DMP du projet "ANR Access ERC Starting 2023 pour le projet HumaNs-MF"

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

Titre du plan	DMP du projet "ANR Access ERC Starting 2023 pour le projet HumaNs-MF"
Version	Version initiale
Objet/périmètre du plan	Le présent plan de gestion de données détaille le mode de production, conservation et protection des données qui seront produites et/ou utilisées dans le cadre du projet ANR Access ERC Starting 2023 "HumaNs-MF" pour la candidature au financement ERC Starting 2025.
Domaines de recherche (selon classification de l'OCDE)	Languages and literature
Langue	fra
Date de création	2024-03-18
Date de dernière modification	2024-03-29
Documents (publications, rapports, brevets, plan expérimental....), sites web associés	• Nom d'humain général personne en français médiéval : l'interface sociale de l'être humain : 10.2143/IG.180.0.3292822 • Nul homme en français médiéval : une grammaticalisation qui n’a jamais eu lieu : • Personne in ne-sentences from 14th to 16th centuries : a contribution of text-type calibrated corpora : • Les termes d’adresse dans les interactions ‘homme-femme’ sur un corpus de fabliaux en ancien français :

Renseignements sur le projet

Titre du projet	ANR Access ERC Starting 2023 pour le projet HumaNs-MF
Acronyme	HumaNs-MF
Résumé	L'objectif du projet ANR Access ERC Starting "Noms d'humain en français médiéval" (HumaNs-MF) est le dépôt d'une candidature compétitive pour le financement ERC Starting Grant 2025 (date limite prévisionnelle de dépôt - octobre 2024). L'activité de la porteuse du projet se construit donc autour de deux axes principaux : • la préparation du projet à soumettre et la rédaction de la proposition de candidature ERC Starting Grant • l'augmentation de l'impact scientifique par le biais de publication d'articles dans des journaux à impact et de participation aux événements scientifiques de diverse nature Par conséquent, l'exécution du projet prévoit essentiellement la production de données textuelles relatives au dossier de candidature, ainsi que de celles qui synthétisent les résultats de recherches individuelles de la porteuse du projet et sont destinées à la diffusion auprès de la communauté scientifique (journaux, événements scientifiques).
Sources de financement	• Agence Nationale de la Recherche : ANR-23-AERC-0007
Date de début	2023-10-01
Date de fin	2025-09-30

Produits de recherche :

1. Default research output

Contributeurs

Nom	Affiliation	Rôles
Geylikman Zinaida		• Personne contact pour les données • Responsable du plan
GEYLIKMAN Zinaida	Université de Paris - 200615317C	• Coordinateur de projet

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 du projet "ANR Access ERC Starting 2023 pour le projet HumaNs-MF"

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 production des données textuelles du projet sera réalisée par la porteuse du projet.

Il s'agira, en premier lieu, du texte de la proposition de candidature ERC Starting Grant structuré selon les normes de rédaction définies sur la page Web dédiée aux financements de la Commission Européenne (EU Funding & Tenders Portal). Par ailleurs, ces données textuelles incluront les publications de recherches individuelles de la porteuse du projet, ainsi que les propositions de communication pour des événements scientifiques (pour un descriptif détaillé des données produites, voir 1b.)

Les résultats des recherches publiés s'appuient sur les données textuelles extraites des dictionnaires et bases textuelles issues de numérisations ou encodées selon diverses pratiques du TAL (annotation syntaxique, lemmatisation). Il s'agit, par exemple, de ressources suivantes :

Guillot-Barbance, Céline, Heiden, Serge et Lavrentiev, Alexei (2017), « Base de français médiéval : une base de référence de sources médiévales ouverte et libre au service de la communauté scientifique », *Diachroniques*, n 7, pp.168-184. halshs-01809581

DMF : *Dictionnaire du Moyen Français*, version 2023 (DMF 2023). ATILF - CNRS & Université de Lorraine. Site internet : http://www.atilf.fr/dmf

Larrivée, Pierre (dir.), 2023, *corpus HIGH-TECH*, v. 1.4, Caen (France), CRISCO (EA 4255).

Larrivée, Pierre et Poletto, Cecilia (dirs), 2023, *corpus MICLE-PREVIEW*, v. 1.0, Caen (France) / Francfort (Allemagne), CRISCO (UR 4255) / Institut für Romanische Sprachen und Literaturen.

L'utilisation des données textuelles et textométriques des ressources ci-dessus dans les articles scientifiques est conditionnée à la citation de la ressource en question dans la Bibliographie ; cette condition sera scrupuleusement respectée. D'autres ressources numériques pourront être utilisées en fonction des besoins de recherche.

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

Dans le cadre du projet seront produits des groupes de données suivants :

1. Données textuelles : articles scientifiques synthétisant les résultats des recherches individuelles de la porteuse du projet qui seront publiés dans les journaux à impact. A ce stade, un article a paru dans l'*Information grammaticale*, un article est accepté par *Zeitschrift für Romanische Philologie,* deux articles sont en préparation. Les versions de travail des articles seront en format .docx, les versions définitives, ainsi que les versions "épreuves" transmises par les journaux, également en format .pdf. Certains articles incluront des figures en format .png.
2. Données textuelles : propositions de communication pour la participation aux événements scientifiques (congrès, colloques, journées d'études *etc.*) en format .docx et .pdf, diaporamas accompagnant les communications orales en format .pptx
3. Données textuelles : diaporamas .pptx présentant la conception du projet à proposer à ERC Starting Grant destinés aux consultations auprès des collègues en TAL, notes en format .txt synthétisant les remarques et suggestions des collègues.
4. Données textuelles : texte de candidature pour le financement ERC Starting Grant. Les versions de travail seront produites en format .docx et pdf. et incluront des fichiers en format .xlsx et .png.
5. Données textuelles : tableau .xlsx récapitulant les diverses activités de la porteuse du projet (publications, participation aux événements scientifiques, consultations auprès des collègues, rédaction de la proposition de candidature ERC Starting Grant) et le stade de leur réalisation. Il s'agit d'un document à usage interne à destination de la porteuse et de la superviseuse du projet.
6. Bibliographie Zotero pour gérer les références bibliographiques sur les sujets abordés dans le projet, maintenue par la porteuse du projet.

---

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 ?

La porteuse du projet respectera les recommandations de nommage des fichiers par l'Université Paris Cité disponibles via le lien https://cloud.u-paris.fr/s/8nLQi9mTZiEqo4A.

Ainsi, le nom du fichier devra inclure toutes ou une partie des métadonnées suivants : sujet/domaine/nature_typedocument_AAAAMMJJ_version_auteur/réviseur

Par exemple, la deuxième version de l'article consacré au nom *personne* en français médiéval à soumettre à l'*Information grammaticale* revu le 05 janvier 2024 par la superviseuse du projet Angelina Aleksandrova sera nommée ainsi : *IG_personneFM_20240105_V2_revuAA*

Sur l'ordinateur personnel de la porteuse du projet, ainsi que sur le NextCloud à usage interne avec accès réservé uniquement à la porteuse et à la superviseuse du projet, les documents seront organisés selon les recommandations de l'Université Paris Cité disponibles via le lien https://cloud.u-paris.fr/s/k6nwS6G8oaMomZz.

Ainsi, les documents seront placés dans le dossier ANR_Access_ERC_2024 et répartis en plusieurs dossiers selon les diverses activités du projet. Voici un exemple du plan de classement :

1. Candidature : 1.1 Partie_A. 1.2. Partie_B. 1.3. Budget
2. Publications : 2.1. *Information_grammaticale* 2.2. *ZfRP* etc. (noms des journaux)
3. Colloques : 3.1. *Trace_Curve_03_2024 3.*2. *Diachro2024* etc. (noms des colloques)
4. Consultations : 4.1. Caen_CRISCO 4.2. Heidelberg_ALMA etc. (ville de l'Université et laboratoire)

Le plan de classement pourra inclure jusqu'à quatre niveaux. Ainsi, le chemin pour aller à la deuxième version de l'article nommée ci-dessus serait :

ANR_Access_ERC_2024/Publications/Information_grammaticale/Revision

Toutes les versions de travail apparaissant dans le processus de l'exécution du projet seront systématiquement sauvegardées.

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

La qualité des productions des données textuelles sera soumise à la validation intellectuelle, une validation experte par la porteuse du projet et par la superviseuse du projet.

Dans le cas d'articles publiés dans les journaux à impact, la qualité scientifique sera également validée par la révision par les pairs assurée par la rédaction des journaux en question. Dans le cas de propositions de communication, la qualité scientifique sera également validée par le comité scientifique des événements en question.

---

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 ?

Les données collectées et produites dans le cadre du projet seront stockées et sauvegardées sur trois supports : l'ordinateur personnel de la porteuse du projet, le NextCloud de l'Université Paris Descartes associé au compte institutionnel de la porteuse du projet et un disque dur externe maintenu par la porteuse du projet.

Les données relatives au projet seront synchronisées automatiquement entre l'ordinateur personnel et le NextCloud ; à cela s'ajouteront des sauvegardes hebdomadaires sur un disque dur externe.

Une sauvegarde mensuelle de la bibliographie Zotero sur l'ordinateur personnel sera également effectuée.

3b. Comment la sécurité des données et la protection des données sensibles seront-elles assurées tout au long

du processus de recherche ?

L'ordinateur personnel de la porteuse du projet est protégé par un mot de passe modifié régulièrement (rappel automatique par Windows).

L'accès du NextCloud est restreint ; uniquement la porteuse et la superviseuse du projet peuvent visualiser les données qui y sont présentes. Dans le cas de partage de certains documents avec d'autres collègues, l'accès leur sera ouvert via un lien spécifique.

Les fichiers contenant les données relatives au dossier de candidature ERC Starting Grant - diverses versions des textes de proposition, tableau xlsx. synthétisant le budget prévisionnel *etc.* - seront considérés comme confidentiels et seront, par conséquent, protégés par un mot de passe. Par ailleurs, jusqu'à ce que la procédure de candidature ERC Starting Grant ne soit terminée, ils ne seront en aucun cas envoyés par email "en clair", mais seront partagés via NextCloud comme décrit ci-dessus. Dans certains cas, un envoi via l'application Renater sera également envisageable.

Le projet ne prévoit pas d'utilisation de données sensibles.

---

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 ?

A ce stade, il n'est pas prévu d'utiliser de données à caractère personnel dans le cadre du projet. Dans le cas éventuel de l'organisation d'un événement scientifique (journée d'étude, colloque), la porteuse du projet respectera le Règlement Général sur la Protection des Données (RGPD) en suivant les recommandations de la Déléguée à la protection des données de l'Université Paris Cité.

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 ?

Les données produites dans le cadre du projet relèveront uniquement de l'activité intellectuelle de la porteuse du projet.

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

Le projet sera mené selon le principe du respect de l'intégrité scientifique.

---

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 ?

Les publications scientifiques résultant du projet seront diffusés par les journaux respectifs suivant les modalités propres à la politique éditoriale de chaque journal en particulier. Il est prévu qu'un à deux articles soient diffusés en OpenAccess. Par ailleurs, d'autres contributions pourront être diffusés via les réseaux sociaux scientifiques tels que Academia ou ResearchGate, dans le respect des restrictions imposées par les journaux éditeurs.

Les résultats des recherches individuelles pourront également être présentées lors d'événements scientifiques de diverse nature (séminaires, journées d'étude, colloques, congrès) dans le cadre imposé par les événements en question

(présentation orale ou par poster).

A ce stade, il n'est pas prévu que le dossier de candidature pour le financement ERC Starting Grant soit partagé en libre accès, même une fois que la procédure de candidature sera terminé. Toutes les données relevant du dossier de candidature seront destinées à un usage interne par la porteuse et la superviseuse du projet, ainsi que d'autres collègues de l'Université Paris Cité et d'autres établissements académiques européens qui apporteront leur conseil à la porteuse du projet.

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 publications scientifiques dans les journaux ont vocation à représenter des données pérennes, la pérennité étant, en principe, garantie par l'éditeur. La conservation des versions "épreuves MEP" des articles sur un disque dur externe dédié est toutefois assurée par la porteuse du projet.

Les décisions sur la conservation et la préservation des données relatives au projet de candidature ERC Starting Grant seront prises plus tard et dépendront de l'issue de la candidature.

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

Le mode d'accès aux publications scientifiques des résultats de recherches dépendra des pratiques éditoriales des journaux en question (OpenAccess/accès restreint en ligne/édition papier). A ce stade, en dehors des publications qui pourront être citées, il n'est pas prévu que les données produites dans le cadre du projet ANR Access ERC Starting HumaNs-MF soient réutilisées.

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

A ce stade, il n'est pas prévu de diffuser les données du projet en dehors des articles scientifiques. Le PID sera attribué à ceux-ci automatiquement par les journaux au moment de la publication.

---

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) ?

La porteuse du projet sera l'unique responsable de la gestion des données.

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) ?