



UN PROJET PORTÉ PAR CLIMERI, DATA TERRA ET PNDP POUR FACILITER LES USAGES DES DONNÉES SYSTÈME TERRE ET ENVIRONNEMENT

Assemblée Générale ODATIS - Projet GAIA DATA 12-13 décembre 2023





EQUIPEX+ /PIA3

OBJECTIF : Développer et mettre en œuvre une infrastructure/plate-forme intégrée de données FAIR et de services distribuées pour l'observation, la modélisation et la compréhension du Système Terre, de la Biodiversité et de l'Environnement

- **sur l'ensemble du cycle de la donnée**, de son **acquisition** (spatiale, sols, in-situ) jusqu'à ses **multi-usages** (qualification/validation, stockage, accès, traitements/croisements de données multi-sources/extraction de connaissances, produits/services)
- **pour la communauté scientifique** contribuant à la connaissance du système Terre, de la biodiversité et de l'environnement ; **acteurs publics et privés**



Equipement Structurant
pour la Recherche /
EQUIPEX+



Budget total : 62 M€ (coûts complets)

Demande ANR-EQUIPEX+PIA3 : 16,2 M€

Forte contribution RH : 339 ETP (soit 4066 p.m.)
personnels permanents + **59 ETP (711 p.m.)** cdd

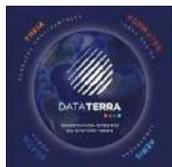
Apports additionnels des organismes : 25 postes
(recrutements cdi, postes permanents, mobilités)

21 Partenaires : CNRS, CNES, IFREMER, IRD, BRGM, IGN, INRAE, Météo-France, MNHN, CEA, IPGP, CINES, Sorbonne Univ., Univ. Grenoble-Alpes, Univ. Lille, Univ. F. Toulouse, UNISTRA, SHOM, OCA, FRB, CERFACS





PROJET PORTÉ PAR TROIS E-INFRASTRUCTURES DE RECHERCHE DU DOMAINE SYSTÈME TERRE ET ENVIRONNEMENT



Data Terra organise l'accès et les traitements intégrés de données d'observation, produits et services couvrant les différents compartiments du système Terre et leurs interactions



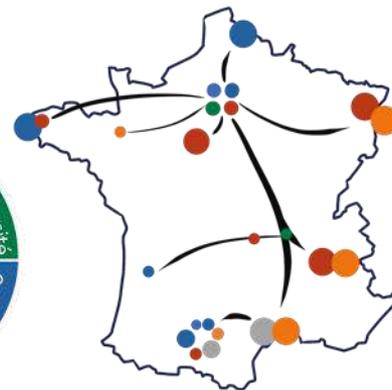
PNDB propose des outils & services pour accompagner et faciliter la compréhension, le partage et l'utilisation des données de biodiversité produites pour et par les communautés de recherche.

> EN 2024 5e pôle de DATA TERRA



CLIMERI-France produit des simulations numériques internationales pour le Programme Mondial de Recherche pour le Climat et met leurs résultats à la disposition de divers utilisateurs en France et à l'étranger.

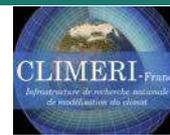
Infrastructure distribuée de services



Mettre en œuvre au plan national, européen et international une infrastructure distribuée de services innovante du domaine système Terre et environnement



CONTEXTE ET ENJEUX



Des systèmes d'information existent : IR **DATA TERRA** pour les données d'environnement, **PNDB** pour la biodiversité et **CLIMERI-France** pour les données de simulations climatiques.

Une organisation par domaine, voire par source de données, avec **des standards et des outils différents**, avec une grande diversité et un large spectre de volumes de données.

VERROUS

- Intégrer des données **hétérogènes**, **complexes**, multidisciplinaires, multi-sites
- S'adapter aux **pratiques interdisciplinaires** d'utilisation de données
- **Gestion « à la demande »** de gros volumes de données en particulier spatiales, services IA
- Prendre en compte la diversité de plateformes et infrastructures **réparties sur le territoire** et **opérées par de nombreux acteurs**
- Concilier / s'intégrer / influencer / contribuer aux dynamiques régionales, nationales, européennes et internationales



ENJEUX

- Mettre en œuvre une **plateforme intégrée** de données et services distribuées soutenues par des centres **d'expertise scientifique** du domaine
- Développer des **services accessibles**, via des portails permettant des recherches et **traitements inter et transdisciplinaires** à partir de données **multi-source** acquises par satellites, navires, avions, drones, submersibles, ballons, dispositifs in situ, inventaires, observatoires, expérimentation et des données issues de simulations de référence
- **Co-construire**, organiser et adapter les services **avec et pour les communautés scientifiques** du domaine système Terre et environnement, les **acteurs publics et socio-économiques**



AMBITION ET OBJECTIFS

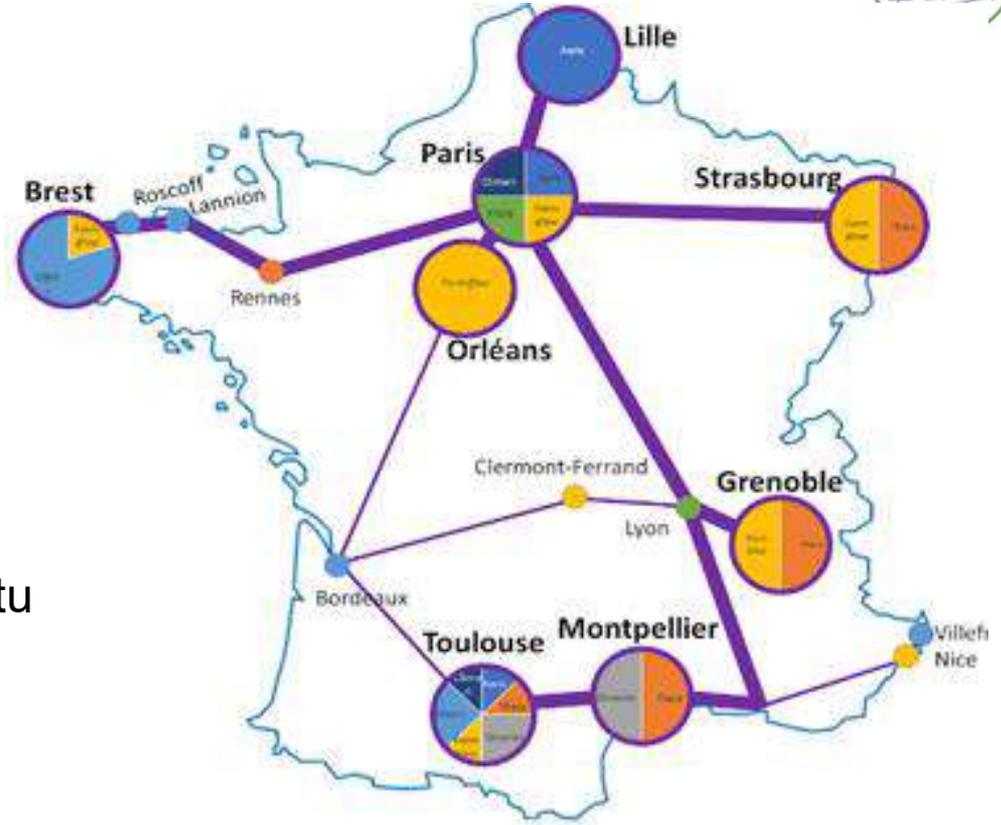


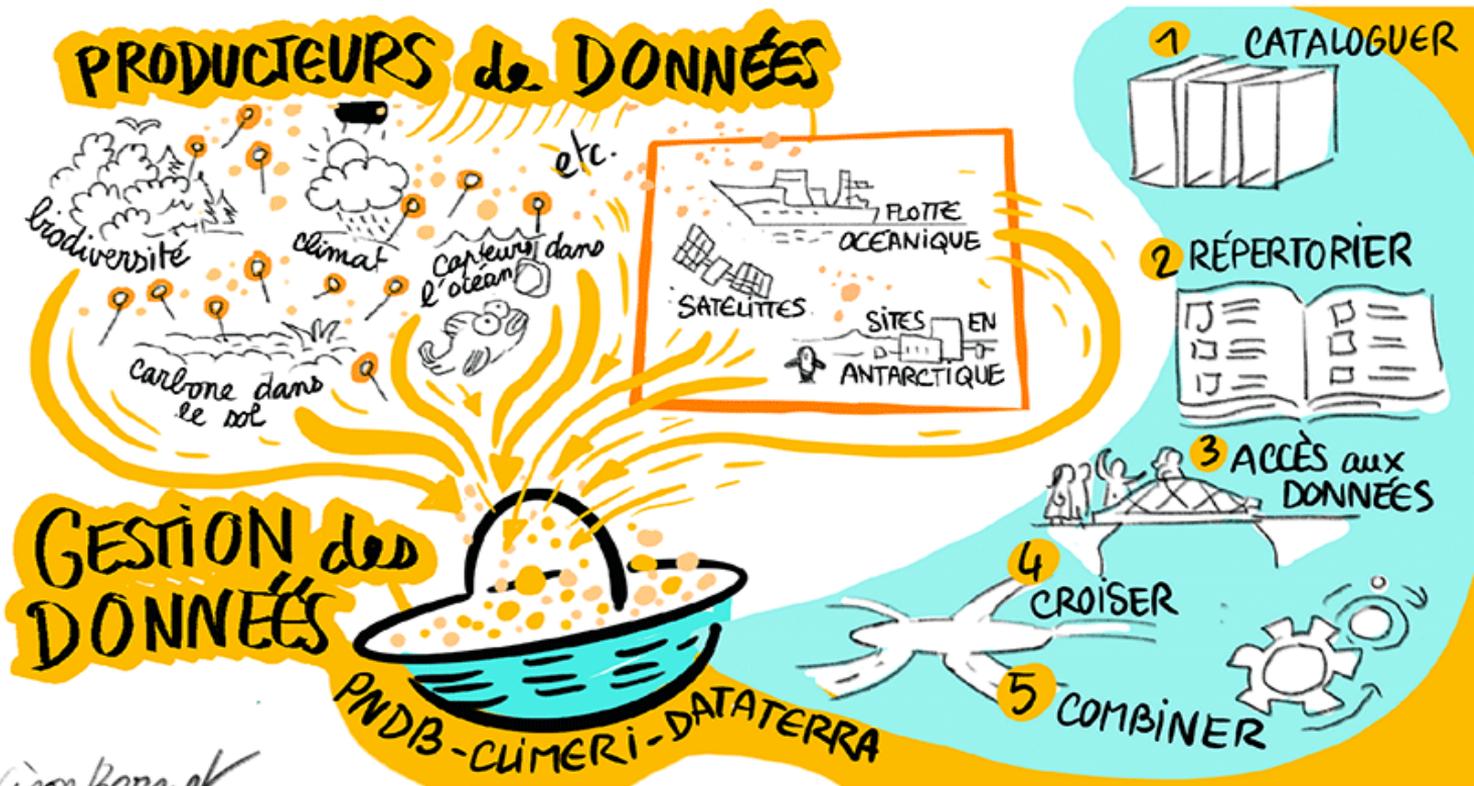
- Travailler, en étroite relation avec l'IR GENCI, les centres nationaux, CINES, IDRIS, CCIN2P3 et centres de données régionaux labélisés/Meso-Centres (**Datarmor, T-Rex**)
- Renforcer les synergies et collaborations avec les IR/IR d'observation (Terre Solide, Atmosphère, Océan, Surfaces continentales, biodiversité, ...) et IR Numérique (**IR ILICO, IR*FOF, FROOS**)
- Contribuer à la souveraineté des données et des connaissances scientifiques et technologiques (préservation des connaissances ; maîtrise de la chaîne de valeur-ajoutée : données – informations – connaissances)
- Contribuer aux initiatives nationales (science ouverte, Infranum,...), européennes (EOSC, Copernicus, DTE, ...) et internationales (GEO, GoFAIR, ONU, coopération Pays du Sud...)
- Participer à la mise en œuvre des jumeaux numériques du système Terre dans le cadre Destination Earth
- Mettre en œuvre des modèles de partenariats permettant d'associer les acteurs publics, de l'innovation, du secteur privé et de la société

8 Centres Ossatures et 30 CDS

Pour Odatis:

- Brest -> Datarmor
- Toulouse -> T-Rex
- L'ensemble des CDS Sat. et In-situ du Pôle Odatis







Les WPs du Projet

Phase de développement : 01/06/2021 – 31/05/2027

Phase d'exploitation : 01/06/2027 – 31/05/2029

Implication d'Odatis

Ifremer

Erwann Quimbert
 Erwan Bodéré
 Clémence Cotten
 Dominique Briant
 Thierry Carval
 Mickaël Dequidt
 Jérôme Detoc
 Sylvain Guyomarch
 Gilbert Maudire
 Julien Meillon
 Frédéric Merceur
 Thomas Monneraye
 Benoit Morin
 Jean-François Piollé
 Christophe Plassard
 Michael Treguer
 Marine Vernet
 Léo Bruvry Lagadec

SHOM

Valerie Cariou
 Claire Fraboul
 Jean-Paul Kirkbride
 Brieg Lailler
 Fanny Lecuy
 Thierry Schmitt
 Xavier Thuault
 Jean-Marie Vient

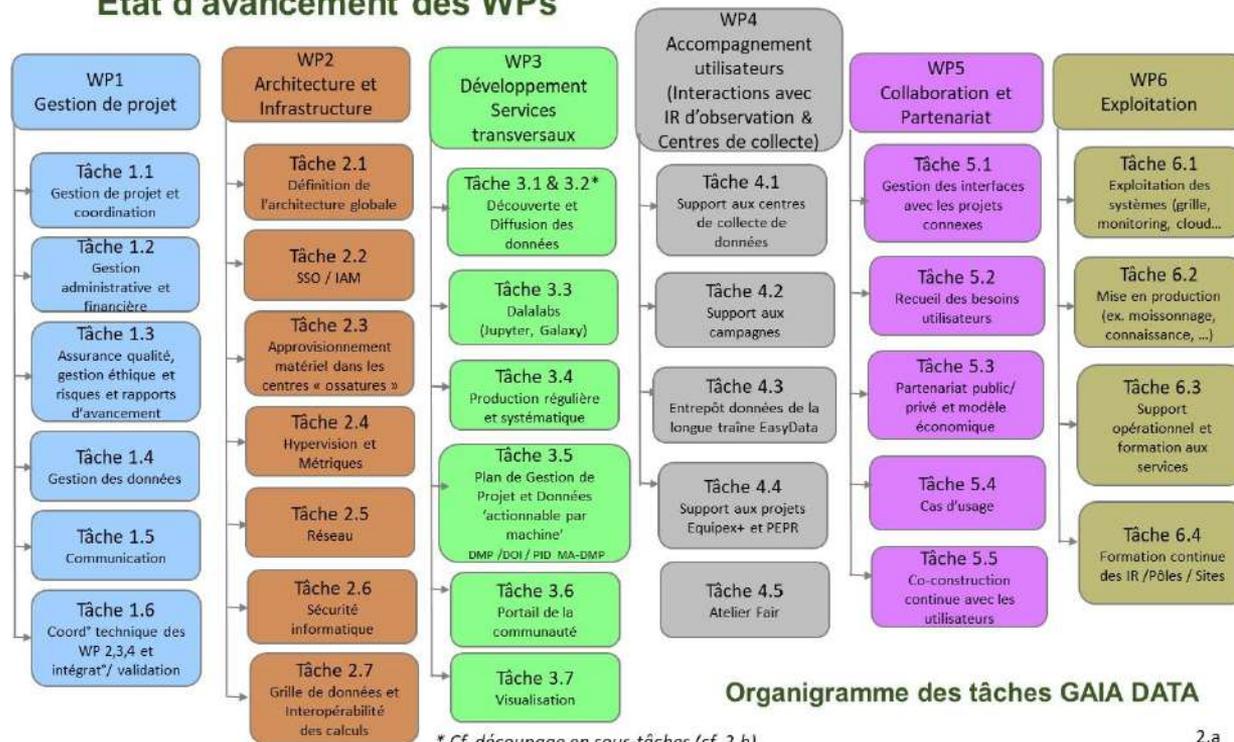
CNRS

Sabine Schmidt
 Catherine Schmechtig
 Joel Sudre

CNES

Cyril Germaineaud
 Gérald Dibarbouré
 Caroline Mercier

État d'avancement des WPs



Organigramme des tâches GAIA DATA

* Cf. découpage en sous-tâches (cf. 2.b)

2.a



Le WP3 Tâche 3.1 & 3.2

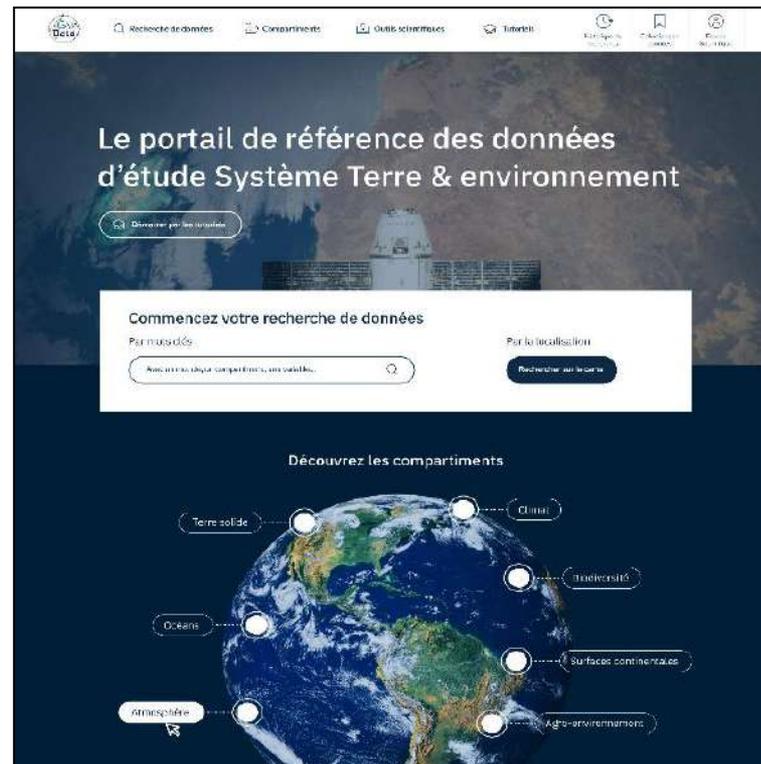
WP3
Développement Services
transversaux

Sous-Tâches ST1 à ST9

- ST1_Analyse des besoins utilisateurs en lien avec le WP5
- ST2_Architecture des services transversaux
- ST3_Identification des 1ères données à prendre en compte
- ST4_Profil d'application (modèle pivot)
- ST5_Colonne vertébrale terminologique
- ST6_Moissonneur, Transformateur, enrichisseur & persistance
- ST7_Catalogue(s)
- ST8_Portail de la connaissance & de découverte des données
- ST9_Pro prospective IA pour la découverte des données

Tâche 3.1 & 3.2*
Découverte et Diffusion des
données

2b



Un premier 'A la Google' : je peux saisir du texte dans une boîte de texte. Le système me propose des suggestions par auto-complétion basées notamment sur le thesaurus de GAIA DATA.

Deuxième mode de découverte, via l'exploration de thesaurus.

Troisième mode de découverte, via des cas d'usage .

J'accède par le portail aux services et aux Environnements de Recherche Virtuel

➤ Point sur les recrutements 2022-2023 (6 recrutements)



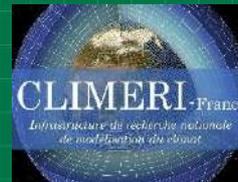
- ❖ **WP2 Services_T2.2_Authentification /Autorisations /Single Sign On**
Développeur web full-stack (CDD 'SSO'/IAM), CNRS UAR2013 IPSL Paris (Jussieu/ Palaiseau), K Ramage / G Brissebrat ; **candidat recruté : M. Terry Caloc, ~19 mois à p. 15/10/2023**
- ❖ **WP3 Services_T3.1&3.2 Découverte et Diffusion des données - ST7 Catalogues (Vocabulaire) et T3.3 DataLabs**
Ingénieur en développement et déploiement logiciels, CNRS IPSL CLIMERI Paris , G Levavasseur, ESPRI-Mod ~Catalogue et Interfaçage ; **candidat recruté : M. Laurent Troussellier, 2 ans à p. 01/06/2023**
- ❖ **WP3 Services_T3.3 DataLabs**
Ingénieur en développement de services d'environnement virtuel de recherche, 'VRE', CNRS UAR2013 Pôle ODATIS, Ifremer, Plouzane (Bretagne), JF Piollé ; **Candidat recruté : M. Gwenaël Caer, 20 mois à p. 02/10/2023**
- ❖ **WP3 Services_T3.3 VRE DataLabs (Données data fibres, VRE)**
Ingénieur-e en ingénierie logicielle E Chaljub, VAP/VRE, CNRS UAR2013 Pôle ForM@Ter (Strasbourg, Grenoble, Paris/IPGP), 'Data-Fibres' ; **Candidate recrutée Mme Margaux Mouchène, 20 mois à p. 01/10/2023**
- ❖ **Transverse WP2-WP3_T 2.7 Grille de données & Interopérabilité des calculs et T 3.3 Services DataLabs (Irods, VRE)**
Architecte services et infrastructure pour les plateformes d'exploitation de données massives du système Terre, CNRS UAR2013 Strasbourg, Jean-Philippe Malet ; **candidat recruté : M. Antoine Lamielle, ~19 mois à p. 23/10/2023**
- ❖ **WP3 Services_T3.1&3.2 Découverte et Diffusion des données-ST7 Catalogues (Web sémantique, Vocabulaire, MTEP)**
Chargée de la standardisation et de la gestion de la connaissance, IRD Montpellier JC Desconnets ; **candidate recrutée : Victoria Agazzi, 2 ans à p. 15/09/2022**

➤ Point sur les recrutements à venir (9 recrutements prévus jusqu'en 2025)





DATA
TERRA



Ce travail a bénéficié d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'Avenir Equipex+.



contact@gaia-data.org

www.gaia-data.org

