sera responsable de la gestion des données (c'est-à-dire le gestionnaire des données) ?**

Chaque projet soutenu doit disposer de son/sa propre responsable pour la gestion des données.

Pour les données créées ou hébergées par la plateforme, celles-ci sont gérées par le coordinateur de la plateforme (G. Porte ARCHE UR3400).

---

**6b. Quelles seront les ressources (budget et temps alloués) dédiées à la gestion des données permettant de s'assurer que les données seront FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable) ?**

La gestion des données froides est prise en charge par la plateforme qui assure en interne leur caractère FAIR sur le long terme.