DMP du projet "Comportement couplé des zones de failles: du laboratoire au terrain"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "PGDR TEST type "projet de recherche" IRSN" fourni par IRSN - Institut de Radioprotection et de Sureté Nucléaire.

Renseignements sur le plan

Titre du plan	DMP du projet "Comportement couplé des zones de failles: du laboratoire au terrain"
Livrable	D2
Version	Version intermédiaire
Objet/périmètre du plan	Le Plan de Gestion des Données décrit le cycle de vie des jeux de données collectés, traités ou générés dans le cadre du projet. Ce document précise comment les données du projet sont gérées, depuis leur création ou collecte jusqu'à leur éventuel partage et leur archivage.
Domaines de recherche (selon classification de l'OCDE)	Earth and related environmental sciences
Langue	fra
Date de création	2022-07-22
Date de dernière modification	2024-01-09

Renseignements sur le projet

Titre du projet	Comportement couplé des zones de failles: du laboratoire au terrain
Acronyme	CHENILLE
Résumé	La compréhension du comportement couplé thermo-hydro-mécanique des zones de faille dans des réservoirs naturellement fracturés est d'une importance fondamentale pour diverses raisons sociétales et économiques, telles que la transition énergétique durable pour une utilisation sûre des ressources naturelles (stockage de l'énergie, stockage des déchets nucléaires ou énergie géothermique). L'objectif global de ce projet est de mieux comprendre les processus physiques résultant d'une charge thermique et hydrique sur une zone de faille existante, identifiée et préalablement caractérisée. Il s'agit ici de réaliser une expérience in situ d'injection de fluide thermiquement contrôlée dans une des failles accessibles depuis le laboratoire souterrain de recherche (URL) de Tournemire. Un système de chauffage sera installé autour de la zone d'injection

et permettra une augmentation incrémentale, précise et contrôlée de la charge thermique. En outre, un important système de surveillance sera installé pour mesurer les événements sismiques et asismiques induits par la charge thermique ou hydraulique. Le système de surveillance sera composé des capteurs acoustiques et des capteurs large bande qui mesureront les événements de faible à moyenne magnitude. Par ailleurs, il est également prévu d'installer un réseau de fibres optiques pour mesurer la température et la déformation de manière distribuée dans la zone d'investigation. Les méthodes sismiques actives seront déployées avant et après l'expérience pour déterminer l'architecture de la faille mais aussi pour détecter l'apparition de nouvelles structures induites par la pressurisation hydro-thermique de la faille. Une série d'expériences en laboratoire sera menée pour comprendre l'évolution chimique et pétrophysique qui se produit dans les zones de faille pendant le chargement thermique et hydraulique. Des expériences thermiques en chambre climatique sur des échantillons de la zone de faille seront réalisées afin de comparer l'évolution de la composition minéralogique des échantillons avec ceux prélevés dans la zone étudiée sur le terrain. Enfin, une étude de la mécanique des roches, de l'échelle microscopique à l'échelle centimétrique avec un suivi des propriétés acoustiques sera réalisée. Cette étude comprendra des expériences de Microscope Electronique à Balayage par spectroscopie à rayons X à dispersion d'énergie (MEB-EDS) permettant d'identifier les mécanismes de localisation des déformations à micro-échelle auxquels il est prévu d'ajouter un système de mesure acoustique. Afin d'étudier l'évolution du comportement mécanique en fonction de l'échelle, des expériences en presse triaxiale, toujours avec un suivi acoustique, sont prévues.

Sources de financement

- Agence Nationale de la Recherche : ANR-21-CE49-0016

Date de début 2022-01-01

Date de fin 2025-12-31

Partenaires

- GFZ Potsdam / Section 4.2 Geomechanics and Scientific Drilling
- Freie Universität Berlin
- Laboratoire de mécanique des solides
- Laboratoire de géologie de lEcole Normale Supérieure

Produits de recherche :

1. DataSet 1 - Données sismiques (Jeu de données)
2. Dataset 2 - Données thermiques par fibres optiques (Jeu de données)
3. Dataset 3 - Images de microscopie optique ou MEB de nucleation et propagation de fissures (Image)
4. Dataset 4 - Données microsismiques sur échantillons dans une presse triaxiale (Jeu de données)
5. Dataset 5 - Base de données sur la propagation des ondes P à différentes échelles dans un milieu argileux (Jeu de données)
6. Dataset 5 - Données de pression, température et fluw du forage d'injection BH1 (Jeu

de données)

Contributeurs

Nom	Affiliation	Rôles
DICK Pierre	Pôle Santé Environnement - Direction Environnement - 201722596S	• Coordinateur du projet • Personne contact pour les données (Dataset1_Sismiques, Dataset5_Velocity, Dataset2_Thermiques, Dataset4_Microsis, Dataset3_Images, Dataset6_Forage) • 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 "Comportement couplé des zones de failles: du laboratoire au terrain"

Informations sur le service de recherche

Nom du service (ou laboratoire, le cas échéant) de recherche

SErvice des Pollutions et des Déchets Radioactifs

Acronyme service (ou laboratoire) et appartenance

SPDR

Structure IRSN de recherche *(selon répertoire national des structures)*

- UREnv 201722596S
- UREnv 201722596S

Partenaires impliqués dans le projet de recherche

Ecole Polytechnique - LMS - Alexandre Dimanov (*alexandre.dimanov@polytechnique*.edu)
ENS Paris - Laboratoire de Géologie - Jérôme Fortin (fortin@geologie.ens.fr)
GFZ Potsdam - Geomechanik und Wissenschaftliches Bohren - Audrey Bonnelye (bonnelye@gfz-potsdam.de)

Responsable(s) de la gestion et de l'ouverture des données

Pierre Dick (IRSN) - pierre.dick@irsn.fr - +33 (0) 1 58 35 73 61

Informations générales sur les données du projet

DataSet 1 - Données sismiques

Description des données

Description des données	Données d'acquisition de campagnes de mesures sismiques
Type(s) *(numériques, textuelles, images, vidéo...)*	Numériques
Format(s)	DAT, TXT
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	96 To

Préciser le mode d'obtention des données produites

- 2- Données expérimentales
- 2- Données expérimentales

Données utilisées par le projet et provenant de sources extérieures

Question sans réponse.

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

Question sans réponse.

"Criticité" des données : que ce passerait-il si les données étaient corrompues ou perdues?

- 4- On ne peut pas les ré-obtenir
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 4- On ne peut pas les ré-obtenir
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé

Dataset 2 - Données thermiques par fibres optiques

Description des données

Description des données	Données d'acquisition de campagnes de mesures thermiques
Type(s) *(numériques, textuelles, images, vidéo...)*	Numériques
Format(s)	DAT, TXT, CSV
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	8 Go

Préciser le mode d'obtention des données produites

- 2- Données expérimentales
- 2- Données expérimentales

Données utilisées par le projet et provenant de sources extérieures

Question sans réponse.

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

THERMA :Distributed temperature analysis software

"Criticité" des données : que ce passerait-il si les données étaient corrompues ou perdues?

- 4- On ne peut pas les ré-obtenir
- 4- On ne peut pas les ré-obtenir

Dataset 3 - Images de microscopie optique ou MEB de nucleation et propagation de fissures

Description des données

Description des données	Images de microspoie optique et MEB
Type(s) *(numériques, textuelles, images, vidéo...)*	images
Format(s)	PNG, TIFF, JPEG
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	1 Go

Préciser le mode d'obtention des données produites

- 1- Données d'observation
- 2- Données expérimentales
- 3- Données computationnelles
- 1- Données d'observation
- 2- Données expérimentales
- 3- Données computationnelles

Données utilisées par le projet et provenant de sources extérieures

Question sans réponse.

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

Question sans réponse.

"Criticité" des données : que ce passerait-il si les données étaient corrompues ou perdues?

- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé

Dataset 4 - Données microsismiques sur échantillons dans une presse triaxiale

Description des données

Description des données	Données microsismiques obtenues sur échantillons
Type(s) *(numériques, textuelles, images, vidéo...)*	Numériques
Format(s)	SEIS
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	10 Go

Préciser le mode d'obtention des données produites

- 2- Données expérimentales
- 2- Données expérimentales
- 2- Données expérimentales
- 2- Données expérimentales

Données utilisées par le projet et provenant de sources extérieures

Question sans réponse.

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

Question sans réponse.

"Criticité" des données : que ce passerait-il si les données étaient corrompues ou perdues?

- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé

Dataset 5 - Base de données sur la propagation des ondes P à différentes échelles dans un milieu argileux

Description des données

Description des données	Données des ondes P obtenues sur des objets de différentes tailles
Type(s) *(numériques, textuelles, images, vidéo...)*	Numériques
Format(s)	SEIS
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	10 Go

Préciser le mode d'obtention des données produites

- 2- Données expérimentales
- 2- Données expérimentales
- 2- Données expérimentales
- 2- Données expérimentales

Données utilisées par le projet et provenant de sources extérieures

Question sans réponse.

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

Question sans réponse.

"Criticité" des données : que ce passerait-il si les données étaient corrompues ou perdues?

- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 2- Elles sont re-générables avec un effort/coût modéré
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé

Dataset 5 - Données de pression, température et fluw du forage d'injection BH1

Description des données

Description des données	Données de pression (packers et chambres d'injection), température (chambres d'injection) et flux d'un fluide provenant d'un système d'injection (multipacker) de forage
Type(s) *(numériques, textuelles, images, vidéo...)*	Numériques
Format(s)	CSV, TXT
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	1 Go

Préciser le mode d'obtention des données produites

- 2- Données expérimentales
- 2- Données expérimentales

Données utilisées par le projet et provenant de sources extérieures

Question sans réponse.

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

Question sans réponse.

"Criticité" des données : que ce passerait-il si les données étaient corrompues ou perdues?

- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 4- On ne peut pas les ré-obtenir
- 3- Elles ne sont re-générables qu'au prix d'un effort/coût élevé
- 4- On ne peut pas les ré-obtenir

Documentation et qualité des données au fil du projet de recherche

Qualification des données au fil du projet de recherche

Pendant le processus de recherche, les données vont suivre un processus de "qualification" les faisant passer de l'état de données brutes ou primaires (B) aux états de données nettoyées /traitées (N) puis analysées (A) et enfin interprétées/figées (I). Ce processus est constitutif du "cycle de vie de la donnée".

Les moyens mis en œuvre pour atteindre ces différents états seront précisés dans la description du protocole expérimental (cahier de laboratoire ou autre dispositif).

Selon le projet de recherche, les données "produits" ou données d'intérêt pourront être des données B, N, A et/ou I.

Règles de nommage des ensembles de données

Les règles de nommage de l'ensemble des données suiveront la convention de nommage entrepris par la NASA (NPARC Verification and Validation Web Site File Naming Conventions Page (nasa.gov))

Arborescence de classement

L'arborescence de classement suivra la convention adopté par GenR (https://genr.eu/wp/towards-a-standardized-research-folder-structure/)

Gestion des données pendant le projet de recherche

Stockage des données tout au long du processus de recherche

Le données générées dans les locaux de l'IRSN seront sauvegardées et stockées sur les serveurs de l'IRSN à Fontenay-aux-Roses.

Les données générées à Polytechnique seront sauvegardées et stockées sur les serveurs de l'Ecole Polytechnique.

Les données générées à l'ENS Paris seront sauvegardées et stockées sur les serveurs de l'ENS Paris.

Les données générées dans les locaux de GFZ seront sauvegardées et stockées sur les serveurs du GFZ.

Pendant le projet, les données seront partagées entre participants :
- au niveau d'un espace "Sharepoint" commun géré par Pierre Dick avec le soutien du service "Développement Numérique et aux Systèmes d'information" (D3NSI) de l'IRSN.

La sécurité des données dans ces espaces est garantie par le contrôle des accès par la personne administratrice de l'espace dédié.

Plan de sauvegarde des données

Les données du projet seront stockées à l'IRSN en utilisant les espaces préconisés par D3NSI, la sauvegarde (et la restauration le cas échéant) est garantie par les procédures d'exploitation fixées dans un référentiel (pour info: le référentiel IRSN/S3A/REF-057).

Conservation des données au terme du projet de recherche

Définition des critères de tri et de sélection des données à conserver

Réflexion à mener sur les critères de tri des données + la documentation de ces critères.
Puis, environ 2 fois par an, nous essaierons de réaliser un tri initial des données produites dans 6 derniers mois pour éviter d'accumuler trop de données

---

Durée de conservation des jeux de données

La durée de conservation des données sera définie en accord et avec l'appui des archivistes du SEARCH (IRSN), elle sera cohérente avec le tableau de gestion des archives (TGA).

Espace de conservation des données

M:\DATA_CHENILLE

Y-a-t-il besoin de ressources supplémentaires particulières pour la conservation des données?

A définir

Référencement des données dans le catalogue des données de l'institut

Les données seront répertoriées dans le catalogue des données de l'IRSN.
Ceci sera fait avec la personne référente du service en charge de la mise à jour du catalogue IRSN sur PADI et avec la cellule de valorisation des données à D3NSI.

Exigences légales et éthiques, code de conduite

Nature des données traitées dans le projet: s'agit-il de ?

Question sans réponse.

Si des données personnelles sont contenues dans les données du projet, peuvent-elles être:

Question sans réponse.

Y -a-t-il besoin de traiter des considérations éthiques ou juridiques par rapport aux données?

A définir avec le DRP/CCE.

Y-a-t-il des questions relatives à la propriété intellectuelle (propriété, droits d'usage,...)?

La règlementation française s'appliquera sur la propriété intellectuelle des données.
Il n'est pas possible de réutiliser les données hors du périmètre du consortium s'il s'agit de résultats partagés, sauf accord écrit préalable des partenaires. La demande devra être signée par tous les partenaires concernés et transmise au coordinateur de projet pour archives.

Partage des données en Open data

Les données du projet de recherche peuvent-elles être publiées en Open Data?

La loi pour une république Numérique de 2016 porte l'obligation d'ouverture des données publiques y compris les données de la recherche. Depuis le décret n°2021-1572 du 3 décembre 2021 *relatif aux exigences de l'intégrité scientifique par les établissements publics contribuant au service public de la recherche,* il a obligation pour les établissements publics de promouvoir la diffusion en Open Data des données produites par de la recherche financée sur fonds publics en les ayant rendues **FAIR : Faciles à trouver, Accessibles, interopérables et Réutilisables.** Le principe d'ouverture qui prévaut est celui des ***"données aussi ouvertes que possible et fermées que nécessaire"***.
Concernant la diffusion en Open Data, la consigne de l'IRSN est d'utiliser les entrepôts thématiques ou les entrepôts définis conjointement avec les partenaires du projet.

Constitution des jeux de données mis en Open data

A définir dans les prochains 6 mois avec les participants du projet

---

Données FAIR

Choix des métadonnées pour des données Faciles à trouver

Non défini à ce stade (fonction des résultats)

Attribution d'un identifiant unique et permanent de chacun des jeux de données

Non défini à ce stade (fonction des résultats)

Qui a créé le(s) jeux de données ?

Non défini à ce stade (fonction des résultats)

Interopérabilité technique des données

Nous utiliserons, autant que possible, des formats ouverts (=libre)

Type de données	Formats
Données brutes	CSV, JSON, XML
Données géographiques	KML, SHP
Texte	DOCX, ODT, RTF, TXT
Image	BMP, GIF, JPG, PNG
Tableur	XLSX, CSV, ODS
Vidéo	MKV, NUT, OGM
Audio	FLAC, OGG, WAV, MP3
Présentations	ODP, PPTX
Archivage	7z (Seven Zip), TAR, ZIP
Dessin	EPS, ODG, SVG

Quelles licences pour les données réutilisables?

Les éditeurs de journaux scientifiques attribuent un DOI aux publications.
Nous prévoyons également de publier au cours du projet un ou plusieurs "data papers" (avec DOI) pour partager certaines données et métadonnées (exemple : Geoscience Data Journal, Earth System Science Data).
Les publications scientifiques archivées dans HAL bénéficient bénéficient également d'un identifiant permanent.

Coût des dépôts de données et coût du maintien des données dans les entrepôts de données?
Ces coûts ont-ils été évalués?
Sont-ils intégrés dans le budget du projet ou non?
Par qui seront-ils portés, le cas échéant?

Coût des dépôts de données et coût du maintien des données dans les entrepôts de données?
--> Non défini à ce stade (fonction des résultats)
Ces coûts ont-ils été évalués?
--> Non
Sont-ils intégrés dans le budget du projet ou non?
--> Non
Par qui seront-ils portés, le cas échéant?
--> A définir ultérieurement