

Gestion des données : DMP OPIDoR et le cycle de vie des données

Paulette Lieby [0000-0002-9289-9652](tel:0000-0002-9289-9652)

DMP OPIDoR

***Attention : la version présentée ici est celle du 12/11/21 :
V2.3.4***

***Très bientôt (mi-novembre 2021) la version maDMP va
être mise en production et sera donc celle avec laquelle
vous travaillerez.***

***Cette version sera présentée dans un des modules de
cette formation.***

Gestion des données : DMP OPIDoR, un outil d'aide au PGD

Complément au module 4

Fair Data formation 15-18 novembre 2021

Renseignements sur le projet

Produits de recherche

Modèle choisi

Rédiger

Partager

Télécharger

tout développer | tout réduire

4. Exigences légales et éthiques, codes de conduite (3 questions)



DMP OPIDoR <https://dmp.opidor.fr/> exigences légales et éthiques

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 ?

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 ?

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

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 ?

B *I* ☰ ▾ ☰ ▾ 🔗 📄 ▾

Utilisation de données publiées (voir la publication pour plus de précision) :

DOI

- Aussi : durée de conservation

ANR

Lorsque vous manipulez des données à caractère personnel, veillez à ce que les lois sur la protection des données (par exemple, RGPD) soient appliquées, notamment :

- Obtenir un consentement éclairé pour la préservation et/ou le partage de données personnelles.
- Envisager l'anonymisation des données personnelles pour la préservation et/ou le partage (des données correctement anonymisées ne sont plus considérées comme des données personnelles).
- Envisager la pseudonymisation des données personnelles (la principale différence avec l'anonymisation est que la pseudonymisation est réversible).
- Envisager le chiffrement des données, qui est considéré comme un cas particulier de pseudonymisation (la clé de cryptage doit alors être stockée séparément des données, par exemple par un tiers de confiance).
- Expliquer si une procédure d'accès spécifique a été mise en place pour les utilisateurs autorisés à accéder aux données personnelles.

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 ?

ANR



- Expliquer qui sera le **propriétaire** des données, c'est-à-dire qui aura le droit d'en contrôler l'accès :
 - Expliquer quelles **conditions d'accès** s'appliqueront aux données. Les données seront-elles librement accessibles, ou des restrictions seront-elles appliquées ? Si oui, lesquelles ? Envisager l'utilisation de **licences** concernant l'accès et la réutilisation des données.
 - S'assurer de couvrir, dans l'**accord de consortium**, ces questions de droits de contrôle d'accès aux données pour les projets multipartenaires et en cas de propriété partagée des données.

- Indiquer si les **droits de propriété intellectuelle** (par exemple la directive bases de données, droits *sui generis*) sont affectés. Dans l'affirmative, expliquer lesquels et comment cela sera traité.
- Indiquer s'il y a des **restrictions sur la réutilisation des données** fournies par des tiers.

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 ?

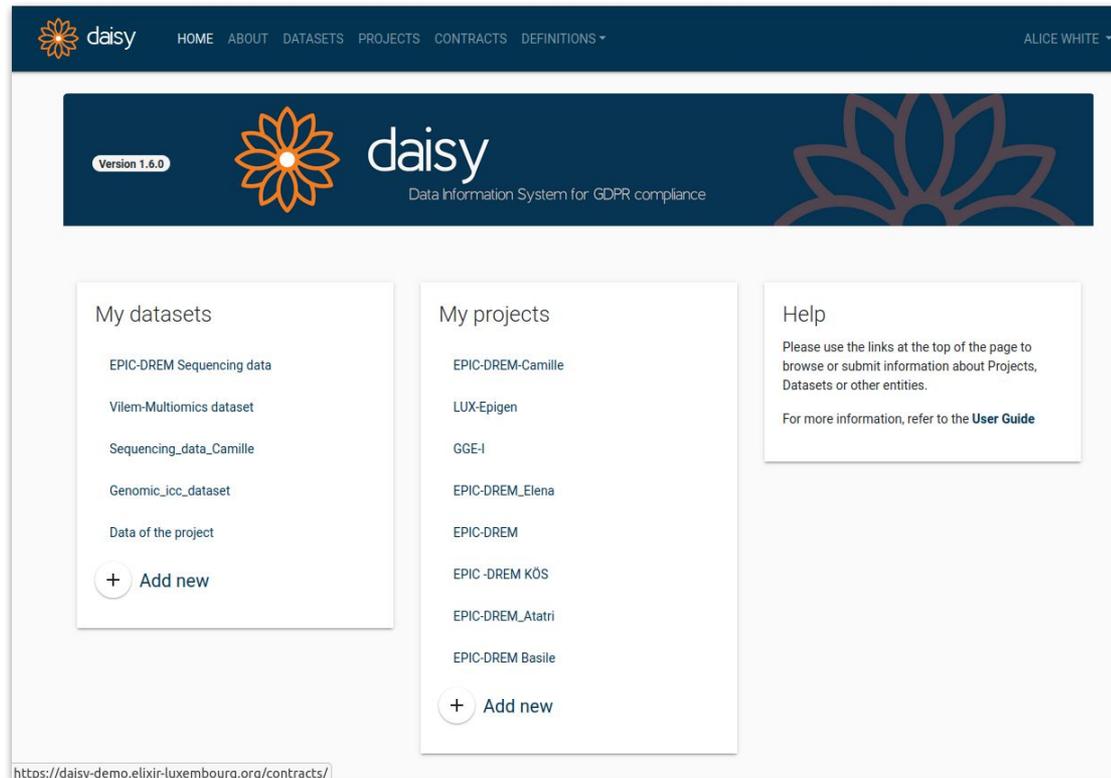
- Se renseigner auprès du DPO

ANR

- Déterminer si les questions d'éthique auront une incidence sur la façon dont les données seront stockées et transférées, qui pourra les voir ou les utiliser et quelles durées de conservation leur seront appliquées. Démontrer que ces aspects sont bien pris en compte et planifiés.
- Adopter les codes de conduite nationaux et internationaux et le code d'éthique institutionnel et vérifier si une revue des pratiques (par exemple par un comité d'éthique) est requise pour ce qui concerne la collecte de données dans le cadre du projet de recherche.

Exigences légales et éthiques -- exemples : DAISY

- Developed by ELIXIR-LU
 - **Free and open source**
- <https://github.com/elixir-luxembourg/daisy/>
- DAISY (Data Information SYstem) is a tool that assists GDPR compliance by keeping a register of personal data used in research
 - See the [User Guide](#)
 - Visit the [demo](#) site



The screenshot displays the DAISY web interface. At the top, there is a navigation bar with the DAISY logo and links for HOME, ABOUT, DATASETS, PROJECTS, CONTRACTS, and DEFINITIONS. The user is logged in as ALICE WHITE. Below the navigation bar is a header section with the DAISY logo, the text "Data Information System for GDPR compliance", and a version indicator "Version 1.6.0". The main content area is divided into three columns: "My datasets", "My projects", and "Help".

My datasets

- EPIC-DREM Sequencing data
- Vilem-Multiomics dataset
- Sequencing_data_Camille
- Genomic_jcc_dataset
- Data of the project
- + Add new

My projects

- EPIC-DREM-Camille
- LUX-Epigen
- GGE-I
- EPIC-DREM_Elena
- EPIC-DREM
- EPIC-DREM KOS
- EPIC-DREM_Atatri
- EPIC-DREM Basile
- + Add new

Help

Please use the links at the top of the page to browse or submit information about Projects, Datasets or other entities.

For more information, refer to the [User Guide](#)

<https://daisy-demo.elixir-luxembourg.org/contracts/>

A tool to assist with GDPR: DAISY

- Sensitivity class

Dataset

Sequencing_data_Camille

Local custodians

Alice White John Black Student Rolin

Project

EPIC-DREM-Camille

Other external id(s)

-

Other comments

-

Added on

Oct. 5, 2021, 10:57 a.m.

Last edit

Oct. 5, 2021, 11:07 a.m.

Sensitivity class

Re-Identifiable, Sensitive

Accession number

-

NOT PUBLISHED

👁️

✎

🗑️

A tool to assist with GDPR: DAISY

- Limits to storage

Data declarations +

Title	Cohort	Obtained from	Datatypes	Special subjects	Storage end	Actions
ChipSeq_data_Camille		Luxembourg Centre for Systems Biomedicine	Samples, Genetic_and_derived_genetic_data	no	Oct. 5, 2024	  
WGS_data_Camille		Luxembourg Centre for Systems Biomedicine	Samples, Whole_genome_sequencing	no	Oct. 5, 2024	  

A tool to assist with GDPR: DAISY

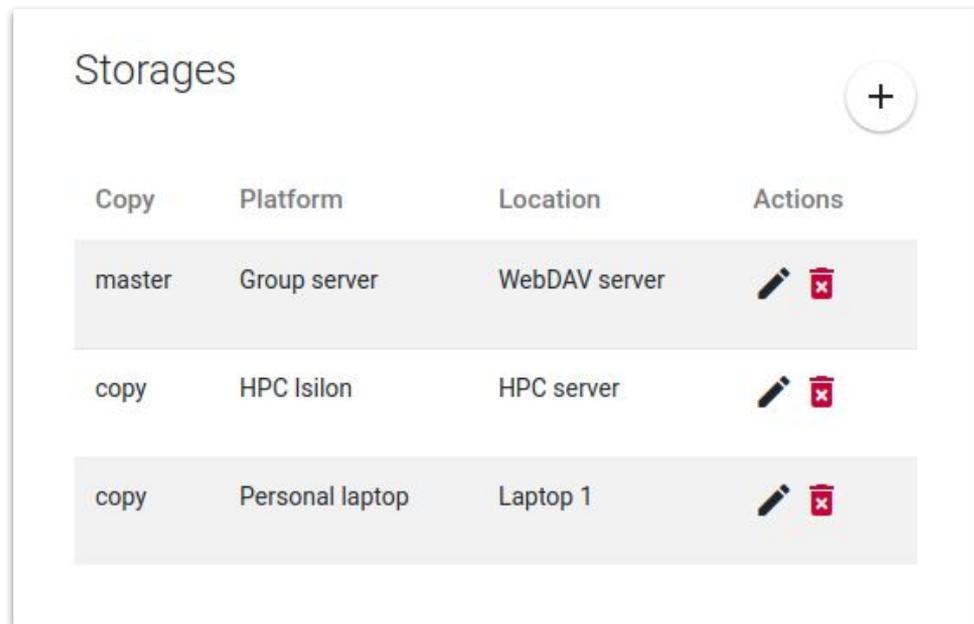
- Consent

Legal bases +

Scope	Legal basis categories	Personal data categories	Remarks	Actions
ChipSeq_data_Camille WGS_data_Camille	Genetic data	Consent [6.1(a)]		 

A tool to assist with GDPR: DAISY

- Storage type and location



The screenshot shows a web interface titled "Storages" with a plus sign in a circle in the top right corner. Below the title is a table with four columns: "Copy", "Platform", "Location", and "Actions". There are three rows of data, each with a light gray background. Each row has a pencil icon and a trash can icon in the "Actions" column.

Copy	Platform	Location	Actions
master	Group server	WebDAV server	 
copy	HPC Isilon	HPC server	 
copy	Personal laptop	Laptop 1	 

A tool to assist with GDPR: DAISY

- Access control

Accesses +

Location	Remarks	Granted on	User	Actions
master - Group server - WebDAV server copy - HPC Isilon - HPC server	Data with other project(s) under the same re...	Oct. 5, 2021	John Black	 
master - Group server - WebDAV server copy - HPC Isilon - HPC server	Data with other project(s) is under the same...	Oct. 5, 2021	Student Rolin	 
master - Group server - WebDAV server	Data with other project(s) is under the same...	Oct. 5, 2021	-	 

Fair Data formation 15-18 novembre 2021

Renseignements sur le projet

Produits de recherche

Modèle choisi

Rédiger

Partager

Télécharger

tout développer | tout réduire

5. Partage des données et conservation à long terme (4 questions)



DMP OPIDoR <https://dmp.opidor.fr/> partage 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 ?

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) ?

5c. Quelles méthodes ou quels outils logiciels seront 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 ?

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 ?

ANR

B *I* ☰ ☰ 🔗 📄

Les données ont été déjà partagées lors de la création de ce PGD training.

Localisation des données :

Note

Habituellement, les données sont rendues publiques dans une banque internationale comme ENA au moment de la publication. Un identifiant de projet unique permet ensuite à la communauté d'accéder aux données.

- Expliquer comment les données pourront être **retrouvées et partagées** (par exemple, par le dépôt dans un entrepôt de données de confiance, l'indexation dans un catalogue, par l'utilisation d'un service de données sécurisé, par le traitement direct des demandes de données, ou l'utilisation de tout autre mécanisme).
- Définir le **plan de préservation des données** et fournir l'information sur la durée d'archivage pérenne des données.

- Expliquer à quel moment les données seront **rendues disponibles**. Indiquer les délais de publication prévus. Expliquer si une utilisation exclusive des données est revendiquée et, dans l'affirmative, pour quelle raison et pour combien de temps. Indiquer si le **partage des données** sera différé ou limité, par exemple pour des raisons de publication, pour protéger la propriété intellectuelle ou le dépôt de brevets.
- Indiquer qui pourra utiliser les données. S'il s'avère nécessaire de restreindre l'accès pour certaines communautés ou d'imposer un accord pour le partage de données, expliquer comment et pourquoi. Expliquer les mesures qui seront prises pour dépasser ou minimiser ces restrictions.

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) ?

ANR

- Indiquer quelles données ne doivent pas être divulguées ou doivent être détruites pour des raisons contractuelles, légales, ou réglementaires.
- Indiquer comment il sera décidé quelles données garder. Décrire les données qui seront à préserver à long terme.
- Décrire les utilisations (et/ou les utilisateurs) prévisibles des données dans un cadre de recherche.
- Indiquer où les données seront déposées. Si aucun entrepôt reconnu n'est proposé, démontrer dans le plan de gestion des données que les données pourront être prises en charge efficacement au-delà de la durée de financement du projet. Il est recommandé de démontrer que les politiques des entrepôts et les procédures de dépôts (y compris les standards de métadonnées, et coûts mis en œuvre) ont été vérifiés.

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

B *I* ☰ ▾ ☰ ▾ 🔗 📄 ▾

Un terminal ou un navigateur web

ANR

- Indiquer si les utilisateurs potentiels auront besoin d'outils spécifiques pour l'accès et la (ré)utilisation des données. Tenir compte de la durée de vie des logiciels nécessaires pour accéder aux données.
- Indiquer si les données seront partagées via un entrepôt, si les demandes d'accès seront traitées en direct, ou si un autre mécanisme sera utilisé.

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

B *I* ☰ ☰ ☰ ☰ ☰ ☰

La banque internationale (ici l'ENA) se charge de la création d'un identifiant unique et pérenne.

ANR

- Expliquer comment les données pourraient être réutilisées dans d'autres contextes. Les identifiants pérennes devraient être appliqués de manière à ce que les données puissent être localisées et référencées de façon fiable et efficace. Les identifiants pérennes aident aussi à comptabiliser les citations et les réutilisations.
- Indiquer s'il sera envisagé d'attribuer aux données un identifiant pérenne. Typiquement, un entrepôt pérenne de confiance attribuera des identifiants pérennes.

Fair Data formation 15-18 novembre 2021

Renseignements sur le projet

Produits de recherche

Modèle choisi

Rédiger

Partager

Télécharger

tout développer | tout réduire

6. Responsabilités et ressources en matière de gestion des données (2 questions)



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) ?

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) ?

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) ?

B I [Menu] [Menu] [Link] [Grid]

Paulette Lieby

- Rôle : gestionnaire de données
- Institution : IFB
- Email : |

ANR

- Décrire les rôles et les responsabilités concernant les activités de gestion des données, par exemple : saisie des données, production des métadonnées, qualité des données, stockage et sauvegarde, archivage et partage des données. Nommer la(es) personne(s) responsable(s) impliquée(s) dans la mesure du possible.
- Pour les projets menés en collaboration, expliquer comment s'effectue la coordination des responsabilités de gestion des données entre partenaires.
- Indiquer qui est responsable de la mise en œuvre du PGD, et qui s'assure qu'il est examiné et, si nécessaire, révisé.
- Envisager des mises à jour régulières du PGD.

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) ?

- Aussi, s'assurer que les ressources couvrent le temps long (par exemple, la conservation à long terme)

ANR

- Expliquer comment les ressources nécessaires (par exemple le temps) à la préparation des données pour le partage/préservation (curation des données) ont été chiffrées. Examiner et justifier soigneusement toutes les ressources nécessaires pour diffuser les données.
- Il peut s'agir de frais de stockage, de coût matériel, de temps de personnel, de coûts de préparation des données pour le dépôt, de frais d'entrepôt et d'archivage.
- Indiquer si des ressources supplémentaires sont nécessaires pour préparer les données en vue de leur dépôt ou pour payer tous les frais demandés par les entrepôts de données. Si oui, précisez le montant et comment ces coûts seront couverts.

Merci !
Questions ?

Remerciements : Hélène Chiapello, Thomas Denecker, Frédéric de Lamotte.