



ENTREPÔT DE DONNÉES

« Un service de dépôt et de diffusion pour les données de la recherche du Système Terre et de l'Environnement »



Fabrique Sciences Ouverte
Lille - 12.05.2022
C. Pierkot



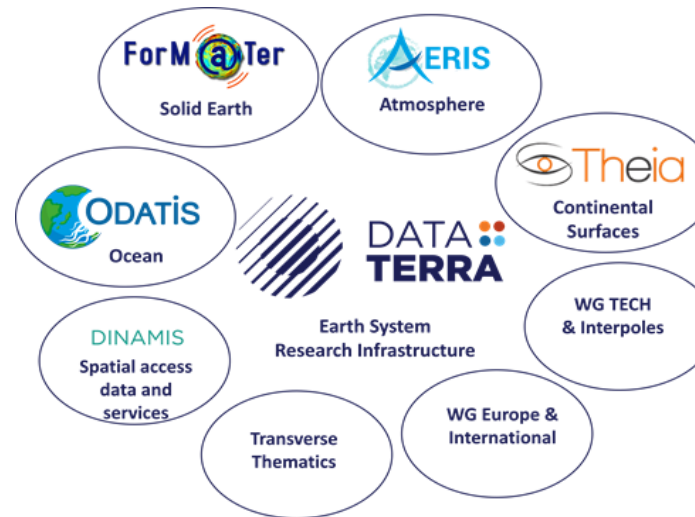
POLES DE DONNÉES ET SERVICES POUR LE SYSTÈME TERRE

Créé en 2016, feuille de route IR MESRI 2018, 2021



Quatre pôles de données correspondant à chacun des grands compartiments du Système Terre :

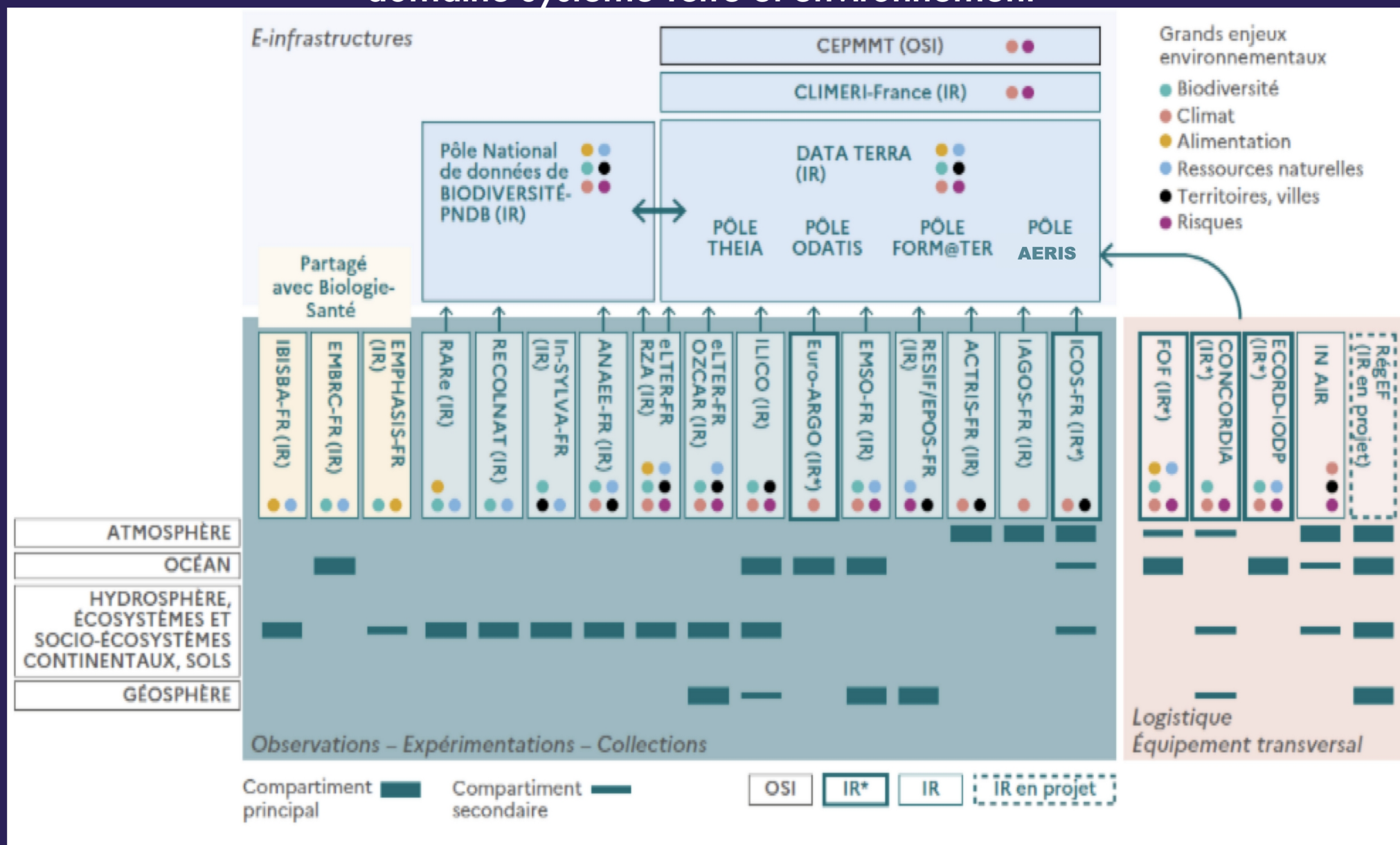
- Surfaces continentales
- Atmosphère
- Océans
- Terre solide



- 26 organismes et universités
- 4 pôles de données
- 6 services (DINAMIS) et groupes de travail transversaux
- 30 Centres de Données et de Services (CDS) et Infrastructures de données spatiales (IDS)



Feuille de route nationale => Positionnement des E-Infrastructures de recherche domaine Système Terre et environnement





Equipex+ GAIA DATA



Vers une plateforme distribuée de données et services FAIR pour l'observation de la Terre

Développer et mettre en œuvre **une infrastructure/plate-forme intégrée de données FAIR et de services distribués pour l'observation et la compréhension du système terre et de l'environnement** sur l'ensemble du cycle de la donnée.

- Faciliter l'accès et l'utilisation des **données** sur l'ensemble des **compartiments du système Terre**
- Développer des services de **visualisation et de traitements**
- Favoriser la **mutualisation, interopérabilité**, émergence d'**approches multi- et inter-disciplinaires**
- Servir les communautés **scientifiques**, les acteurs de l'**action publique et de l'innovation**
- Mettre en œuvre une **stratégie nationale**, européenne et internationale

Entrepôt



Tâche du WP3 : Services transversaux



SEANOE

Gestion d'un entrepôt de données thématique

la Fabrique de la science ouverte
Lille, le 12 mai 2022

LE PROJET

Équipe projet

- Véronique Bertrand – RESIF
- Hélène Bressan – BRGM
- Christelle Pierkot - Data Terra
- Marine Vernet – Data Terra

Avec le soutien de :

- Sandrine Dalmar – Data Terra
- Joel Sudre – Data Terra

Version test
de l'entrepôt



2021



A l'origine du projet

- Analyse des besoins
- Spécifications
- Définition du modèle de métadonnées

Lancement
de l'entrepôt



2022

- Tests
- Thesaurus
- Formation et communication
- Référents et modérateurs

2023

- Amélioration continue du service
- Moissonnage d'autres entrepôts
- Formation/Support/Accompagnement
- Référencement dans catalogue de la PNF
- Métriques et statistiques



Objectifs



Capitaliser les données de la recherche et leur attribuer un DOI

Pourquoi déposer dans l'entrepôt Data Terra ?

Pérennisation et visibilité des données

- Stockage adapté et diffusion de données de qualité
- Attribution de DOI aux données déposées permettant leur citation
- Référencement des données déposées dans le catalogue de la plateforme nationale fédérée

Simplicité du dépôt

- Authentification via EduGAIN/Renater,
- Prise en compte d'identifiants chercheurs,
- Dépôt en quelques étapes
- Modération et accompagnement par des spécialistes du domaine

Qualité des métadonnées et interopérabilité

- Norme ISO 19115-3,
- Interface adaptée aux spécificités des données géo-référencées,
- Thesaurus communautaires

Services associés : possibilité d'associer des ressources aux données (publication, code source, site web projet, ...)

Qui peut déposer ?

- Les chercheurs affiliés à Data Terra ou à une IR partenaire
- Les chercheurs partenaires d'un projet porté par Data Terra

Quoi déposer ?

Les données de recherche du Système Terre et Environnement :

- Produits à valeur ajoutée issus de données d'observation, en vue de publication
- Données issues de projets nationaux ou internationaux dont Data Terra est coordinateur

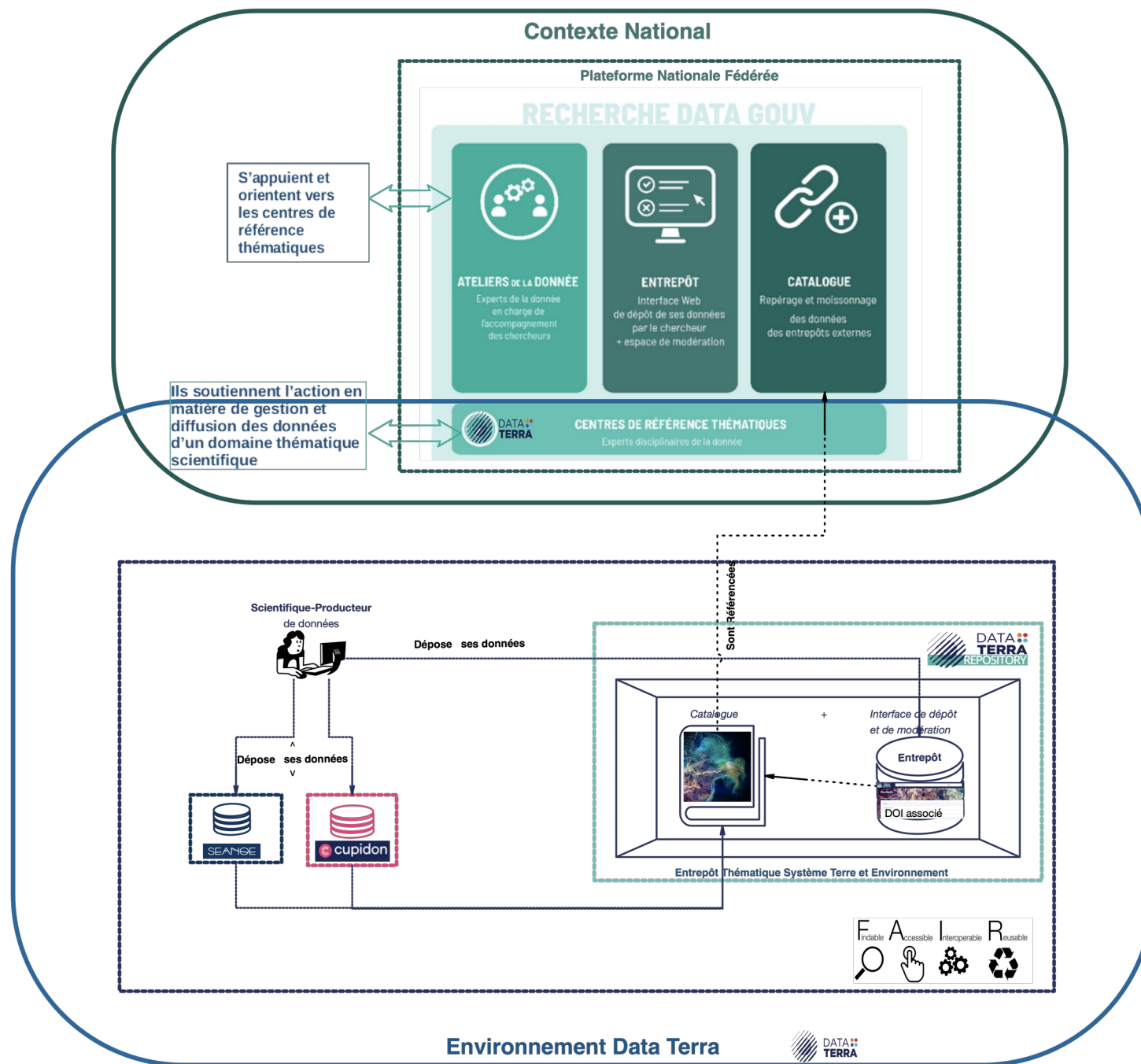
Lien avec les autres entrepôts

Va moissonner d'autres entrepôts thématiques et/ou institutionnels

- Seanoë
- Cupidon
- ...

Entrepôt thématique national de Recherche Data Gouv

L'entrepôt DataTerra dans l'écosystème de la Science Ouverte



Fonctionnalités de l'entrepôt

TECHNO / INFRA

Techno sous jacente	Geonetwork 4. (surcouche implémentée)
Volumétrie	5Go par fichier, 100Go pour le dépôt

PERSONNALISATION

Langue	Fr et En
Auteurs	ORCID, id HAL, ResearchID, Scopus
Niveaux de granularité	2 niveaux : Dépôt et les JDD associés
Mots clés	Thesaurus et vocabulaires avec URI

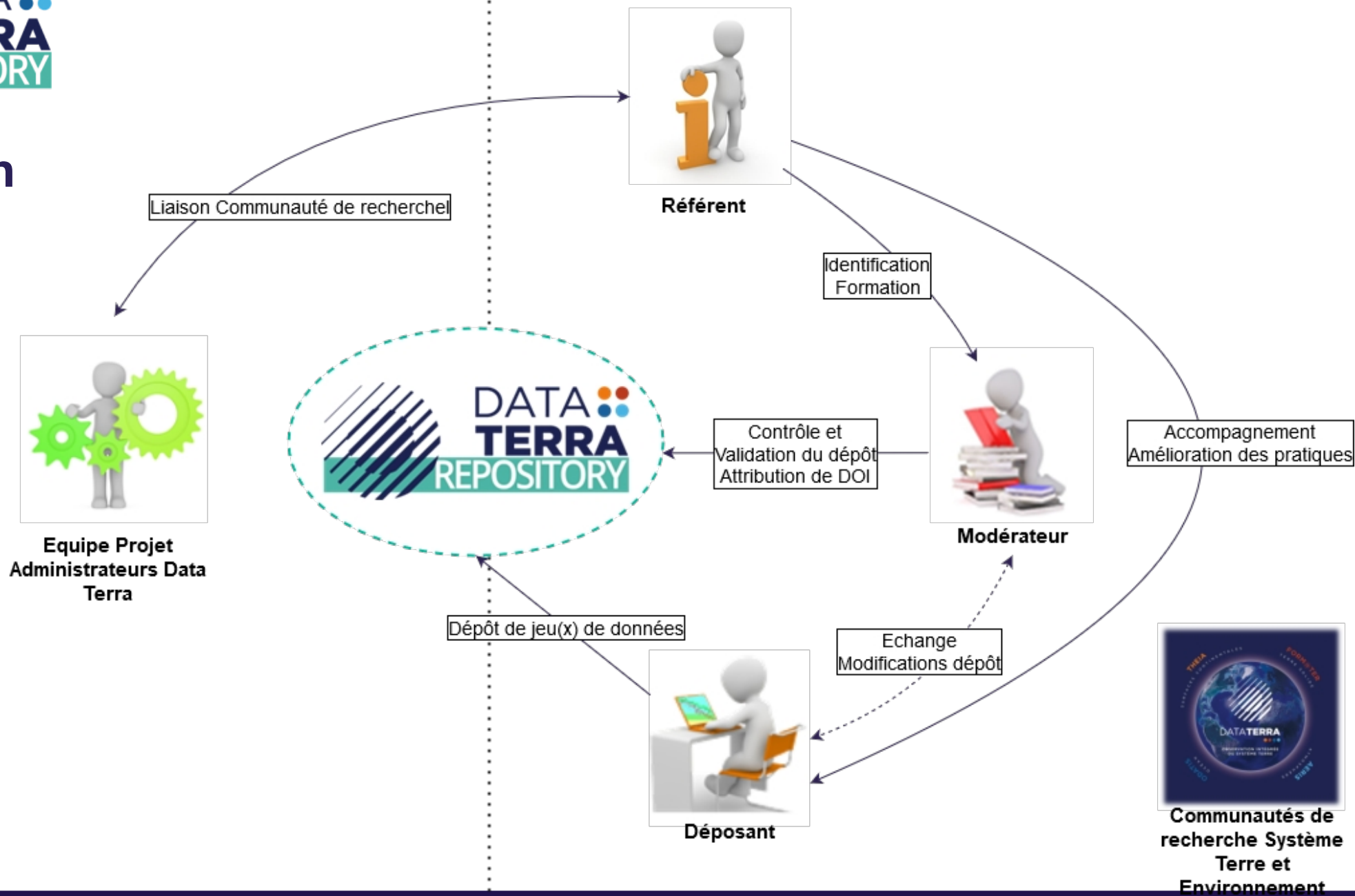
FONCTIONNALITÉS DE DÉPÔT

DOI	Oui, préfixe_DataTerra : 10.34852
Citation générée automatiquement	Oui

CONDITIONS ET LICENCES

Licences d'utilisation	Creative commons + Etalab. CC-BY 4.0 par défaut
Contraintes d'accès	Ouvert avec embargo possible Métadonnées tjs disponibles Données sensibles sous la responsabilité du déposant
Embargo	2 ans max, à partir du moment où le dépôt est publié. Choix et durée sous la Responsabilité du déposant

Organisation



Accompagnement / Communication

Support

- Guide de dépôt
- Guide de modération

Promotion

- Page web : <https://www.data-terra.org/donnees-services/entrepot-de-donnees-data-terra/>
- Vidéo promotionnelle
- Flyer

Ateliers

- Atelier RESIF
- Université de Lille 12/05 – Fabrique de la Science ouverte
Session Entrepôts de données disciplinaires
- Ateliers de la donnée ?



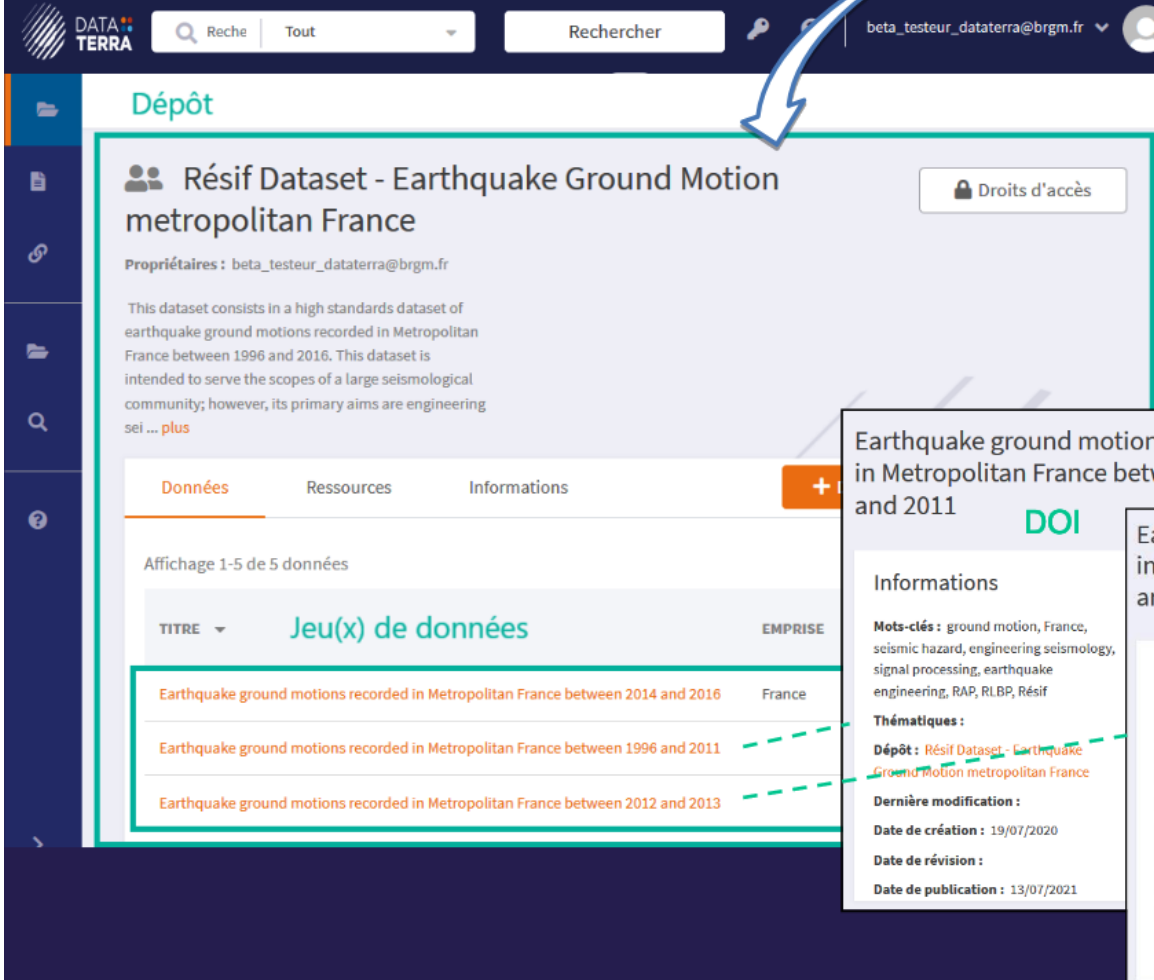


PROCESSUS DE DEPOT

Niveaux de granularité

2.1 Interface utilisateur – mes dépôts





L'entrepôt Data Terra propose **deux niveaux de granularité** :

- Un niveau **Dépôt** qui englobe un ou plusieurs jeux de données
- Un niveau **Jeu de données** sur lequel le DOI est attribué

Ainsi, le déposant peut choisir, en fonction de ses besoins de citation, le niveau de granularité souhaité pour son DOI. Par exemple :

- 1 dépôt avec 1 JDD global associé = 1 DOI
- 1 dépôt avec plusieurs JDD associés partageant des attributs communs (ex: type de données, capteur, localisation, etc.) mais ayant leur propre spécificité (ex: période de collecte) = 1 DOI par JDD

A savoir: Un jeu de données peut être composé de plusieurs fichiers.

Earthquake ground motions recorded in Metropolitan France between 1996 and 2011

DOI

Informations

Mots-clés : ground motion, France, seismic hazard, engineering seismology, signal processing, earthquake engineering, RAP, RLBP, Résif

Thématiques :

Dépôt : Résif Dataset - Earthquake Ground Motion metropolitan France

Dernière modification :

Date de création : 19/07/2020

Date de révision :

Date de publication : 13/07/2021

Earthquake ground motions recorded in Metropolitan France between 2012 and 2013

DOI

Informations

Mots-clés : ground motion, France, seismic hazard, engineering seismology, signal processing, earthquake engineering, RAP, RLBP, Résif

Thématiques :

Dépôt : Résif Dataset - Earthquake Ground Motion metropolitan France

Dernière modification :

Date de création : 18/07/2020

Date de révision :

Date de publication : 13/07/2021

Earthquake ground motions recorded in Metropolitan France between 2012 and 2013

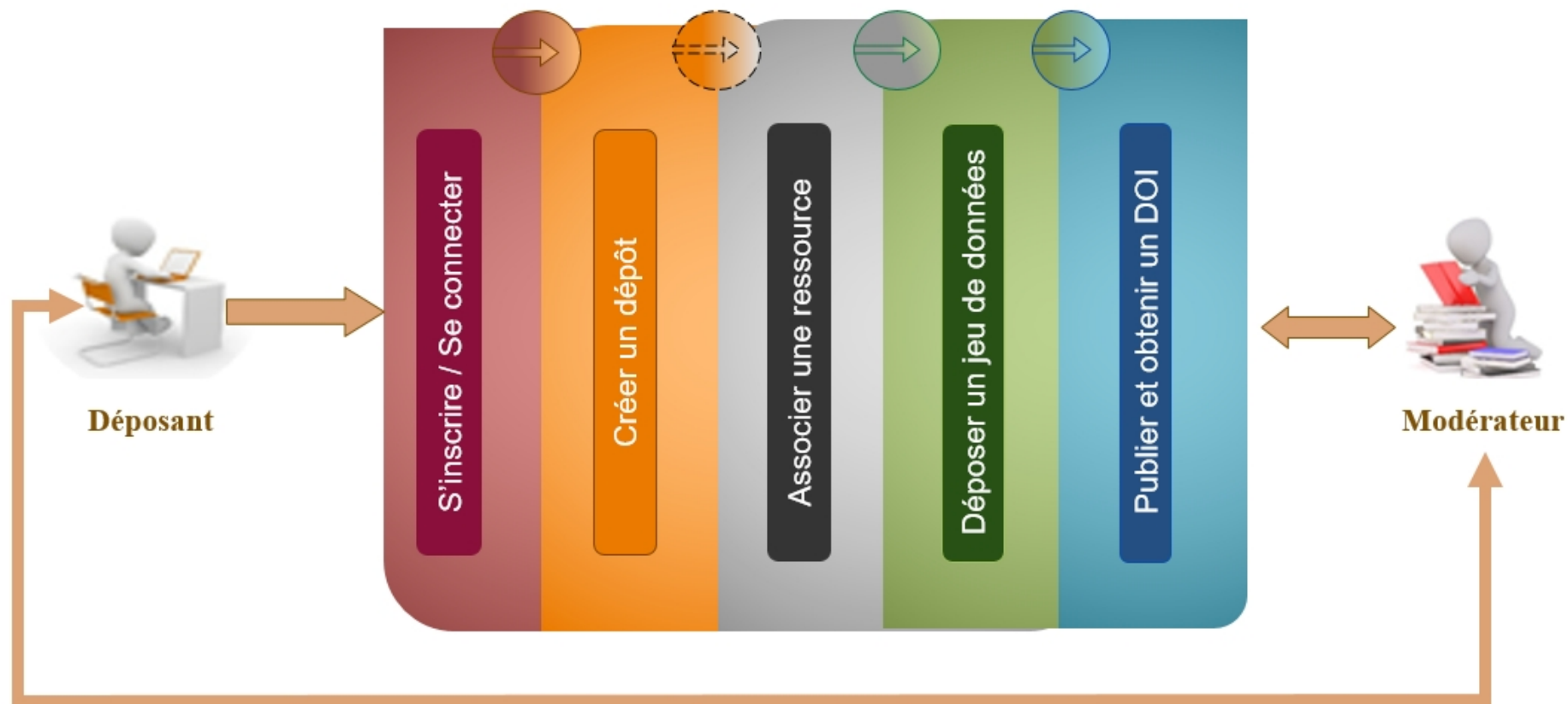
DOI

Emprises

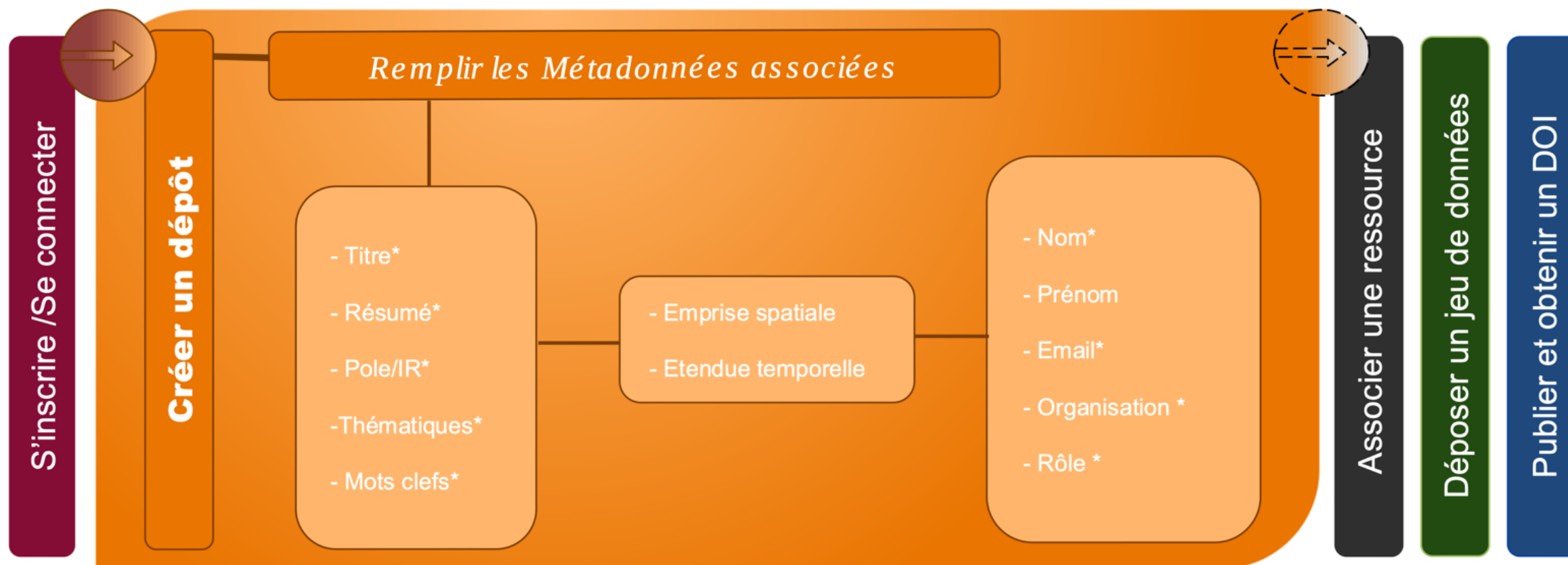
(aucune description)



Workflow - Processus de publication

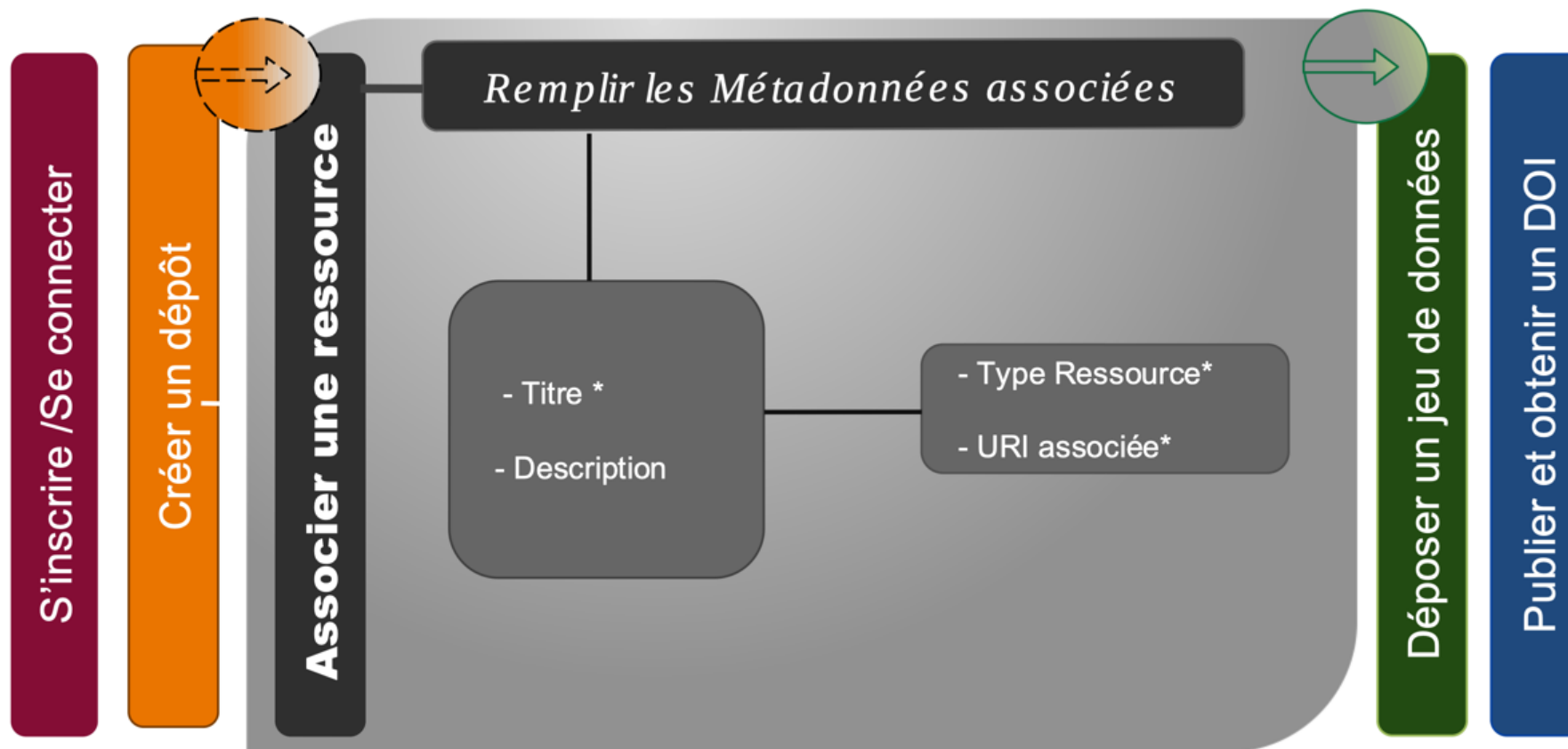


Workflow – Créer un dépôt

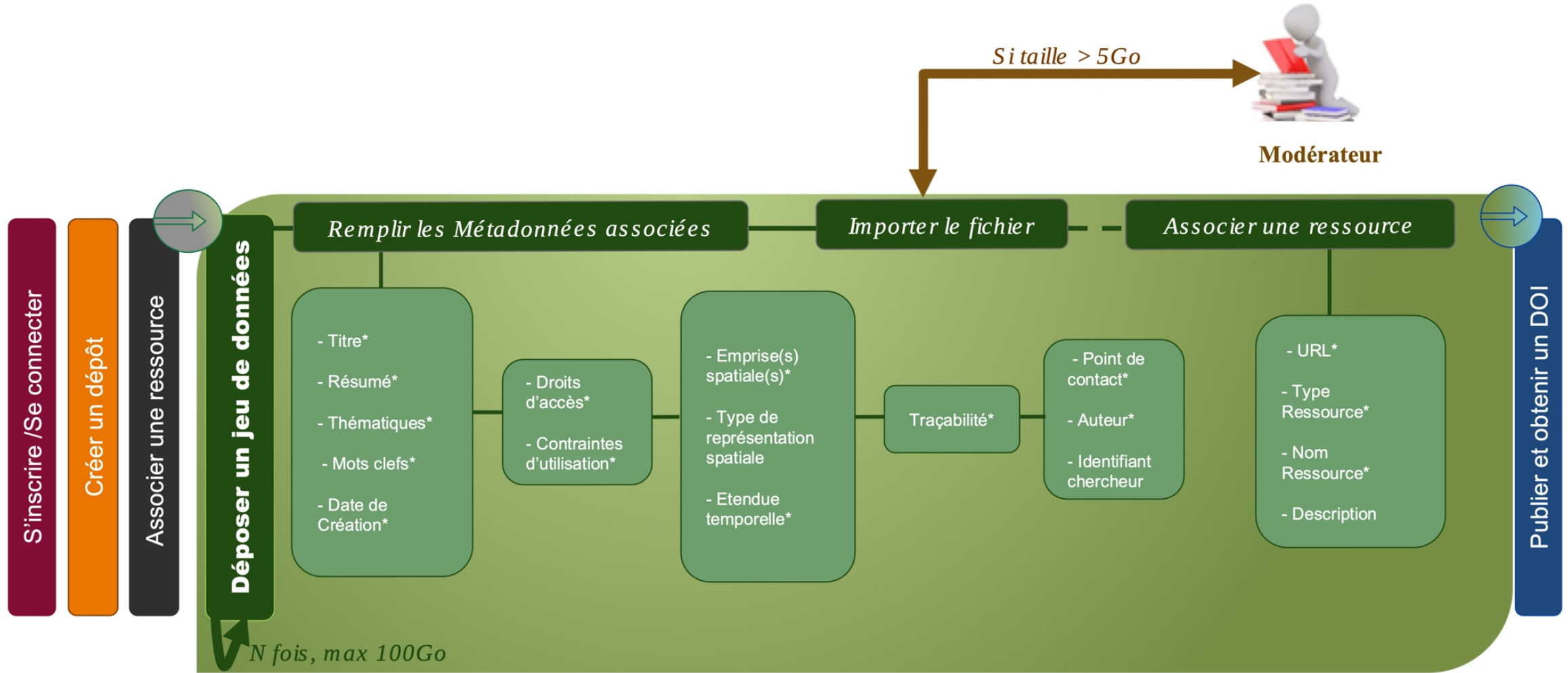


*Les champs notes avec une * sont des champs obligatoires*

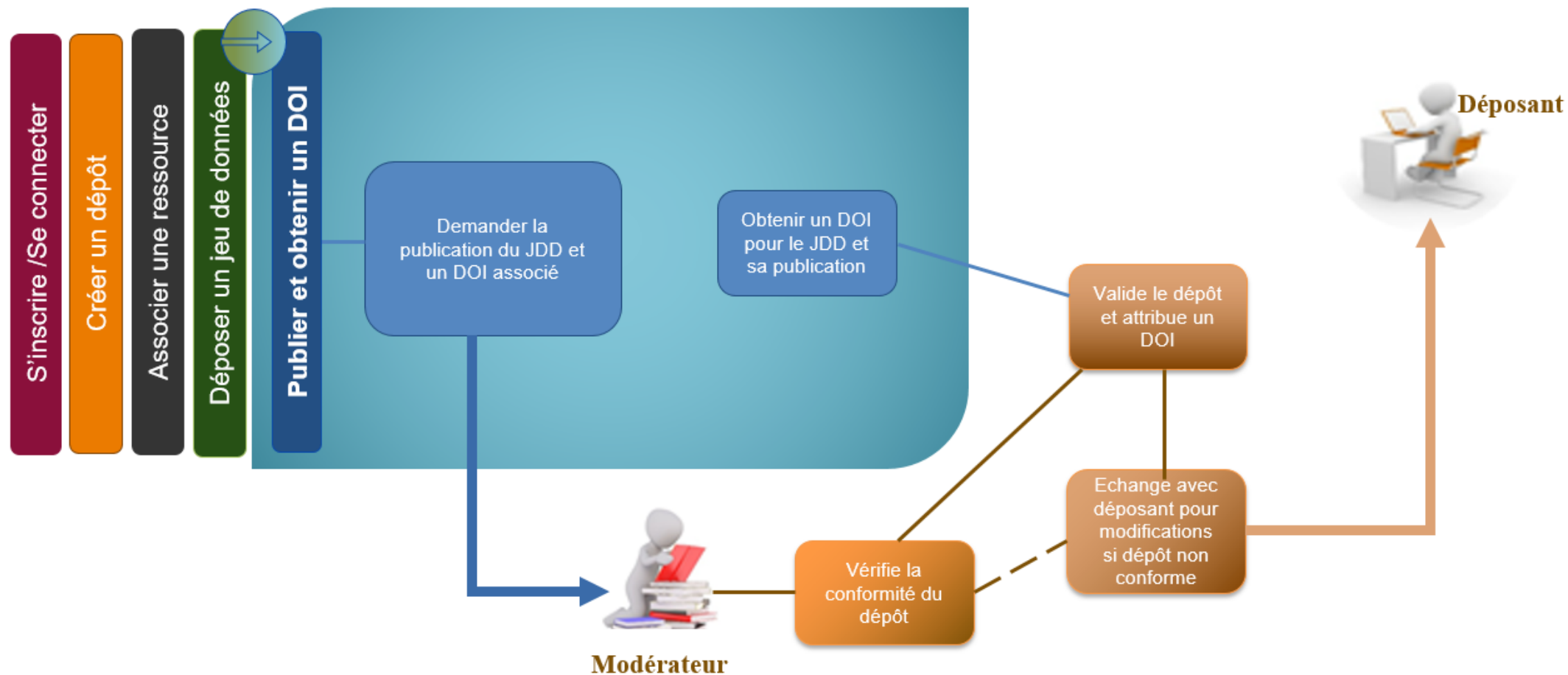
Workflow – Associer une ressource à un dépôt



Workflow – Déposer un jeu de données



Workflow – Publier un jeu de données et obtenir un DOI





DEMONSTRATION

Merci de votre attention

Pour nous contacter : entrepot@data-terra.org

Liste de diffusion :

https://groupes.renater.fr/sympa/subscribe/entrepot_dataterra

Site web : <https://www.data-terra.org/donnees-services/entrepot-de-donnees-data-terra/>



Entrepôt thématique
des données du système Terre et Environnement

Données géoréférencées

Métadonnées normalisées

Science ouverte

DOI

FAIR

Visibilité

Citation

Réutilisation

**DATA
TERRA
REPOSITORY**

Pour qui ?

- Les scientifiques du système Terre et Environnement
- Les partenaires nationaux, européens et internationaux

Quels types de données ?

Données géoréférencées :

- Issues de recherches en vue de publication
- Issues de projets

Un processus en 4 étapes :

- 1 Créer un dépôt
- 2 Associer les jeux de données
- 3 Associer des ressources
- 4 Attribuer un DOI

→ Accompagnement par les équipes ressources

www.data-terra.org/dataterra-repository/

Ne pas jeter sur la voie publique.