
DMP du projet "Education & Research to Advance Societal Missions through Engagement"

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 "Education & Research to Advance Societal Missions through Engagement"
Version	Version initiale
Objet/périmètre du plan	Ce plan de gestion de données (PGD) est la version initiale du PGD demandé par l'ANR, dans le cadre du financement attribué au projet ERASME en réponse à l'appel à projets ExcellencES (ANR-21-EXES-0009). Ce document, produit à partir de l'outil DMP OPIDoR, sera amené à évoluer au fur et à mesure que les activités du projet seront mises en place et que les données seront produites.
Domaines de recherche (selon classification de l'OCDE)	
Langue	fra
Date de création	2023-02-01
Date de dernière modification	2023-03-22

Renseignements sur le projet

Titre du projet	Education & Research to Advance Societal Missions through Engagement
Acronyme	ERASME
Résumé	<p>Au cours des vingt dernières années, la compétition entre les universités au niveau mondial s'est fortement accrue, conduisant à un écart croissant entre les institutions dites de recherche intensive et celles se concentrant sur l'enseignement de premier cycle et l'enseignement professionnel. A l'Université Paris-Est Créteil, nous sommes convaincus que d'autres options sont envisageables par la combinaison d'une formation centrée sur l'étudiant, d'une recherche de pointe et d'un engagement fort au service des besoins du territoire et de ses communautés. Nous visons ainsi un modèle alternatif d'université engagée.</p> <p>L'impact dramatique qu'a eu la pandémie a remis l'accent sur la nécessité d'un nouveau modèle d'université plaçant une recherche de niveau mondial au service de ceux en ayant le plus besoin. Notre modèle d'université engagée confronte les principaux défis sociétaux, non seulement en tant qu'objets d'études mais également en tant que composantes intégrantes de notre institution, en engageant à la fois les</p>

communautés extérieures à l'université mais également nos communautés internes pour identifier les problématiques auxquelles nous devons répondre et pour co-concevoir les réponses à y apporter. Nous sommes convaincus que ce modèle d'université engagée, construit tout au long de nos 50 ans d'existence, a un véritable potentiel de transformation, non seulement pour l'UPEC mais également pour d'autres universités en France et à travers le monde. Mais nous sommes également conscients que cette approche requiert une profonde transformation de nos modes de fonctionnement, d'interaction et de planification. Cette transformation est au coeur du projet ERASME - éducation & recherche pour faire avancer les missions sociétales par l'engagement.

La première étape de notre processus de transformation a été définie en 2019 dans notre Projet Stratégique portant notre vision. Celle de "devenir une université engagée de référence, centrée sur les défis de transformation sociale et environnementale et les aspirations à plus de justice sociale et d'équité".

Aujourd'hui, grâce à l'opportunité que nous offre l'appel à projets ExcellencES, nous poursuivons notre transformation en répondant à quatre objectifs majeurs qui nous permettront de renforcer notre lien avec le territoire :

- Repenser la formation ;
- Focaliser et partager la recherche ;
- Contribuer à une société plus juste ;
- Engager nos partenaires et communautés.

Pour concrétiser la réalisation de ces objectifs, deux conditions préalables devront être mises en œuvre :

- Engager nos personnels et étudiants ;
- Renforcer nos services support.

En lien avec ces objectifs et conditions, nous avons identifié une série d'actions phares nous conduisant chacune à nous transformer. Nous cherchons à transformer : notre fonctionnement (en impliquant et en engageant nos communautés internes comme externes), la réalisation de nos missions (en ayant un impact social positif au-delà du simple impact scientifique), nos prises de décisions (en impliquant nos partenaires principaux et en prenant en compte les besoins du territoire). Grâce à ERASME, nous pourrions faire participer d'autres universités et institutions au niveau national et international à un dialogue sur la façon dont les écosystèmes d'enseignement supérieur et de recherche d'un territoire peuvent tirer profit d'interactions entre des universités aux rôles et aux modèles différents. Nous cherchons également à renforcer notre position d'université engagée, en encourageant d'autres universités, en France et ailleurs, à partager leurs bonnes pratiques en la matière. En prenant en considération les caractéristiques de notre territoire et de l'écosystème d'ESRI francilien, nous sommes dans une position parfaite pour accueillir et encourager ces réflexions et changements à venir.

Notre mission est simple mais ambitieuse : nous souhaitons être, dans 10 ans, l'université engagée de référence pour notre territoire et en France.

Sources de financement

- Agence Nationale de la Recherche : ANR-21-EXES-0009
- Université Paris Est-Créteil Val-de-Marne :
- Agence Nationale de la Recherche : ANR-21-EXES-0009

Date de début 2022-05-01

Date de fin 2032-04-30

Partenaires

- Centre national de la recherche scientifique ()
- Institut national de la santé et de la recherche médicale ()
- INSERM Délégation Paris IDF Centre-Est ()
- CNRS Paris Villejuif ()

Produits de recherche :

1. WP2 - Focaliser et partager la recherche - Observatoire de l'impact de la recherche (Jeu de données)
2. WP2 - Focaliser et partager la recherche - Graduate Programs (Jeu de données)
3. WP4 - Contribuer à une société plus juste - Appel à projets transformant (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
DUBOIS-RANDE Jean-Luc - 0000-0003-4968-5728	UPEC	
DUBOIS-RANDÉ Jean-Luc	Université de Paris XII (Val de Marne - Paris Est Créteil)	<ul style="list-style-type: none"> • Coordinateur du projet
Kenadid Ilyas		<ul style="list-style-type: none"> • Personne contact pour les données (WP4 - AAP , WP2 - Recherche, WP2 - Recherche)
NOEL Marie-Hélène	UPEC	<ul style="list-style-type: none"> • Responsable du plan de gestion de données

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 "Education & Research to Advance Societal Missions through Engagement"

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 ?

Les données de recherche sont produites par des équipes de l'Université Paris-Est Créteil et de ses partenaires (CNRS, INSERM), qui sont financées dans le cadre du projet ERASME. Les données seront produites dans le cadre d'objets financés tels que:

- un appel à projets annuel à forte valence interdisciplinaire ;
- des graduates programs
- un contrat post-doctoral préparatoire à l'observatoire de l'impact de la recherche.

Les PGD individuels mis en place par chacun des responsables des projets financés doivent documenter :

- Le mode de recueil des données (dépouillements de sources primaires, enquêtes, mesures de terrain, expérimentations, analyses, modélisation, etc.)
- Les arbitrages technologiques effectués, qui devront être justifiés, pour chaque catégorie de données, en regard des possibilités : état de l'art et ressources dont dispose le projet (financières, techniques, humaines)
- L'éventuelle ré-utilisation de données préexistantes, avec mention impérative des sources d'où proviennent ces données, leur durée et conditions d'accessibilité, ainsi que leur lieu de stockage

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

Les projets de recherche sélectionnés et financés dans le cadre d'ERASME pouvant être de nature variée, les données produites sont également plurielles, par exemple : données de dépouillements, données d'enquêtes, données statistiques, données démographiques, données géophysiques, données bibliométriques, données photographiques, etc.

La caractérisation de ces données doit être documentée dans chacun des PGD individuels, avec une attention particulière apportée, pour chacun des jeux de données :

- à l'utilisation d'un nom unique et explicite permettant à la direction exécutive du projet d'en opérer un suivi efficient;
- à la mise en œuvre d'outils et de standards pertinents (modernes, interopérables) pour chaque catégorie de données produites ou mobilisées (ex. données géomatiques, édition scientifique de sources, bases de données relationnelles, graphes de connaissances, données statistiques, données pour l'analyse de réseaux, sources numérisées, corpus bibliographique, thesaurus, etc.)
- à la finalité de la gestion des données : production de connaissances, science reproductible, préservation de données à valeur patrimoniale, diffusion auprès du public, etc.
- au périmètre thématique (santé, sciences et techniques, lettres et sciences humaines...)
- au périmètre géographique
- à la nature des données (dépouillement de sources primaires, transcription d'entretiens, relevés de fouilles, relevés pédologiques, relevés de températures, images satellites, etc.)
- aux types de données (numérique, textuel, image, audio, modèle, ressource interactive, etc.)
- aux formats techniques (formats ouverts à privilégier)
- au volume de stockage requis pour ces données

Pour garantir l'accès aux données et leur pérennité, chacun des projets financés doit privilégier l'utilisation de formats standards et ouverts. Se référer lorsque possible, pour le choix du format de fichier, au récapitulatif élaboré par DORANum sur les formats ouverts et fermés.

Pour les projets ne pouvant répondre aux exigences de Science ouverte, des précisions complémentaires seront apportées.

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 ?

Les jeux de données issus des projets de recherche financés par ERASME sont de différentes natures, types et formats. C'est pourquoi, afin d'assurer une description uniforme, dans le but de faciliter le recensement des données, leur identification et leur réutilisation

par la communauté scientifique, chaque responsable de projet financé doit fournir une notice de métadonnées en respectant la trame fournie ultérieurement par la direction exécutive ERASME.

La compilation de ces informations, fournies par chacun des responsables de projets financés, permet d'établir un catalogue global des données de recherche produites dans le cadre de ERASME, au travers de l'ensemble des projets individuels de recherche.

Sont ainsi récapitulées, pour chacun des jeux de données :

- l'identifiant unique du jeu de données, de type DOI
- le nom du fichier de données, unique et explicite
- les finalités scientifiques de la collecte et de la conservation de ces données
- le périmètre thématique
- la nature des données
- le format des données
- le type de fichier correspondant
- la personne ou l'équipe contact, responsable de ces données
- le lieu de stockage du jeu de données
- les droits d'accès
- les principales publications associées au jeu de données (avec leur DOI)
- la documentation éventuellement disponible (en plus du fichier de type "Lisez-moi" présenté ci-dessous)

En plus des métadonnées, chacun des responsables de projets financés doit fournir un fichier de type "Lisez-moi" (au format .txt) comportant une présentation du jeu de données et la description de son contenu, incluant :

- l'identifiant unique du jeu de données, de type DOI
- le nom du fichier de données, unique et explicite
- l'ensemble des éléments pertinents pour la compréhension scientifique du contexte de production ou de collecte des données
- la méthodologie de collecte ou de production des données
- une documentation appropriée à la nature des données : par exemple, dans le cas des données statistiques, une documentation de type "codebook" explicitant les codes, variables, unités, etc.
- la méthode d'analyse et de traitement des données et toute autre information s'avérant nécessaire à la compréhension des données, comme par exemple :
- les éléments liés aux traitements de données personnelles, qui documentent la fiche du registre des activités de traitement,
- la présentation des procédures d'anonymisation et/ou de pseudonymisation de données personnelles,
- la mention, lorsque nécessaire, du logiciel spécifique à utiliser pour la visualisation et/ou le traitement des données (avec indication du site où télécharger le dit-logiciel - Voir points 3a et 5c.)

La mise à disposition de ce fichier "Lisez-moi" se fait également à travers le formulaire de dépôt des métadonnées produit ultérieurement comme précisé plus haut.

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

Les jeux de données sont produits par différentes équipes de recherche, dans le cadre de leurs activités et sujets respectifs. Par conséquent, la qualité de chaque jeu de données est du ressort des responsables respectifs de chacune de ces équipes, tant pour la description de ces données que pour les données elles-mêmes

Les outils utilisés et les procédures de contrôles mis en place doivent être précisés dans les PGD de chacun des projets financés.

Les équipes de recherche pourront bénéficier de l'appui du pôle "management des données" de la Direction de la transformation de l'UPEC.

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 seront stockées et sauvegardées sur les ordinateurs configurés par l'établissement et sur un espace sécurisé.
Si des projets de recherche en santé sont financés par le programme ERASME, nous hébergerons les données sur un serveur certifié "hébergement de données de santé".
En cas de nécessité, un processus d'archivage intermédiaire sera prévu pour les projets le nécessitant.

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

Si des projets de recherche en santé sont financés par le programme ERASME, nous hébergerons les données sur un serveur certifié "hébergement de données de santé".
Si d'autres données sensibles (opinions religieuses, politiques; origines ethniques etc.) sont traitées lors de projets de recherche nous mettrons en place les sécurités techniques et organisationnelles nécessaires à leur protection.
Les postes de travail nécessitent une authentification et peuvent se connecter via un VPN. Des mises à jour régulières des antivirus et des programmes sont réalisées. Les salles serveurs sont accessibles par accès restreint. Nous disposons de firewalls.
Si des données personnelles venaient à être traitées, un chiffrement des fichiers sera mis en place.

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 ?

Si des données personnelles venaient à être traitées, un chiffrement des fichiers sera mis en place.
Nous préconiserons la pseudonymisation des données personnelles pour limiter les risques. Une information des personnes concernées sera réalisée précisant leurs droits en matière de protection des données. Un consentement sera éventuellement demandé si nécessaire.
Un processus de réponse aux demandes d'exercice de droit sera mis en place afin de respecter les délais réglementaires de réponse. Dans les cas qui le nécessiteraient, une analyse d'impact de protection des données sera effectuée.
Les postes de travail nécessitent une authentification et peuvent se connecter via un VPN. Des mises à jour régulières des antivirus et des programmes sont réalisées. Les salles serveurs sont accessibles par accès restreint. Nous disposons de firewalls.

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 ?

Dans le cas où plusieurs équipes partenaires sont impliquées dans la genèse des jeux de données, les droits de chacune de ces équipes sur les données et le matériel doivent être spécifiés dans l'accord de consortium s'il existe, et/ou à défaut dans le plan de gestion de données individuel correspondant au projet en question. Doivent être notamment spécifiés : les droits de propriété intellectuelle sur les données, qui en contrôle l'accès, les conditions d'accessibilité, les éventuelles restrictions appliquées, etc.
Afin de s'assurer si les jeux de données nécessitent ou non des formalités spécifiques liées par exemple à l'Accès et au Partage des Avantages (Protocole de Nagoya), ou à la nature spécifique de certains matériels, la Direction de la recherche et de la Valorisation de l'UPEC peut être sollicité lorsque nécessaire : direction-recherche@u-pec.fr
Les membres des équipes des projets financés par ERASME s'engagent à respecter la législation en vigueur tout au long du projet (RGPD, droit à l'image, propriété intellectuelle, etc.) et à accompagner la diffusion et le dépôt de leur jeu de données d'un contrat de licence. Les licences Etalab ou Creative Commons (CC-BY) sont à privilégier, avec un niveau d'ouverture de la licence CC à adapter aux données produites et au niveau d'ouverture souhaité. (Possibilité d'ajouter si nécessaire deux contrats de licence.)

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

Chaque équipe de recherche financée dans le cadre de ERASME doit s'interroger sur les données qu'elle produit et sur les problèmes d'éthique ou de protection de la confidentialité pouvant éventuellement se poser (données sensibles, données de santé, données personnelles, etc.).

Pour évaluer les situations au cas par cas, chaque responsable de projet peut, préalablement à l'élaboration de son plan de gestion de données individuel, prendre l'attache du Comité d'éthique, de déontologie et d'intégrité scientifique (CEDIS) ou des comités d'éthique et de la recherche (CER) d'autres universités via la fédération des CER, afin de veiller à prendre les mesures nécessaires de protection en terme de récolte de données, stockage, transfert et conditions d'accès. Celles-ci doivent alors être explicitement détaillées dans les PGD individuels des projets financés. Dans le cas où les jeux de données produits ne soulèvent aucun problème d'éthique, cela doit également être spécifié dans les PGD individuels.

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 données seront soumises au principe FAIR (Findable, Accessible, Interoperable, Reusable). Dans les conditions de l'article 30 de la Loi «Pour une République numérique», les responsables des projets financés s'engagent par ailleurs à déposer les publications scientifiques (texte intégral) issues de leur projet de recherche, de développement ou d'innovation sur le portail HAL-UPEC sauf nécessité impérieuse de confidentialité.

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 équipes de chaque projet financé par le programme ERASME devra définir ses propres règles en matière de conservation et de préservation des données sur le long terme.

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

Chaque PGD individuel viendra préciser les méthodes et outils logiciels nécessaires pour accéder et utiliser les données.

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

L'attribution d'un DOI pour chaque jeu de données est en cours de réflexion à l'échelle de l'établissement.

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) ?

Les porteurs de projets devront définir dans les PGD individuels les responsables en matière de gestion des données. La Direction de la transformation sera en charge de la coordination du programme ERASME pour l'UPEC et assure la responsabilité de la mise en œuvre et de la révision période du présent PGD cadre.

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) ?

Les financements alloués dans le cadre du programme ERASME ont permis le recrutement et la structuration d'un pôle "management de la donnée" au sein de la Direction de la transformation. Le pôle est constitué de 3 personnes dont une DPO, responsable du pôle.