
DMP du projet "Mettre le plan de gestion des données automatisé dans les mains des biologistes"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe : modèle structuré" fourni par Science Europe.

Renseignements sur le plan

Titre du plan	DMP du projet "Mettre le plan de gestion des données automatisé dans les mains des biologistes"	
Domaines de recherche (selon classification de l'OCDE)		
Langue	fra	
Date de création	2023-04-18	
Date de dernière modification	2023-07-19	
Identifiant	https://opidor-preprod.inist.fr/plans/2943	
Type d'identifiant	URL	
Licence	Nom	Etalab Open License 2.0
	URL	http://spdx.org/licenses/etalab-2.0.json

Renseignements sur le projet

Titre du projet Mettre le plan de gestion des données automatisé dans les mains des biologistes

Acronyme maDMP4LSn

Résumé L'Institut Français de Bioinformatique (IFB), avec ses deux infrastructures informatiques et ses 30 plates-formes, est une structure essentielle pour les Sciences de la Vie et ses applications (santé, agriculture, environnement), fournissant un environnement de production, d'analyse et de gestion de données pour les communautés d'utilisateurs. Bien que ces communautés puissent être considérées comme bien dotées d'outils de structuration et de gestion des données, des problèmes subsistent pour mettre pleinement en œuvre les bonnes pratiques et gérer les données, de leur production à leur préservation, en assurant leur accessibilité, leur reproductibilité et leur réutilisation. Le Plan de Gestion des Données (PGD ou DMP pour Data Management Plan) est considéré comme un élément-clé qui facilitera la mise en œuvre des principes FAIR (Findable, Accessible, Interoperable, Reusable). Le DMP contient les éléments qui guident le processus de gestion des données tout au long de leur cycle de vie et fournissent également des traces de la provenance des données (inclus tous les résultats de recherche dérivés des données brutes). L'Inist-CNRS propose l'outil Web DMP-OPIDoR qui facilite la rédaction des DMP et contribue également à l'harmonisation des bonnes pratiques en fournissant conseils, exemples et modèles spécifiques de certaines institutions. L'évolution du DMP OPIDoR vers un DMP automatisable (« machine-actionable DMP ») est actuellement à l'étude sur la base des retours d'expérience des utilisateurs, des cas collectés et des travaux de RDA sur le DMP actif. L'outil DMP OPIDoR « machine actionable » sera conforme au modèle commun RDA DMP mais inclura également une extension afin de servir les différents acteurs et besoins disciplinaires validés par un large groupe d'utilisateurs.

Sources de financement

- Agence Nationale de la Recherche : ANR-19-DATA-0017

Date de début 2023-04-19

Date de fin 2023-04-20

Partenaires

- Institut de l'information scientifique et technique

Produits de recherche :

1. jeu 1 (Modèle)
2. jeu de données 2
3. New research output 3

Contributeurs

Nom	Affiliation	Rôles
Busin Anne	Inist Institut de l'information - 198822A	<ul style="list-style-type: none"> • Coordinateur de projet • Personne contact pour les données (Research Output 3) • Responsable du stockage des données (jd1)
Cosserat Françoise		<ul style="list-style-type: none"> • Personne contact pour les données (jd1, jd 2) • Responsable du plan
Jacquemot Marie-Christine		
Rassinoux Laurent		<ul style="list-style-type: none"> • Responsable de la production ou de la collecte des données (jd1)

Budget

Type de coût (Titre)	Montant	Etape du cycle de vie
Autre (cout réutilisation donnees jd1)	100 EUR	Coûts liés à la réutilisation de données existantes - jd1

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 "Mettre le plan de gestion des données automatisé dans les mains des biologistes"

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

jeu 1

1.1 Description générale du produit de recherche

Nom	jeu 1
Description	nnnn
Type	Modèle
Thésaurus INRAE	<ul style="list-style-type: none">• ADN complémentaire (Thésaurus INRAE)
Subject	Earth and Environmental Sciences, Engineering
Langue	fra

1.2 Est-ce que des données existantes seront réutilisées ?

Données réutilisées	<ul style="list-style-type: none">• Nom complet jd 1 :
Coûts	<ul style="list-style-type: none">• Autre : 100 EUR

1.3 Comment seront produites/collectées les nouvelles données ?

Equipements, plateaux techniques utilisés	<ul style="list-style-type: none">• AGATA : https://cat.opidor.fr/index.php/AGATA
--	---

jeu de données 2

1.1 Description générale du produit de recherche

Nom	jeu de données 2
------------	------------------

1.2 Est-ce que des données existantes seront réutilisées ?

1.3 Comment seront produites/collectées les nouvelles données ?

New research output 3

1.1 Description générale du produit de recherche

Nom New research output 3

Mots clés (texte libre)

1.2 Est-ce que des données existantes seront réutilisées ?

Question sans réponse.

1.3 Comment seront produites/collectées les nouvelles données ?

Question sans réponse.

2. Documentation et qualité des données

jeu 1

2.1 Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Question sans réponse.

2.2 Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Question sans réponse.

jeu de données 2

2.1 Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Question sans réponse.

2.2 Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Question sans réponse.

New research output 3

2.1 Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Question sans réponse.

2.2 Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Question sans réponse.

3. Exigences légales et éthiques, code de conduite

jeu 1

3.2 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 ?

Question sans réponse.

3.3 Quels sont les aspects éthiques à prendre en compte lors de la collecte des données ?

Question sans réponse.

jeu de données 2

3.2 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 ?

Question sans réponse.

3.3 Quels sont les aspects éthiques à prendre en compte lors de la collecte des données ?

Question sans réponse.

New research output 3

3.2 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 ?

Question sans réponse.

3.3 Quels sont les aspects éthiques à prendre en compte lors de la collecte des données ?

Question sans réponse.

4. Traitement et analyse des données

jeu 1

4.1 Comment et avec quels moyens seront traitées les données ?

Question sans réponse.

jeu de données 2

4.1 Comment et avec quels moyens seront traitées les données ?

Question sans réponse.

New research output 3

4.1 Comment et avec quels moyens seront traitées les données ?

Question sans réponse.

5. Stockage et sauvegarde des données pendant le processus de recherche

jeu 1

5.1 Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Description	blabla jd1
Identification de la ressource	<ul style="list-style-type: none"> • Mésocentre Clermont Auvergne : https://cat.opidor.fr/index.php/Mésocentre_Clermont_Auvergne • GenOuest : https://cat.opidor.fr/index.php/GenOuest
Politique de sauvegarde	<ul style="list-style-type: none"> • L'espace utilisateur sur le cluster de calcul est un espace de stockage non sauvegardé. Seuls les espaces de stockage en mode "projet" sont sécurisés. Le nombre de réplicats est défini avec le responsable du projet et est au maximum de 3. Les réplicats sont tous localisés physiquement dans le même datacentre. Les modes d'accès aux données proposés sont un tableau de bord (interface graphique) et un accès en ligne de commande pour une connexion au cluster de calcul avec les clients comme goofys et s3cmd. L'accompagnement des utilisateurs se fait par le système de gestion de tickets Jira. • Des snapshots sont créés toutes les heures sur les « homes ». Ceux-ci sont conservés pendant 5 semaines. Sur les répertoires projets, des snapshots sont créés tous les jours et conservés pendant 5 semaines. Aucun snapshot n'est réalisé sur les espaces de calcul.
Mesures de sécurité	L'espace utilisateur sur le cluster de calcul est un espace de stockage non sauvegardé. Seuls les espaces de stockage en mode "projet" sont sécurisés. Le nombre de réplicats est défini avec le responsable du projet et est au maximum de 3. Les réplicats sont tous localisés physiquement dans le même datacentre. Les modes d'accès aux données proposés sont un tableau de bord (interface graphique) et un accès en ligne de commande pour une connexion au cluster de calcul avec les clients comme goofys et s3cmd. L'accompagnement des utilisateurs se fait par le système de gestion de tickets Jira.

jeu de données 2

5.1 Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Identification de la ressource	<ul style="list-style-type: none"> • GenOuest : https://cat.opidor.fr/index.php/GenOuest
Politique de sauvegarde	<ul style="list-style-type: none"> • Des snapshots sont créés toutes les heures sur les « homes ». Ceux-ci sont conservés pendant 5 semaines. Sur les répertoires projets, des snapshots sont créés tous les jours et conservés pendant 5 semaines. Aucun snapshot n'est réalisé sur les espaces de calcul.

New research output 3

5.1 Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Question sans réponse.

6. Partage des données et conservation à long terme

jeu 1

6.1 Comment les données seront-elles partagées ?

Question sans réponse.

6.2 Comment les données seront-elles conservées à long terme ?

Question sans réponse.

jeu de données 2

6.1 Comment les données seront-elles partagées ?

Question sans réponse.

6.2 Comment les données seront-elles conservées à long terme ?

Question sans réponse.

New research output 3

6.1 Comment les données seront-elles partagées ?

Question sans réponse.

6.2 Comment les données seront-elles conservées à long terme ?

Question sans réponse.