

IR DATA TERRA : E-Infrastructure de données (spatiales, sols, in-situ) et de connaissances pour une observation intégrée du Système Terre et de l'environnement

Réseau national des observatoires du trait de côte



"Valorisation des données : vers une généralisation des données téléchargeables ?"

Frédéric Huynh (IRD), directeur IR Data TERRA

Emmanuel Chaljub (UGA-CNRS), directeur pôle **Form@Ter**; Sébastien Payan (Sorbonne Univ.), directeur pôle **AERIS**

Erwann Quimbert (IFREMER), directeur pôle **ODATIS**; Anne Puissant (CNRS/UNISTRA), directrice pôle **THEIA**

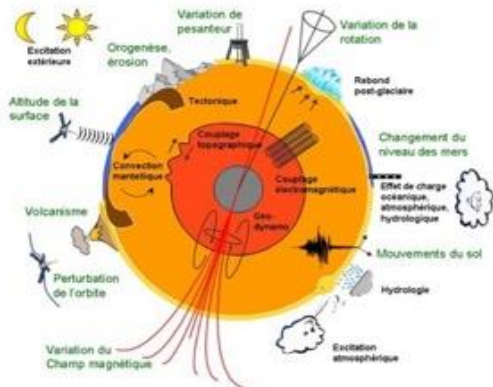
Jian-Shen Sun (MNHN) directeur **PNDB**, Karim Ramage (CNRS), **directeur technique pi**

Jean-François Faure (IRD), secrétaire exécutif **DINAMIS**; Isabelle Biajotti (CNRS), responsable **ART**



La Terre, un système complexe dynamique :

- processus géophysiques et environnementaux,
- différentes échelles spatiales et temporelles,
- interactions permanentes entre les compartiments Terre solide, surfaces continentales, océan, atmosphère et l'anthroposphère.



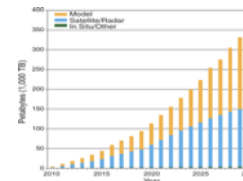
D'après O. de Viron

Comprendre ces processus nécessite l'accès et l'analyse de données complexes, nombreuses, volumineuses et d'origines multiples (satellites, in situ, campagnes, observations à long terme, résultats d'expérimentation, modèles, etc ...)

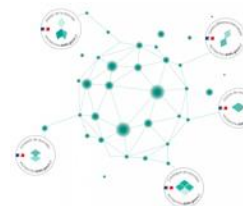
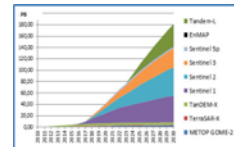
Nécessité d'un dispositif pérenne d'accès à des **données, produits et services** permettant **d'observer, comprendre et prévoir** de manière **intégrée** l'**histoire, fonctionnement et évolution** du système Terre soumis aux changements globaux

Besoins scientifiques et technologiques et principaux défis

- **Evolution des besoins des communautés scientifiques**
données **multi-sources**, **multi-capteurs**, services d'accès aux données, traitements, analyse/modélisation, IA, approches intégrées des interactions et complexité, observation à long terme
- **Augmentation exponentielle du nombre de données, diversités des sources, complexités,**
...
Spatiales, in-situ, modèles => besoins d'analyse/réanalyse, traitements intelligents
- **Méthodes et outils de FAIRisation des données spécifiques des domaines**
Découverte, méta-données enrichies, vocabulaires partagés, ...
=> écosystèmes d'acteurs
- **Co-construire une infrastructure distribuées de services spécifiques du domaine système Terre et Environnement**
=> enjeux scientifiques et données transverses
- **Concilier structuration nationale transversale à l'Europe**
=> modèle d'IR Data Terra en Europe et à l'international
- **Concilier recherche d'excellence et développement de partenariats avec les acteurs publics et économiques**
- **Contribuer efficacement aux politiques publiques, aux transitions systémiques des territoires, à l'innovation ; aux enjeux sciences et sociétés**



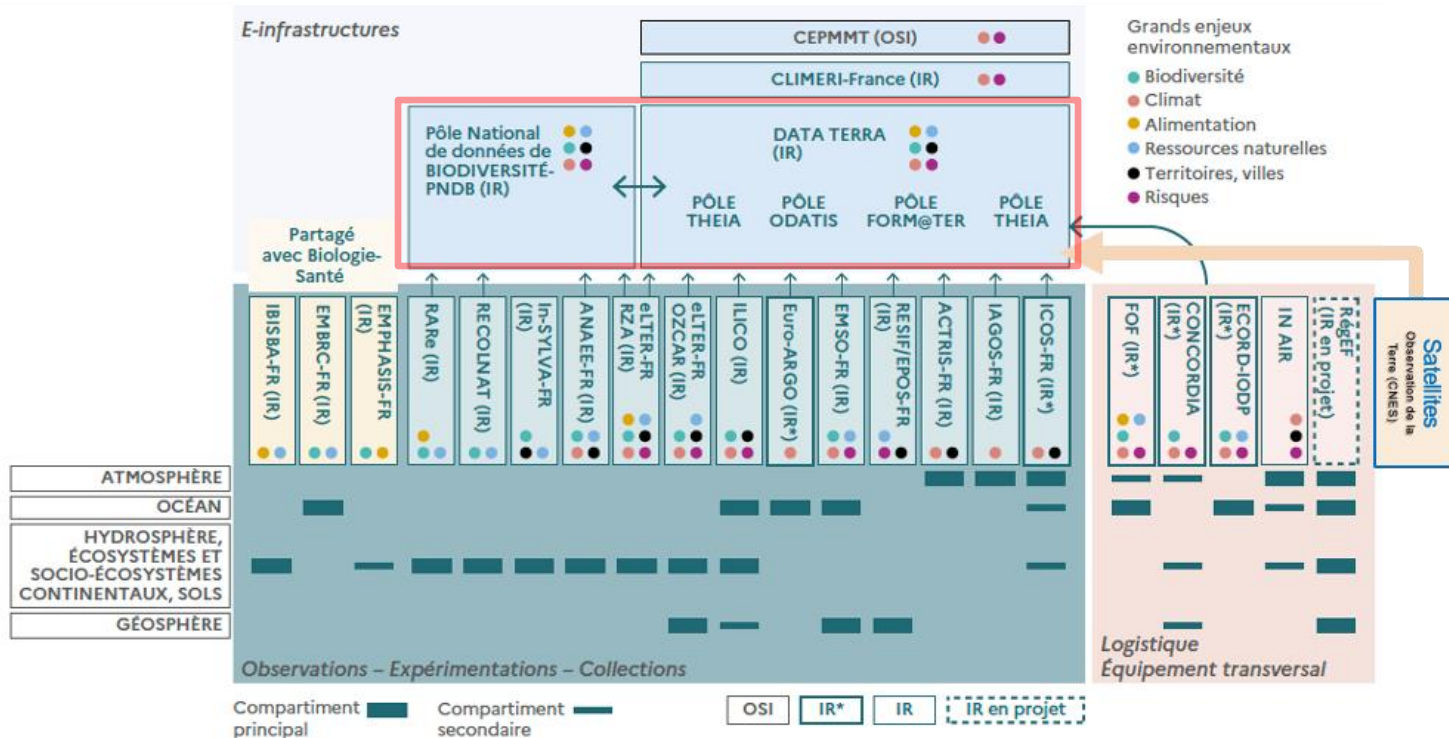
Projected increase in global climate data for climate models, remotely sensed data, and in situ instrumental/proxy data. From Overpeck et al. Science, 2011



BIG DATA

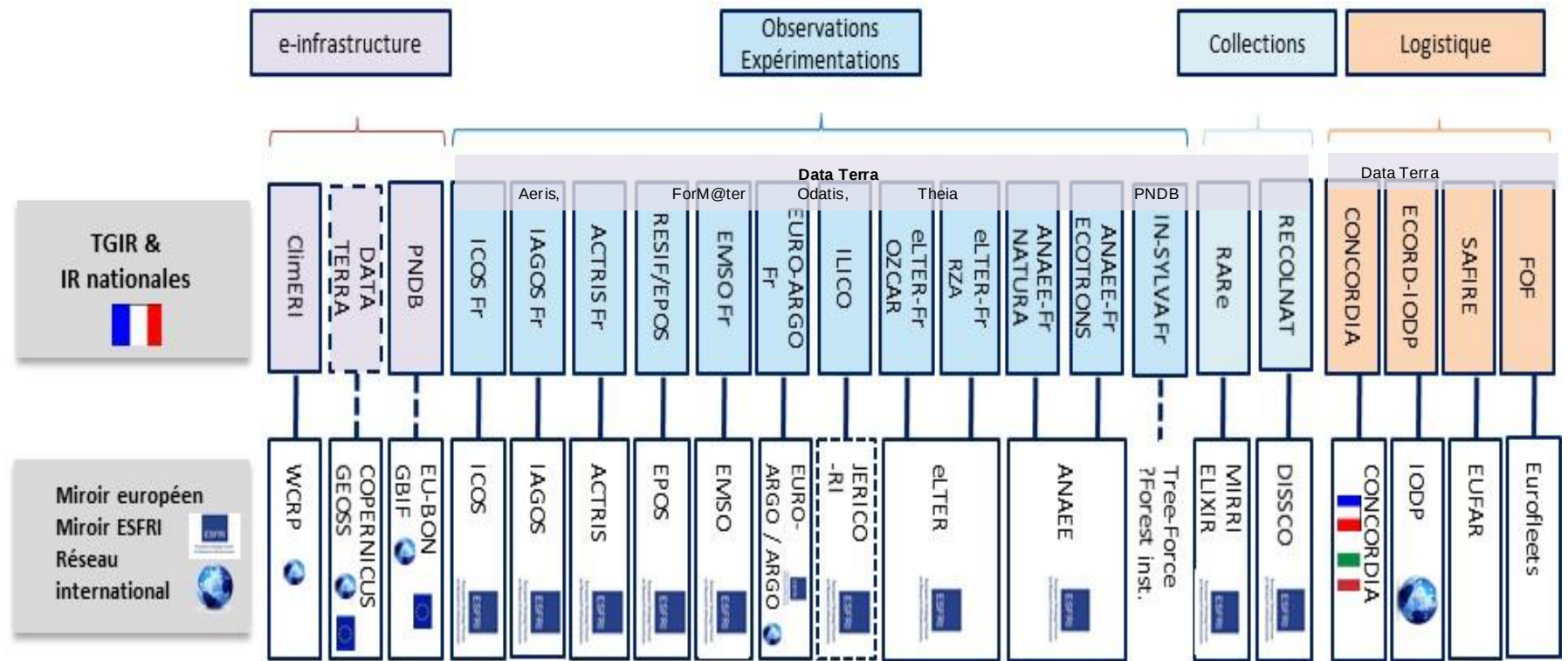


Positionnement de l'IR Data Terra - Feuille de route des IR/IR* E-Infrastructure du domaine environnement / système Terre



MESRI - FRN 2022-2025 - HC-IR, mai 2021

Positionnement de l'IR Data Terra dans la Feuille de route ESFRI des IR européennes et programmes



TGIR & IR nationales

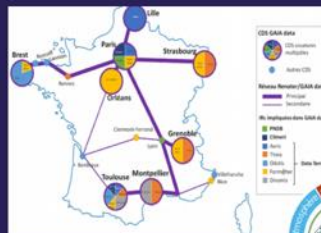

Miroir européen
Miroir ESFRI
Réseau international




Integrated platform for Earth system sciences data and services

Develop a global system of access to data, products and treatment services to analyse, understand, model and predict history, functioning and evolution of the Earth system subject to global changes

- 26 Research Organizations and Universities
- 30 Data and Services Centers (CDOS)
- 32 Consortium of Scientific Expertise (CES)
- 230 FTE / 450 scientists, engineers



DATA & SERVICES HUBS

	Atmosphere
	Ocean, coastal
	Land Surfaces
	Solid Earth
	Biodiversity
	Climat models, data simulation

CROSS-CUTTING SERVICES

	HR Spatial Imagery
	FAIR Data and Services Networking Communities



Long term observation data



constellation satellites
optic, radar...



In situ : physical and biological variables



Data from models
Climat, water, ecosystem



€42m/year



1000 products & services



15,000
permanents
users



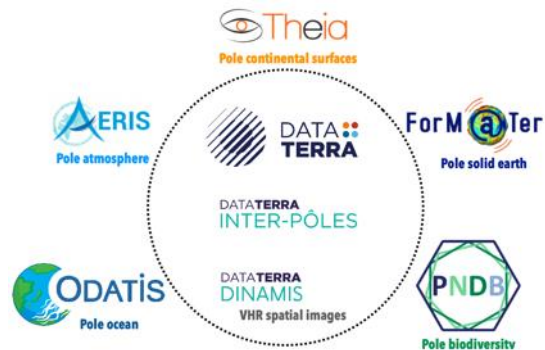
> 100 PB
(2022/2023)

Data Terra: The French Research Infrastructure for Earth system & environment sciences

AMBITION AND GOALS

A **Federated science-driven system** enabling access to and use of open **trusted** multi—domain data, products and AI-powered cross-analysis services for observing, understanding and predicting in an integrated way the history, dynamics and evolution of the Earth system, environment and biodiversity subject to global changes

IMPLANTATION DE DATA TERRA ET SES CENTRES DE DONNEES ET DE SERVICES



- 34 Research Organisations and Universities
- 30 Observing Data and Services Centres (CDOS)
- 32 Consortium of Scientific Expertise (CES)
- 450 scientists, engineers and technicians

Through the CNRS-INSU/INEE, CNES, IRD, INRAE, Univ. :

Services Nationaux d'Observations (SNOs) articulated by the national network of the Observatoires des Sciences de l'Univers (OSUs)

DATA TERRA 2024 : 5 PÔLES nationaux de données et services - compartiments du système Terre et dispositifs transversaux



Faciliter l'accès et l'utilisation des données et produits de qualité sur l'ensemble des compartiments du système Terre (Données spatiales, aéroportées, sols, in-situ)

Développer des services de visualisation et de traitements adaptés aux besoins, à l'accroissement de la volumétrie et aux avancées technologiques

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



PROJETS SCIENTIFIQUES ET TECHNIQUES

Nationaux, Européens et
Internationaux

- EOSC :
- FAIR-EASE, ...
- GAIA DATA
- PEPR ...



GROUPE DE TRAVAIL
SCIENCE, TECH, FORMATION,
EUROPE & INTERNATIONAL,
COMMUNICATION,
COOPÉRATION Pays du Sud





Services découverte, Accès et Gestion de données

Découverte (s'y retrouver parmi des milliers de jeux de données)

Catalogue (métadonnées, vocabulaires, ontologies)

Consultation et accès aux données via des web services interopérables

Visualisation avancée

Accompagnement des communautés pour la **FAIR**isation



Services transversaux pour faciliter les travaux transdisciplinaires

Grille de données, cloud, portail connaissances, SSO, métriques, support utilisateurs & formation – animation communautés

Support aux campagnes

Analysis Ready Data
Datacubes, ...



Earth Analytics Lab exploration de la donnée, bac à sable

Virtual Analysis Platform - VAP : écosystème Notebook/PANGEO/STAC

Datacubes

Capacité à se connecter directement sur les centres Traitements à la demande (WPS)

NoCode/Low Code: Galaxy-E, FG/VIP, ~Matlab/Simulink

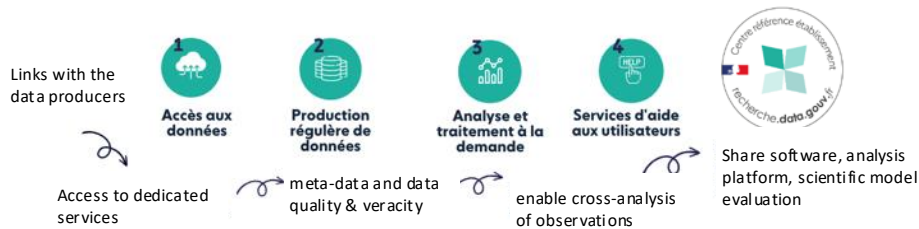


Services de production réguliers

Optimisation des traitements (outils orchestration) et formats de données (Zarr, CoG, Dask, ...)

Supporté sur un continuum d'infrastructures partagées





Atmosphere



Solid Earth



Ocean



Continental Surfaces



VHR transversal Imagery system



Thematic Data and services catalogue and scientific expertise for each pole and for transversal system Dinamis



aeris-data.fr



poiterresresolide.fr



odatis-ocean.fr



theia-land.fr



dinamis.data-terra.org

Services to data producers, products, software
Thesaurus and ontology
Support for campaigns
On demand computing services
Newsletter, project added values

Services to data producers, products, software
Thesaurus and ontology
Support for campaigns
On demand computing services
Newsletter, project added values

Services to data producers, products, software
Thesaurus and ontology
Support for campaigns
On demand computing services
Newsletter, project added values

Services to data producers, products, software
Thesaurus and ontology
Support for campaigns
On demand computing services
Newsletter, information bulletins

On-line adhesion
Services to data users
On demand computing services
Valorisation, use cases, added value data products

Service transversal

EaSy Data

Entrepôt Données longue traîne

Caractériser et analyser l'atmosphère

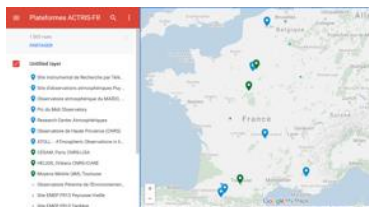
dynamique, physique et chimie atmosphérique

évolution du climat, qualité de l'air, ...

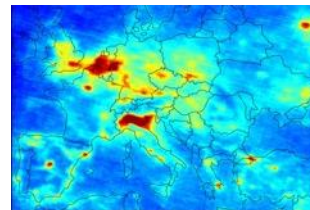
Le pôle de données et de services pour l'atmosphère **AERIS** fédère au niveau national des activités de gestion de données et d'expertise scientifique en atmosphère.

Les données proviennent de plateformes instrumentées, de réseaux d'observation, d'instruments à bords de satellites, ballons, avions, de mesures de laboratoires, inventaires, modèles, données de campagnes scientifiques.

Quelles thématiques pour les territoires ? Exemple de la qualité de l'air



**Mutualiser la
surveillance
de la qualité
de l'air**

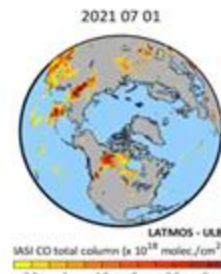


**Créer un
réseau
européen de
surveillance
de la qualité
de l'air**

**Estimer et
quantifier les
émissions de
carbone et de
particules**



**Suivre les
effets des feux
de forêts pour
la qualité de
l'air**



Le pôle de données et de services pour la terre solide Form@ter fédère au niveau national des activités de gestion de données et d'expertise scientifique en matière de terre solide.

Les données proviennent de missions satellitaires, d'instruments in situ et de résultats d'expérimentations, de modélisations et de simulations. Des services de calcul sont également disponibles.

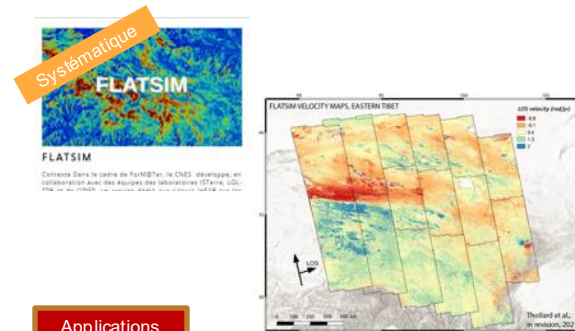
Quelles thématiques pour les territoires ?



Calcul de déformation du sol par imagerie spatiale



Suivi spatio-temporel de régions critiques



Observer et comprendre les dynamiques des océans

Le pôle de données et de services pour l'océan **Odatis** fédère au niveau national des activités de gestion de données et d'expertise scientifique pour les océans.

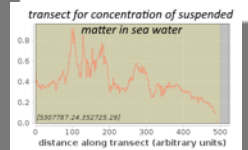
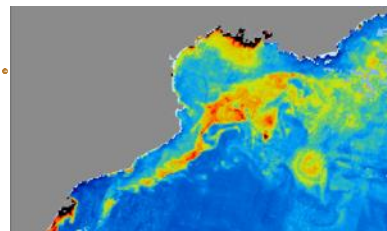
Les **données** proviennent de missions satellitaires, d'instruments in situ (plateformes fixes, flotteurs, gliders, radars, campagnes en mer, mesures en laboratoires, ...)

Quelles thématiques pour les territoires ?

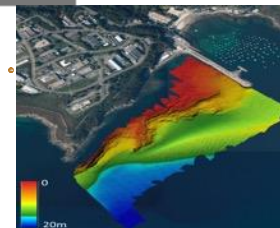


Suivi
du trait
de côte

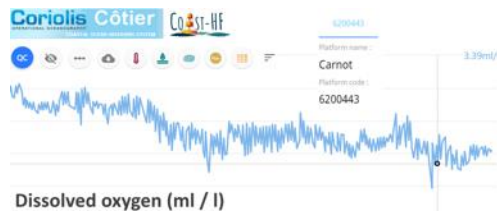
Qualité
de l'eau



Bathymétrie
LASER et
LIDAR



Oxygénation
des eaux



Observer les dynamiques des surfaces continentales et ressources naturelles

Le pôle de données et de services pour les surfaces continentales **Theia** fédère au niveau national des activités de suivi des écosystèmes et des ressources environnementales.

Les **données** proviennent de missions satellitaires et d'instruments in situ. Elles soutiennent le développement de produits, modèles et logiciels à valeur ajoutée dans dix thématiques : agriculture, biodiversité, climat, eau, forêt, littoral, neige & glace, risques naturels, santé, urbain.

Thématiques



Occupation
des sols



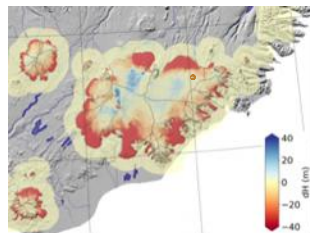
Tâche
urbaine



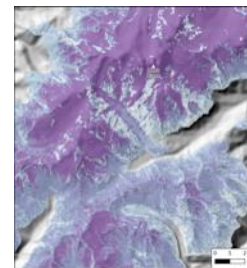
Production
de
biomasse



Humidité
des
parcelles



Changement
d'altitude
des glaciers



Couverture
neiguse

Its main objective is to provide a coherent set of tools and services for the description, access, validation, analysis and reuse of biodiversity data.

The data comes from **satellite missions and in situ instruments**. They support the development of value-added **products, models and software in ten themes**: agriculture, biodiversity, climate, water, forest, coast, snow & ice, natural risks, health, urban.

STUDY CASE

The screenshot shows the RStudio interface. The top pane displays a web browser with the URL 'RStudio - R for the Biological Data Science'. The browser content includes a list of publications and a map of the United Kingdom. The left sidebar shows the RStudio environment with a list of files and folders.

Information on species communities

Use scripts on a cluster **with 0 programming skills!** Create, edit, share, reuse workflows

Analysis and production of indicators

<https://ecology.univ-paris7.eu/>

Create & share interactive visualisations,
Rapid display and share together results on a Web using open standards

[illegible]

Permettre et faciliter l'accès et les usages des images satellitaires de très haute résolution spatiale

Le Dispositif Institutionnel National d'Approvisionnement Mutualisé en Imagerie Satellitaire, DINAMIS, s'adresse à 3 types d'utilisateurs :

Utilisateurs institutionnels autorisés français

- Pour la recherche scientifique, l'enseignement supérieur et les acteurs des territoires

Utilisateurs institutionnels autorisés scientifiques étrangers

- Pour la coopération au Sud, dans l'Union Européenne et internationale

Utilisateurs privés (France)

- Pour les porteurs de projets R&D et les prestataires de commandes publiques



Couverture métropolitaine 2020. Copyright ADS - IRD, INRAE, IGN



Littoral Sénégal, Pléiades
Copyright CNES, ADS



Montpellier. Pléiades © CNES 2020, distribution Airbus DS

Une offre institutionnelle

Couvertures spécifiques Pléiades

littoraux sableux métropolitains et guyanais, DOM-TOM...

Couvertures millésimées SPOT 6-7 France métropolitaine

Acquisitions ad-hoc Pléiades et SPOT 6-7 Monde entier - Exprimées par les UIA, quel que soit leur lieu d'affectation.

Imageries complémentaires haute résolution

Relais vers le programme CNES Spot World Heritage (Spot 1-5), plateforme CNES PEPS (Sentinel 2), Geosud, programme CNES Kalideos : RapidEye, CosmoSkyMed, TerraSar, -X, Aster...

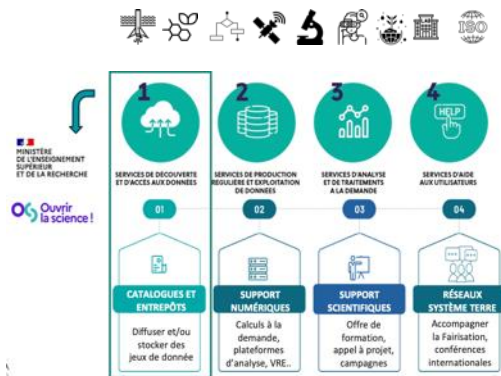
Data Terra: the national thematic reference centre for earth system & environment

Data Terra and PNDB: national thematic reference centres for Earth System, Environment and biodiversity data

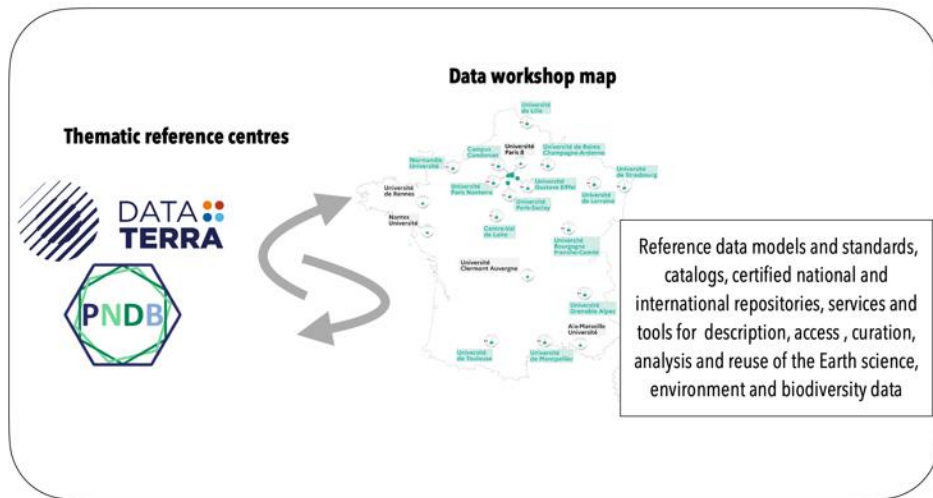
- Reference data model harmonisation (across disciplines) and standards (in disciplines)
- Thematic thesaurus and ontology
- Interoperability standards
- Machine readable catalog/registries of trusted thematic repositories, data and software
- Actionable thematic Data Management Plan models
- French contribution to the international Earth system, environment, biodiversity ecosystem

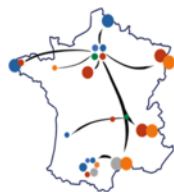
Research Data Gouv (RDG)

The national platform for sharing and opening FAIR research data and a French contribution to the EOSC ecosystem



Tools and services adapted for the RDG missions

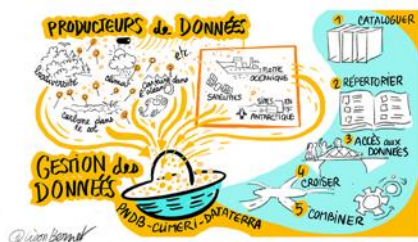
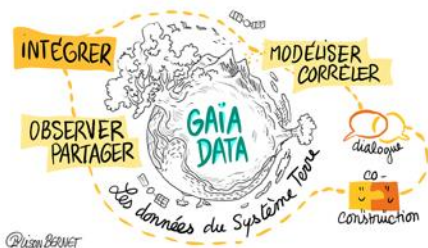




EQUIPEX+ /PIA3

OBJECTIF : Développer et mettre en œuvre une infrastructure/plateforme 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, simulation) 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€

Durée : 8 ans (06/2021 -> 06/2029)

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)

3 Infrastructures de Recherche : Data Terra, Climeri-France, PNDB

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

Equipe Projet :

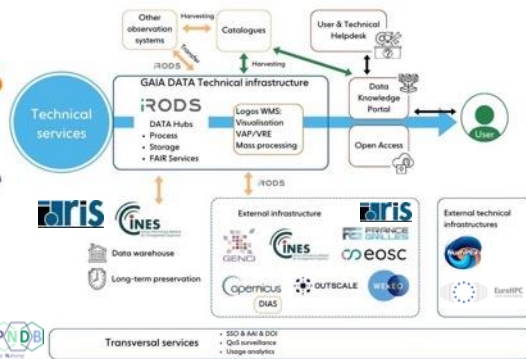
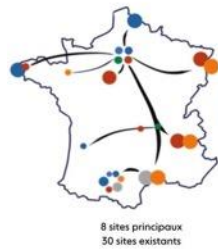
Responsable Scientifique et Technique : Frédéric Huynh (IRD/Data Terra)

Coordination Technique : Joel Sudre (CNRS/Data Terra)

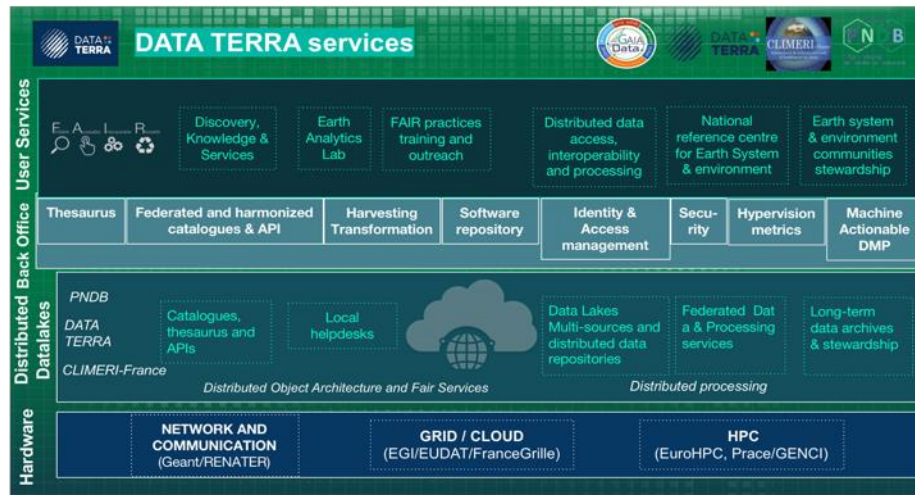
Responsables WP : K. Ramage(CNRS)/E. Chaljub(UGA) (WP2), R. Moreno (CNES)/G. Levasseur (SU) (WP3), JF Piolle (Ifremer) (WP4), C. Blanke (WP5)



Data Terra: The national E-infrastructure operating services developed through the national GAIA Data project



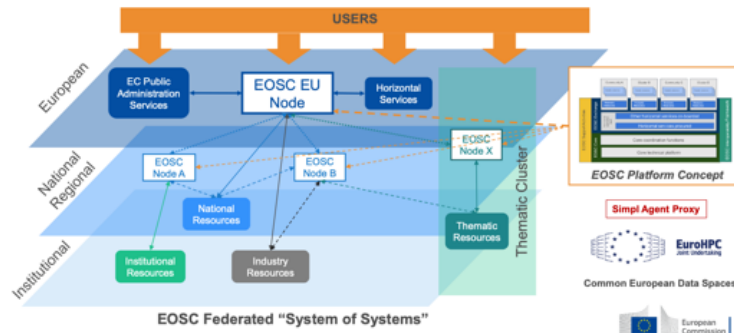
- ◆ Interoperable FAIR multi-domain data and compute services across a continuum of infrastructures
- ◆ Open AI-powered models for Earth system and environment science: e.g.,
 - ▶ foundation models for science, image analysis, event detection, data assimilation, physics-based models, Graph GNNS, surrogate models ...
- ◆ Science-driven Digital Twins projects: e.g.,
 - ▶ Extreme events, Climate, Digital Ocean, Hydrology, Geo-hazards, Territory
- ◆ Science-driven applications: e.g.,
 - ▶ connecting Copernicus, EO data, socio-economic information, IT information, scenarios ...
 - ▶ for decision making and economic activities, risk mitigation, sustainable urban areas, energy transition, water ...
- ◆ Support to National Priority Programs: e.g.,
 - ▶ One Water: “Water, Common Good”,
 - ▶ FairCarbonN: Carbon in continental ecosystems, trajectories for carbon neutrality
 - ▶ Sous-Sol: géothermie, H2 and CO2 storage, water
 - ▶ TRACCS: Transforming Climate modelling for climatic services



Data Terra: The French EOSC thematic community Resources/node for earth system, environment and biodiversity

Operate the French resources contribution to an Earth System, Environment and Biodiversity Thematic EOSC Cluster

- In synergy with ESFRI Clusters (ENVRI Fair) and other **European national thematic cluster initiatives**: e.g., NFDI@Earth and NFDI@Biodiversity (Germany), ITINERIS (Italy), UK environmental services ...
- **Sustainable science-driven thematic resources and services** integrated into EOSC portal and leveraging EOSC core services
- **Open and trusted FAIR long-term high-quality curated multi-domain and multi-component data**: observation (space, airborne, in-situ), simulation and experiments
- **Thematic reference data models and standards (data, metadata)**, machine readable catalogs, certified thematic repositories
- **Science-driven expertise and resources in-and-across integrated sub-domain**: AERIS, ODATIS, Form@Terre, Theia, PNDB, CLIMERI, DINAMIS
- **High-quality and trusted open data corpus** with standardised thesaurus and ontology, and knowledge-based evaluation and certification methodologies:
 - **ML/DL models for multi-modal data cross-analysis**
 - **Foundation models** (unimodal and multimodal) for Earth science, environment, biodiversity discovery,
 - **Science-based Digital Twins** in Earth science, environment and biodiversity



Thematic community Nodes: Other candidates





DATA**TERRA**



contact@data-terra.org

+33 (0)4 67 54 87 08

www.data-terra.org