

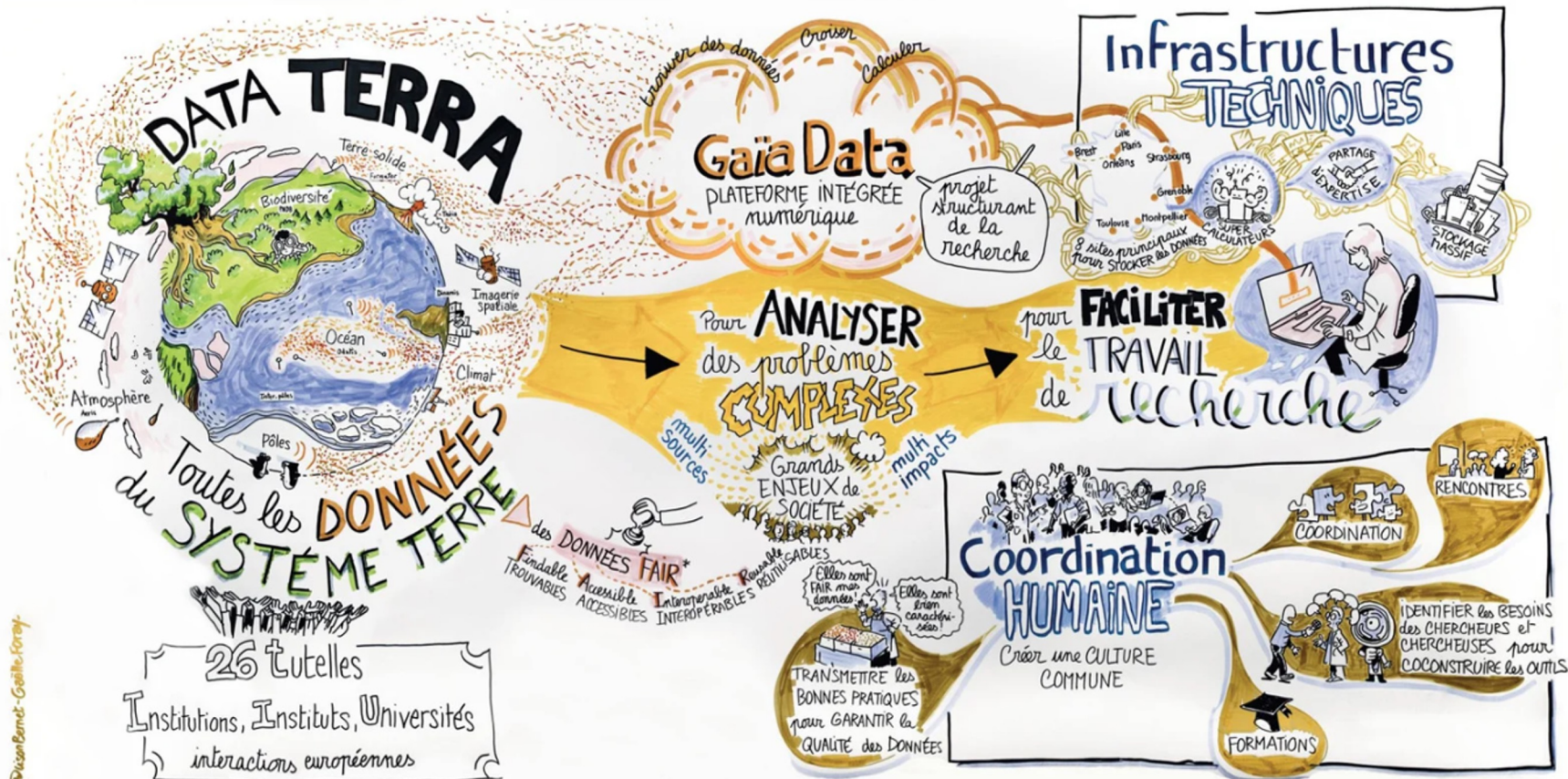


## IR DATA TERRA : E-Infrastructure de données (spatiales, in-situ) et de connaissances pour une observation intégrée du Système Terre et de l'environnement

**Frédéric Huynh** (IRD), *directeur IR Data TERRA*

Emmanuel Chaljub (UGA-CNRS), *directeur pôle Form@Ter* ; Patrice Henry (CNES), *directeur pôle AERIS* ; Erwann Qumbert (IFREMER), *directeur pôle ODATIS* ; Anne Puissant (UNISTRA), *directrice pôle THEIA* ; Richard Moreno (CNES), *directeur technique* ; Jean-François Faure (IRD), *secrétaire exécutif DINAMIS*



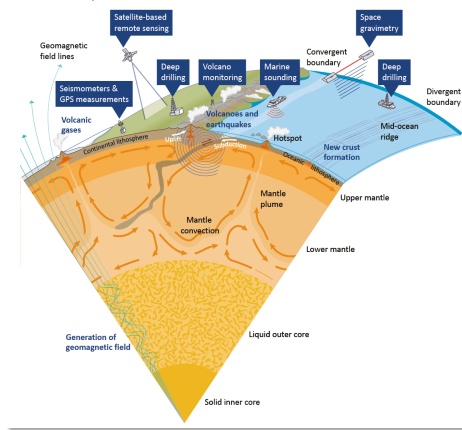


## A complex dynamical system

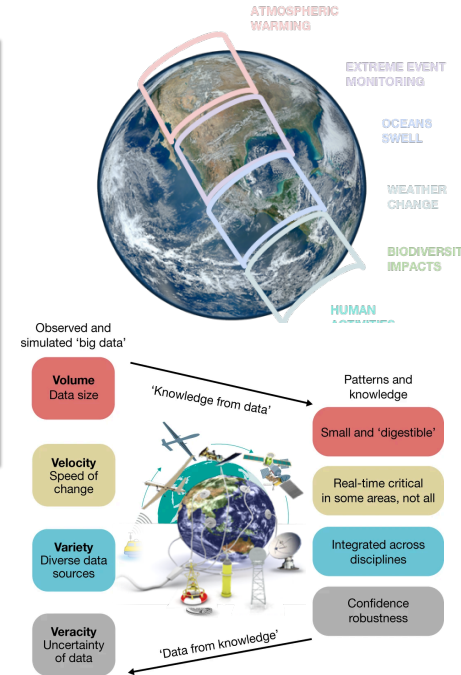
- Coupled geophysical, bio-chemical and environmental processes,
- With different spatial and temporal scales,
- And permanent interactions between the solid Earth, continental surfaces, and fluid envelopes (ocean, atmosphere),
- As with the anthroposphere.

## Evolution of the needs

- Integrated approaches to complexity and uncertainty quantification
- Multi-physics & multi-scale stochastic simulations
- Multi-domain, multi-sensor data enabling multi-wavelength analysis (*spatial, spectral, temporal & textual*)
- In-situ & space data coupling: e.g. data assimilation, inference
- Long-term data series and services
- Integrated Artificial intelligence and big data services
- Inter- and trans-disciplinary science/societal challenges



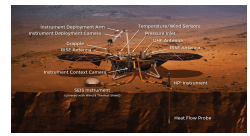
Requires **access to and cross-analysis of diverse, diverse and voluminous multi-source data** (satellites, in situ, campaigns, long-term observations as well as experimental, numerical models results, etc...)



Scientists and decision makers need to have an **integrated interdisciplinary approach** for multi-domain data access and services together with on-demand AI-enabled data analysis and modelling of increasing resolution, uncertainty quantification, and extreme natural events characterisation

# Explosion of Data fluxes and diversity

## Data Ubiquity



InSight mission



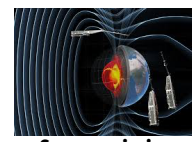
IAGOS



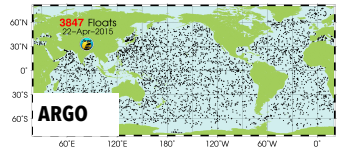
Copernicus/Sentinel



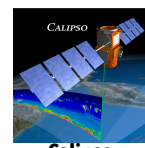
Balloon



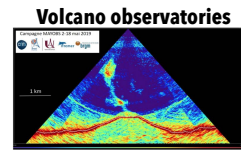
Swarm mission



ARGO

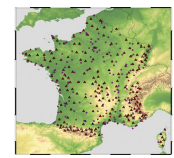


Calipso



Volcano observatories

## Seismic/geodesy



## Data explosion (rate, volume, diversité):

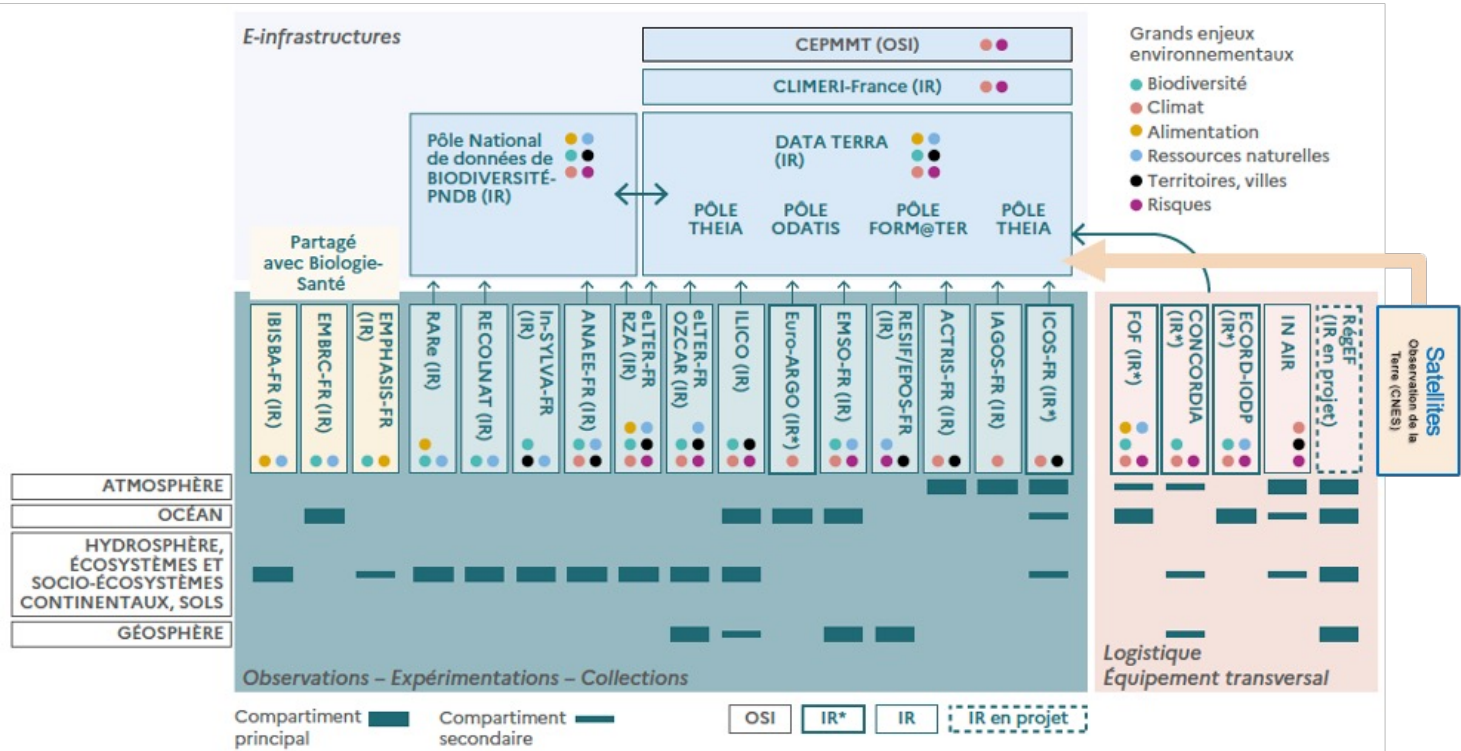
- **Edge environments:** observation, monitoring systems
- **Centralised environments (Cloud, HPC):** large ensemble of simulations, data assimilation, High-end Data Analytics

## New generation of discoveries relies upon:

- **Data acquisition:** streaming data processing/reduction/compression, primary data delivery
- **Observation continuity:** long time / on-demand short time duality
- **Data Archiving and management:** long-term archiving, curation (interoperable data reference models, provenance, distribution)
- **Data FAIR services:** multi-source virtual observatory services, DOA

- **High-performance statistical data analysis:** AI-based distributed multi-source data analysis
- **ELarge ensemble simulations:** multi-physics and multi-scale systems, uncertainty quantification, extreme events
- **Probabilistic inversion/assimilation:** high-dimensional probabilistic inference methods
- **Wide-area workflows (HPC/HDA):** data logistics in a multi-provider Edge-to-Centralised infrastructure continuum

# Positionnement des Infrastructures de Recherche du domaine système Terre et environnement - France



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

# Missions dans le domaine de Data Terra : E-Infrastructure du domaine environnemental / système Terre

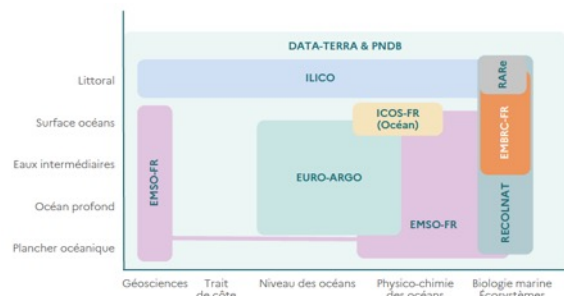


Figure 4 : infrastructures de recherche françaises dans le domaine « Océan et littoral ». L'encadré orange (EMBRIC-FR) indique une infrastructure commune avec le domaine Biologie-Santé. Les infrastructures de logistique pour l'accès au milieu océanique ne figurent pas sur cette figure.

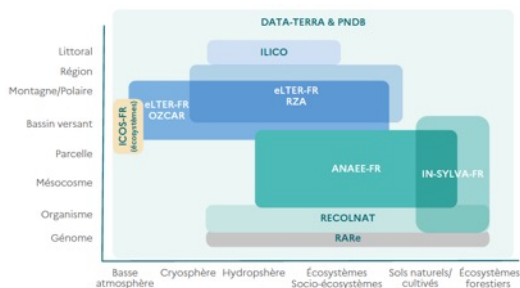


Figure 5 : Infrastructures de recherche françaises dans le domaine « Hydrosphère, écosystèmes et socio-écosystèmes continentaux et sol ». Les infrastructures de logistique pour l'accès au milieu océanique ne figurent pas sur cette figure.

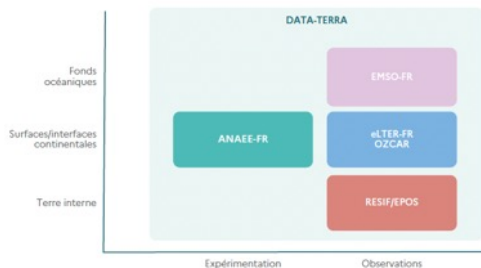


Figure 6 : Infrastructures de recherche françaises dans le domaine « géosphère ». Les infrastructures de logistique pour l'accès au milieu océanique ne figurent pas sur cette figure.

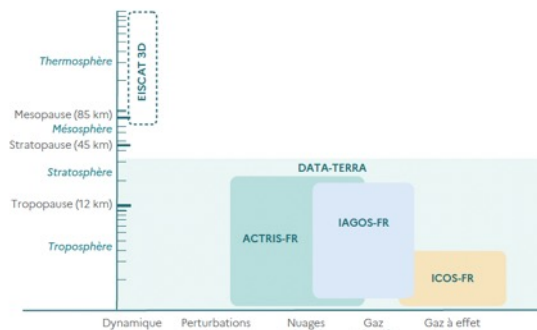
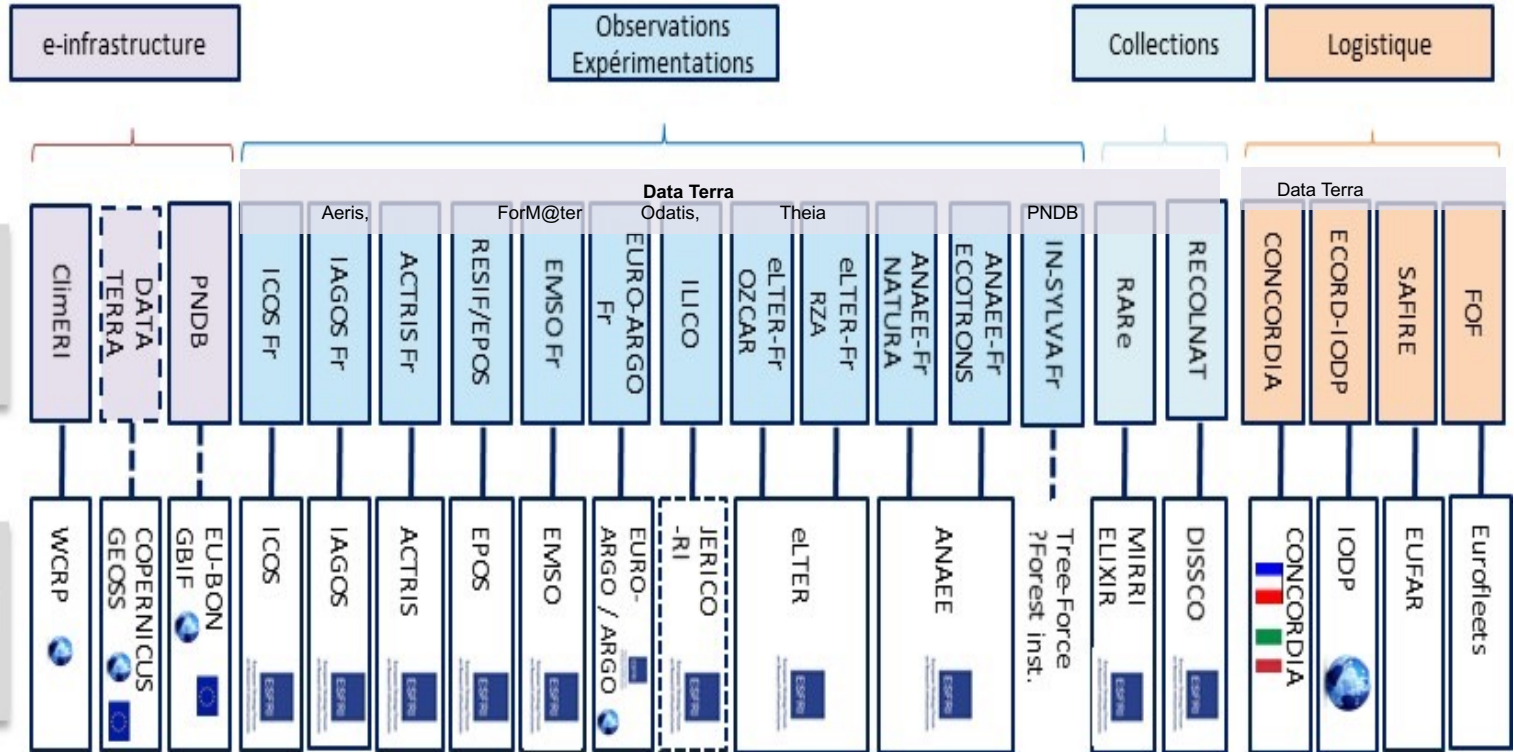


Figure 3 : Infrastructures de recherche françaises et européennes dans le domaine « Atmosphère ». L'infrastructure EISCAT-3D est une infrastructure ESFRI sans participation française. Les infrastructures de logistique pour l'accès au milieu atmosphérique ne figurent pas sur cette figure.

## e-infrastructures

**Le Pôle de données et services du système Terre DATA TERRA est une infrastructure qui se structure pour offrir un portail unique vers les quatre pôles (AERIS pour l'atmosphère, ODATIS pour l'océan, FORM@TER pour la Terre solide et THEIA pour les surfaces et interfaces continentales) ainsi que des services et outils. Il permettra de contribuer aux grands programmes européens comme COPERNICUS et internationaux comme GEOSS**

# Positionnement des Infrastructures de Recherche du domaine système Terre et environnement - Europe



TGIR & IR nationales

Miroir européen  
Miroir ESFRI  
Réseau international

# Construction et vision à long terme

## Processus d'intégration inter-organismes

AERIS  
THEIA  
ODATIS  
Form@ter

SATMOS (1986)  
Pôle thématique ETHER (1996)

Pôle thématique ICARE  
MEDIAS/SEDOO  
Postel  
CADTS

Création du pôle thématique surfaces continentales

Rapport "Podaire"  
Mission de préfiguration des pôles

Rapport « Waiteufel »



1996

2012

2013

CIO Pôle système Terre  
Nomination des 3 chargés de mission et Directeur pour Theia

Création AERIS 16/12/2014  
Comités Directeurs et BE

Création ODATIS  
Création du CS  
Premier portail Form@ter

Assemblée Générale  
Création de GTs  
Succès AO PIA3 ETALAB

Signature 2<sup>ème</sup> convention Theia  
Lancement site web  
Equipe de direction  
Portail FORM@TER

REVEX AERIS  
Avis très positif du groupe de revue  
Service Opérationnel FLATSIM  
Imagerie radar et optique

Signature convention FORM@TER

Signature de convention ODATIS prévue



2013

2014

2015

2016

DINAMIS

2017

2019

2020

GT Inter-pôles

Création des différentes instances : CD, BE, CS, CES, ART

2015 – 2018 Nouvelles tutelles : ONERA Cerema AgroParisTech

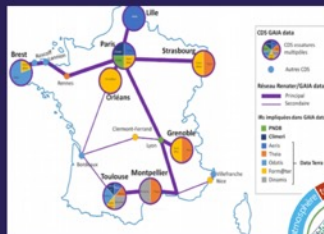
Portail OZCAR - THEIA





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

- AERIS** Atmosphere
- ODATIS** Ocean, coastal
- Theia** Land Surfaces
- ForM@Ter** Solid Earth
- PNDDB** Biodiversity
- CLIMERL** Climat models, data simulation

## CROSS-CUTTING SERVICES

- DATA TERRA DINAMIS** HR Spatial Imagery
- DATA TERRA INTER-PÔLES** 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 : 4 PÔLES compartiments du système Terre et dispositifs transversaux



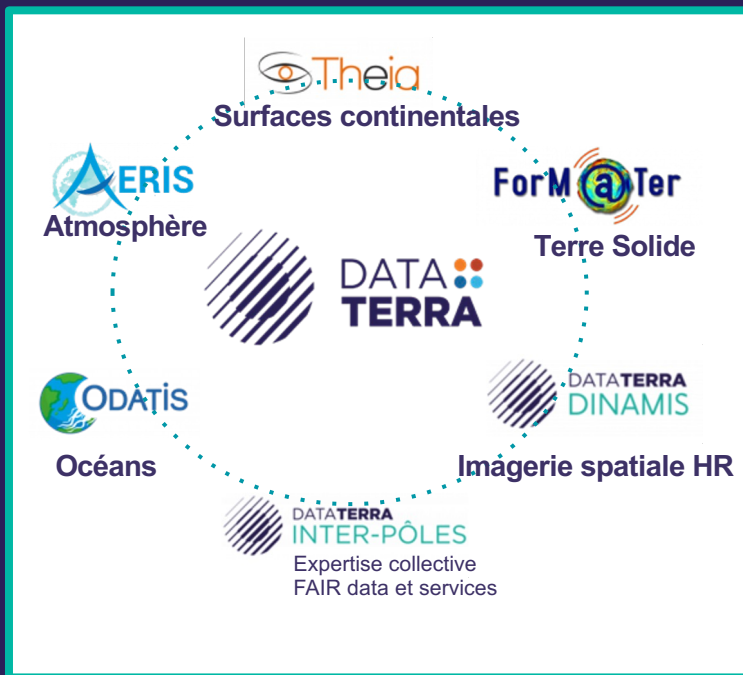
Faciliter l'accès et l'utilisation des données et produits de qualité sur l'ensemble des compartiments du système Terre (Données spatiales, aéroportées, sols, in-situ)

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

Favoriser la mutualisation, l'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 ...



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



DATA  
TERRA



# DATA TERRA, une e-Infrastructure de Recherche dédiée au système Terre

L'IR DATA TERRA propose des services aux données d'observation du système Terre et de l'environnement

Liens avec les producteurs de la donnée



**Accès aux données**

Accéder à des services dédiés



**Production régulière de données**

Améliorer la qualité des méta données



**Analyse et traitement à la demande**

Faciliter le croisement des observations



**Services d'aide aux utilisateurs**



Plateforme d'analyse, Évaluation modèle

**service transversal**

**EaSy Data**

Entrepôt Données longue traîne

## En 2024, 5 pôles de données et de services

<b>AERIS</b>	<b>FORM@TER</b>	<b>ODATIS</b>	<b>THEIA</b>	<b>PNDB</b>
Atmosphère	Terre solide	Océan	Surfaces continentales	Biodiversité

**dispositif transversal**

**DINAMIS**

Accès données satellites THRS

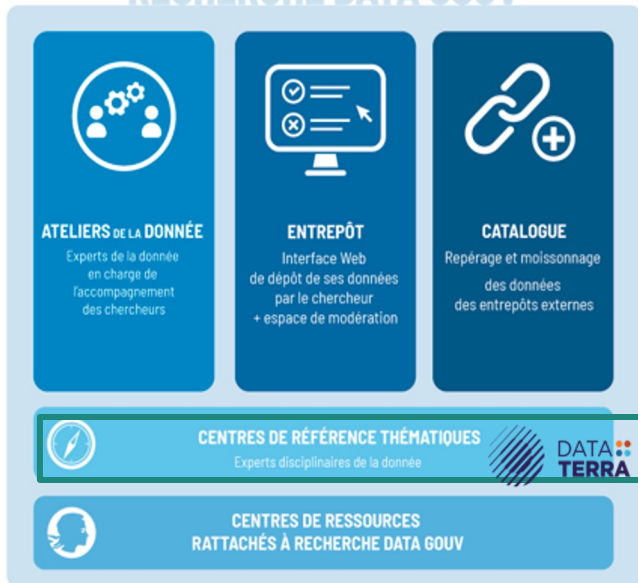
# Recherche Data Gov

Data Terra => Centre national de Référence Thématique système Terre et Environnement



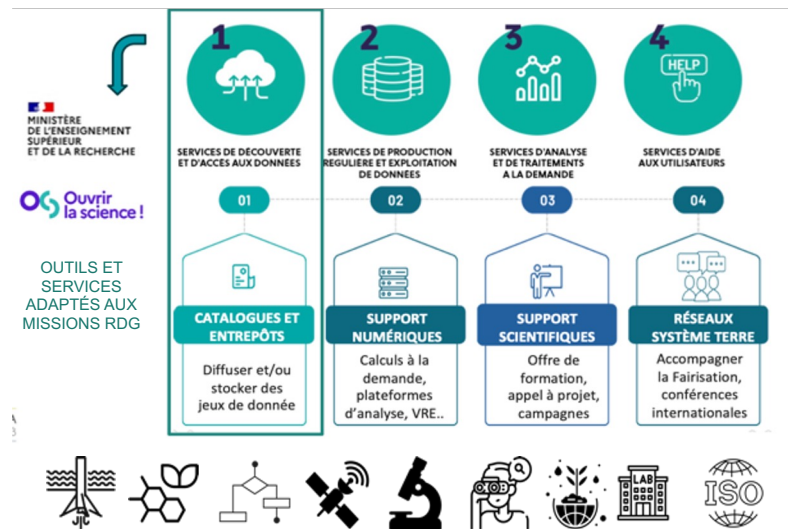
Un écosystème au service du partage et de l'ouverture des données de recherche FÉDÉRER, ACCOMPAGNER, PARTAGER, OUVRIR, RÉUTILISER

## RECHERCHE DATA GOUV



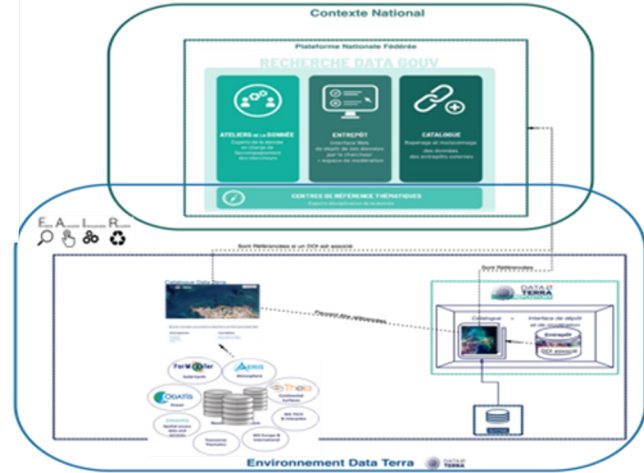
### Les centres de référence thématiques de Recherche Data Gov :

- Définissent les normes internationales de description des données
- Définissent les normes internationales de diffusion des données (ouverture, période d'embargo, accès restreint)
- Définissent et diffusent les bonnes pratiques de collecte, documentation, traitement, et diffusion des données
- Définissent la liste des entrepôts de données de référence de leur domaine thématique (nationaux et/ou internationaux) vers lesquels orienter les chercheurs pour le dépôt des données et que Recherche Data Gov moissonnera
- Contribuent à la définition de l'arborescence thématiques des données de l'entrepôt Recherche Data Gov
- Soutiennent l'articulation entre les dispositifs thématiques spécialisés et Recherche Data Gov





## DATA TERRA : accès unifiés pour accès aux données et services d'observation du système Terre et de l'environnement



Atmosphère

Terre solide

Océan

Surf.continentales

### 📌 Catalogue de données et de services



### 📌 Services associés

Support aux campagnes, services aux producteurs de données, cluster de calculs, logiciels et produits thématiques, ateliers thématiques, appel à projets, newsletter, adhésion en ligne, etc

Recherche Data Gov.  
Data Terra : Centre de Référence Thématique



**Données orphelines- longue traîne**



# L'ENTREPÔT DES DONNÉES DE LONGUE TRAÎNE DU SYSTÈME TERRE ET DE L'ENVIRONNEMENT

## POUR QUI ?

- ✓ LES SCIENTIFIQUES DU SYSTÈME TERRE ET ENVIRONNEMENT
- ✓ LES PARTENAIRES NATIONAUX, EUROPÉENS ET INTERNATIONAUX

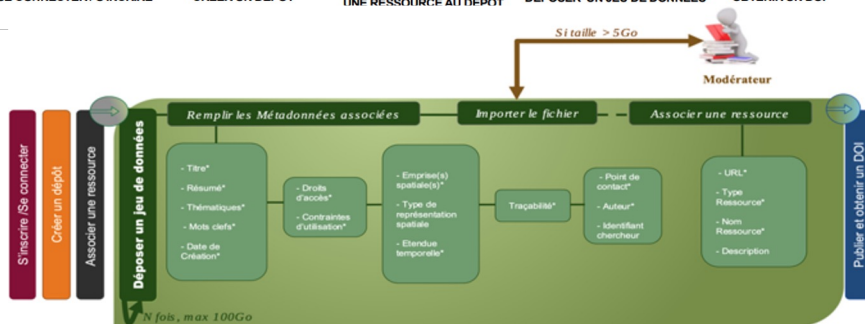
## QUEL TYPE DE DONNÉES ?

- ✓ DONNÉES LONGUE TRAÎNE
- ✓ ISSUES DE RECHERCHE EN VUE DE PUBLICATION
- ✓ ISSUES DE PROJETS

## Un procédé simple en 5 étapes clés



## Un modérateur pour accompagner le déposant



## Pérennisation et visibilité des données :

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

## Simplicité du dépôt :

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

## Qualité des métadonnées et interopérabilité :

- Norme ISO 19115-3,
- Interface adaptée aux spécificités des données géo-référencées
- Thesaurus communautaires
- Services avec possibilité d'associer des ressources aux données (publication, code source, site web projet, ...)

# Caractériser et analyser l'atmosphère

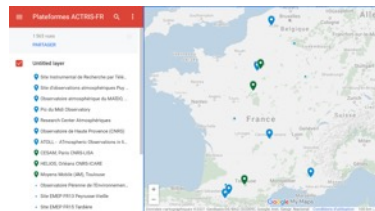
## dynamique, physique et chimie atmosphérique

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

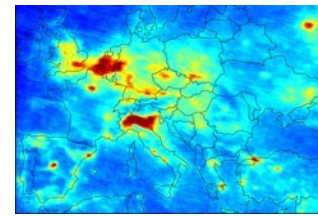
Le pôle de données et de services pour l'atmosphère **AERIS** fédère au niveau national des activités de gestion de données et d'expertise scientifique en atmosphère.

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

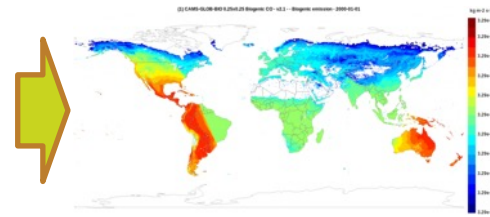


Mutualiser la surveillance de la qualité de l'air

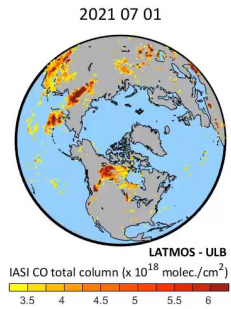


Créer un réseau européen de surveillance de la qualité de l'air

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 ?

**A la demande**

Merci de vous connecter avant de commencer.  
(Lors de la première connexion, suivre les instructions)

**GDM-OPT**  
GDM-OPT signifie Ground Deformation Monitoring using OPTical image time series

**DSM-OPT**  
DSM-OPT signifie Digital Surface Models from OPTical stereoscopic very-high

**GDM-SAR**  
GDM-SAR signifie Ground Deformation Monitoring using inSAR image time series

Calcul de déformation du sol par imagerie spatiale

**Applications**

Activation de l'initiative CREST<sup>2</sup>

ERUPTION DE LA SOUFRIERE SAINT-VINCENT: EVALUER LES ALEAS VOLCANIQUES EN TEMPS REEL GRACE A L'IMAGERIE OPTIQUE HAUTE-RESOLUTION

**Systématique**

**FLATSIM**

Comme dans le cadre de ForM@Ter, le CNRS développe, en collaboration avec des équipes des laboratoires ITRF, IGN, IRD, un service dédié aux territoires critiques.

FLATSIM VELOCITY MAPS, EASTERN TIBET

Thöniard et al., 2016

**Applications**

Déformations de la surface de la Terre à l'échelle continentale

Suivi spatio-temporel de régions critiques

→ grandes zones de failles actives et systèmes magmatiques, glissements de terrain et déformation lithosphérique à l'échelle de grands massifs...



The solid earth data and services center Form@ter brings together **data management activities** and **scientific expertise** in solid earth at the national level.

The data comes from **satellite missions**, **in situ instruments** and the **results of experiments**, **modeling** and **simulations**. **Calculation services** are also available.

### STUDY CASE

#### GROUND DEFORMATION CALCULATION BY SPATIAL IMAGERY

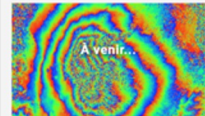
A la demande



**GDM-OPT**  
GDM-OPT signifie Ground Deformation Monitoring using OPTical image time series



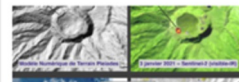
**DSM-OPT**  
DSM-OPT signifie «Digital Surface Models from OPTical stereoscopic very-high



**GDM-SAR**  
GDM-SAR signifie Ground Deformation Monitoring using INSar image time series

#### SPATIAL TEMPORAL MONITORING OF CRITICAL REGIONS

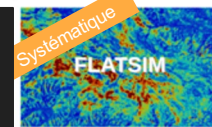
Activation de l'initiative **CIEST<sup>2</sup>**



ERUPTION DE LA SOUFRIERE SAINT-VINCENT: EVALUER LES ALEAS VOLCANIQUES EN TEMPS REEL GRACE A L'IMAGERIE OPTIQUE HAUTE-RESOLUTION

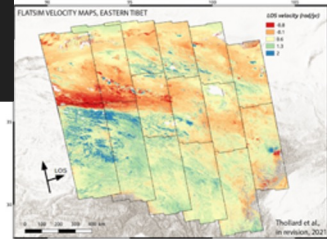
#### APPLICATIONS

Large active fault zones and magmatic systems, landslides and lithospheric deformation on the scale of large massifs.



**FLATSIM**

Commence Dans le cadre de ForM@Ter, le CNRS développe, en collaboration avec des équipes des laboratoires (Sterre, UCL)



FLATSIM VELOCITY MAPS, EASTERN TIBET



# ForM@Ter : services de calcul



## Traitement (intensif) de données

Suivi de la déformation du sol par imagerie spatiale

- **GDM-OPT** (SLIDE, ETQ, ICE)

Corrélation optique Sentinel 2

- Visualisation série temporelle (TIO, Volat et al., 2017)

Possibilité de rendre les résultats publics

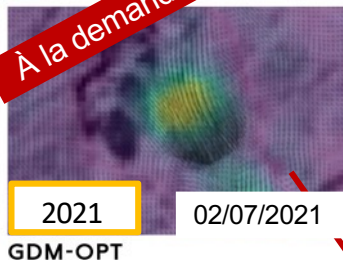
- **DSM-OPT** : production de **MNS** à partir d'images Pléiades

- paramètres prédéfinis suivant le type de paysages de la zone d'étude
- requêtes images stéréo directement via DINAMIS
- catalogue MNS
- automatisation dans le cadre de la CIEST
- Modèle économique

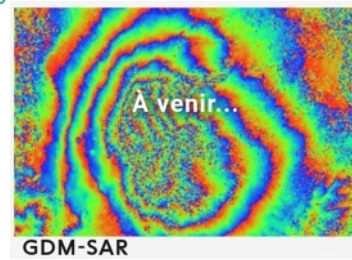
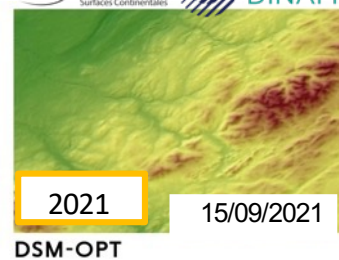
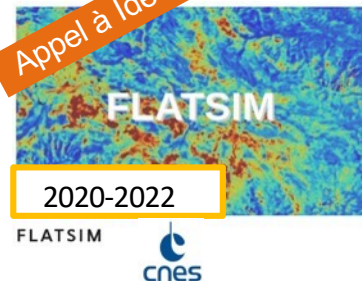
- **FLATSIM** : inSAR sur archive complète Sentinel-1

- 2ème AI 07/2022 superficie > 10 Millions km2
- Début de production au CNES

À la demande



Appel à Idées



Activation



8 activations en 2021  
3 en 2022

6 (bientôt 7) services de calculs (4 à la demande, 2 sur décision)

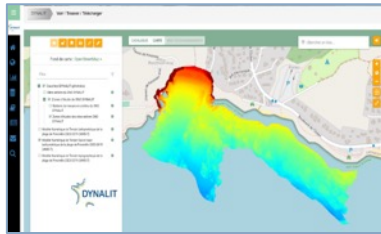
- **GDM-SAR** : calcul à la demande données Sentinel-2  
ouverture imminente

# 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.

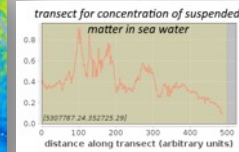
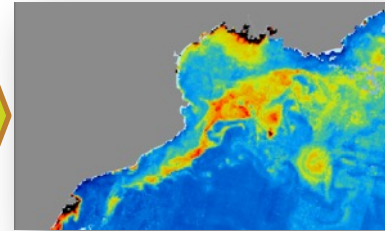
Les **données** proviennent de missions satellitaires, d'instruments in situ (plateformes fixes, flotteurs, gliders, radars, campagnes en mer, mesures en laboratoires, ...)

## Quelles thématiques pour les territoires ?



Suivi  
du trait  
de côte

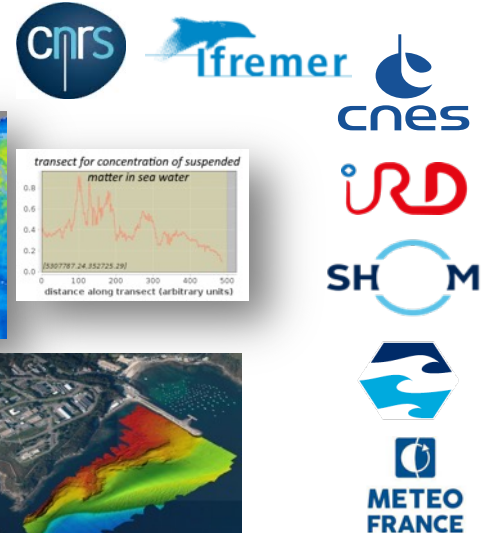
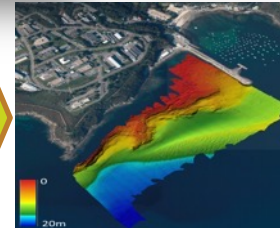
Qualité  
de l'eau



Oxygénation  
des eaux



Bathymétrie  
LASER et  
LIDAR

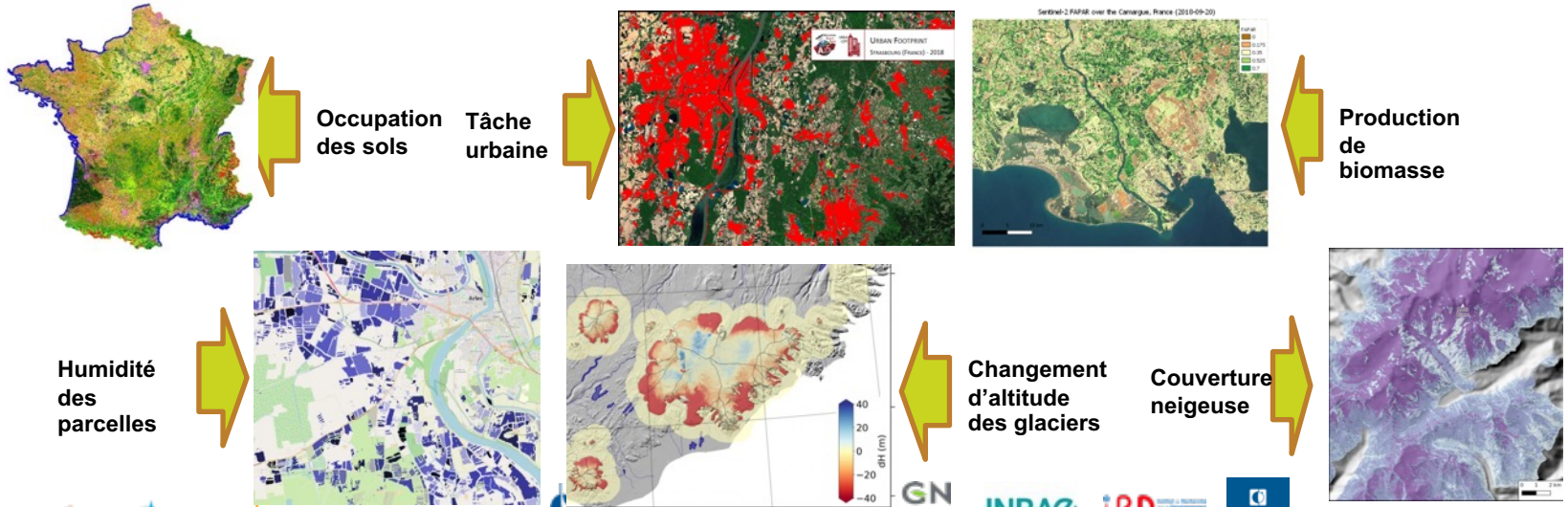


# Observer les dynamiques des surfaces continentales et ressources naturelles

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.

## Quelles thématiques pour les territoires ?



# LAND SURFACES

## Observe the dynamics of land surfaces and natural resources

The Theia land surfaces data and services hub federates **ecosystem** and **environmental resource monitoring activities** at the national level.

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

SOILS OCCUPATION



URBAN PLOTS



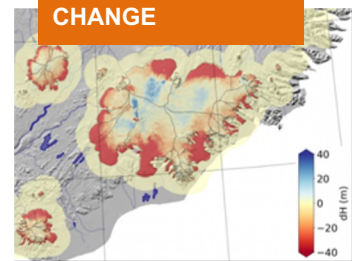
BIOMASS PRODUCTION



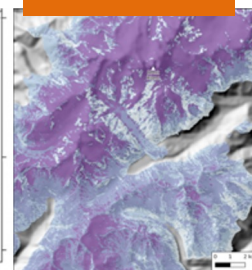
HUMIDITY PLOTS



GLACIER ALTITUDE CHANGE



SNOW COVER



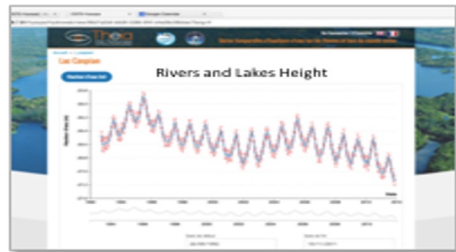
[www.theia-land.fr](http://www.theia-land.fr)



## HYDROWEB : WATER LEVELS OF RIVERS AND LAKES

Continuous, long-duration time-series of the levels of **64 lakes** and **248 virtual stations on rivers** (500 virtual stations at the end of 2018)

**Satellite altimetry is used to measure the water level** currently Jason-3, Sentinel-3A and in the future Sentinel-3B, Jason-CS and SWOT



## SOIL MOISTURE MAP VERY HR RESOLUTION

### Source data

**Sentinel 1** : main source for radar signal inversion with neural networks

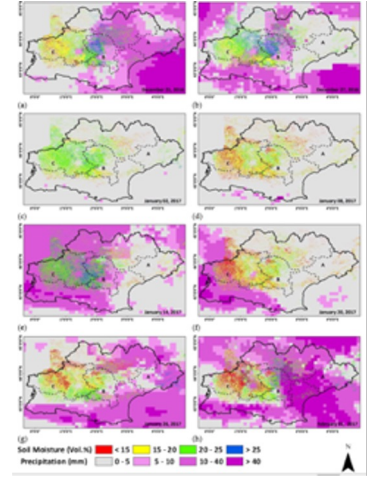
**Sentinel 2** : for contribution of vegetation to the total radar signal

**Theia Land Cover map** : to extract agricultural parcels

**Scale** : sub-parcel scale

**Production by Theia** : 6 days update over some regions

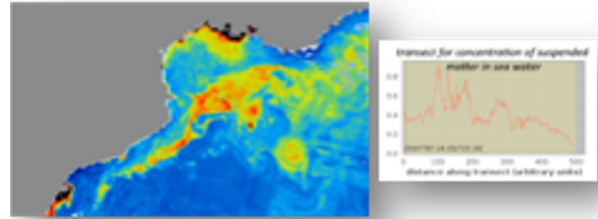
**Facilitation** : N. Baghdadi (INRAE, TETIS) et M. Zribi (CNRS, Cesium)



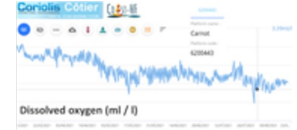
# CONTINENTAL WATER COLOR

Distribution of the **first water quality products** (SPM suspended particulate matter) in 2018 through THEIA

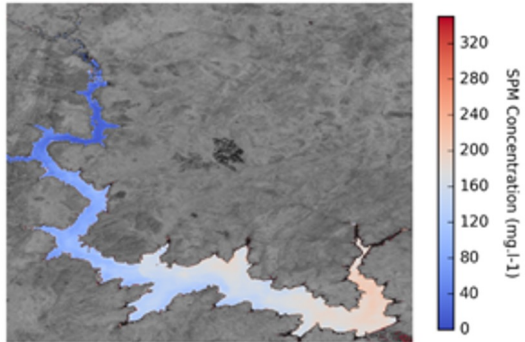
We expect to cover the **30 foremost watershed** (in terms of discharge) and **20 lakes** in the world



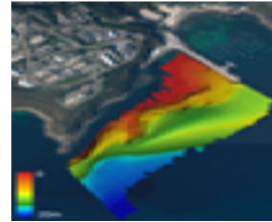
WATER QUALITY



WATER OXYGEN



BIOMASS PRODUCTION



BATHYMETRY LASER and LIDAR



COASTLINE MONITORING

# 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...



# ÉRUPTION VOLCAN TONGA

**14 janv. 2022** l'éruption volcanique sous-marine a entraîné des vagues qui ont frappé l'Australie, la Nouvelle-Zélande, le Japon ainsi que les côtes ouest de l'Amérique du Nord et du Sud durant plusieurs jours.

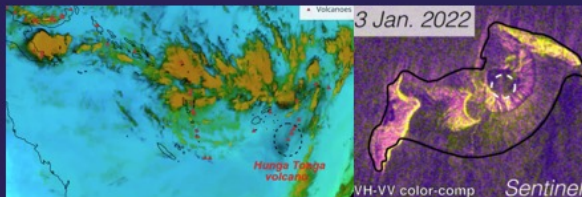
Toutes les informations recueillies par les scientifiques **facilitent ainsi les interventions d'urgence** et aident à **mieux comprendre et anticiper ce type de catastrophes naturelles à l'avenir.**



*Les îles Tonga sont composées d'environ 170 îles*

AERIS

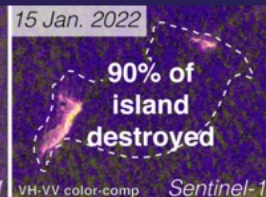
Atmosphère



Vues par les satellites géostationnaires

FORM@TER

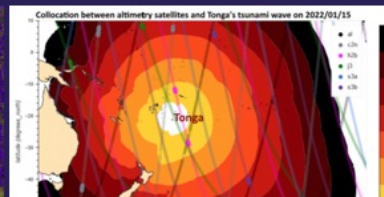
Terre solide



Évolution de la morphologie de l'île volcanique Hunga Tonga vue par imagerie satellite

ODATIS

Océans



Mesures de la hauteur du tsunami au large, par les satellites altimètres et via les marégraphes côtiers

THEIA

Surfaces continentales



Cartographier les dommages pour guider les secours

DINAMIS

Images satellites HD

# VAGUE DE CHALEUR EN EUROPE

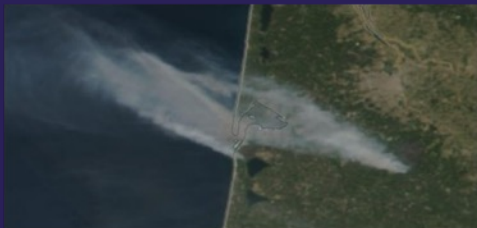
**Été 2022** Les scientifiques travaillent à **comprendre les vagues de chaleur en Europe, leur fréquence ainsi que leur intensité** afin de permettre aux sociétés civiles de mieux se préserver et de réussir à inverser la tendance.

La chaleur extrême persistante amène également à mieux appréhender la notion de climat au sein de la société civile.



AERIS

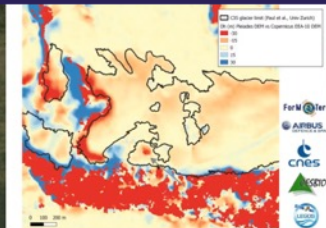
Atmosphère



Feux de forêt et vague de chaleur – un été particulièrement intense

FORM@TER

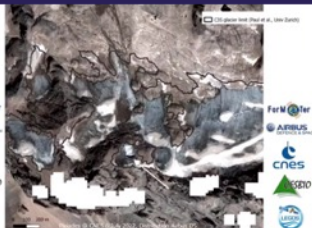
Terre solide



Estimation du volume de glace détaché lors de l'effondrement du glacier de la Marmolada

THEIA

Surfaces continentales



Estimation du volume de glace détaché lors de l'effondrement du glacier au Kurdistan

ODATIS

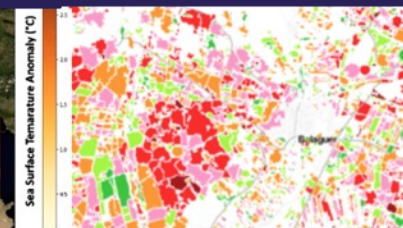
Océans



Vague de chaleur juin 2022 en mer Méditerranée

DINAMIS

Images satellites HD



L'ampleur de la sécheresse rendue visible par les satellites

# European positioning



# European projects

## EARTH SCIENCE

### MARINE

EMODNet, SEADATANET,  
BLUE CLOUD,  
MARINE COPERNICUS

### SURFACES CONTINENTALES

Copernicus Land, Climat C3S

### ATMOSPHERE

Copernicus CAMS

...

## TRANSVERSAL SERVICES

### DATA GOVERNANCE

ENVRI-FAIR, FAIR EASE, EOSC  
Pillar, FAIR Impact

### HIGH PERFORMANCE CALCULATION

PHIDIAS, EURO-HPC

### OPERATIONAL SERVICES

Copernicus Service C3S, CAMS, Land,  
Marine



Figure 5. Work package structure and work flow of EOSC-FAIR.



# Participations actives à des projets nationaux, européens et internationaux

## ENVRI – FAIR : FAIRisation multi-domaines

Implication au travers des IR/TGIR ESFRI : ACTRIS, IAGOS, EPSO, EURO-ARGO, ANAEE ; 37 part., 19M€ - 16% Fr (CNRS, Ifremer, BRGM, INRA)

**CEF – PHIDIAS** (CEF – DG Connect, 3 M€) : service à la demande « données spatiales et environnementales » et HPC : **CINES** CNRS, CSC, IFREMER, IRD, MARIS, Univ. Louvain, ...

**EOSC-Pillar** : gouvernance nationale et Eur. EOSC (INFRA-EOSC, 12 M€), coordination WP6 : 10 études de cas (FAIR Data and Services)

Seadatanet - SeadataCloud, projet d'appui à l'Infra (10 M€)

Blue Cloud : EOSC ocean (10 M€) Ifremer/Odatis, CNRS, Sorbonne Univ. Mercator-Ocean, ...

FAIR EASE (INFRA-EOSC 01-04) : 5 M€, 25 partenaires, coordination Data Terra

FAIR IMPACT (INFRA-EOSC 01-05)

Copernicus : services C3S, CAMS, Land, Marine, ...

**EURO-HPC** (CEA, GENCI, ...) : 2 cas d'usages (Hydrologie, Urbain)

**CLUSSTER** (ATOS, GENCI, CNRS, CS-Group, OVH,...) : partenaire associé

ANR Flash (open science) : 4 projets : FAIRisation, certification

GO-FAIR IN : GAIA Data In (global integrated Data Impl. Net.)

GEO : programme Board, Working Group, initiatives, ...

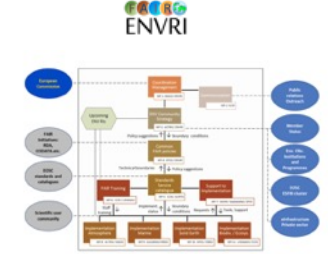


Figure 1. Work package structure and work flow of ENVRI FAIR.



# DATA TERRA au sein d'EOSC

=> constituer le noeud EOSC français dans les thématiques du système Terre, environnement et biodiversité  
=> Ensemble de services intégrés au portail EOSC

## Impliqué dans de nombreux projets EOSC via le CNRS-INSU

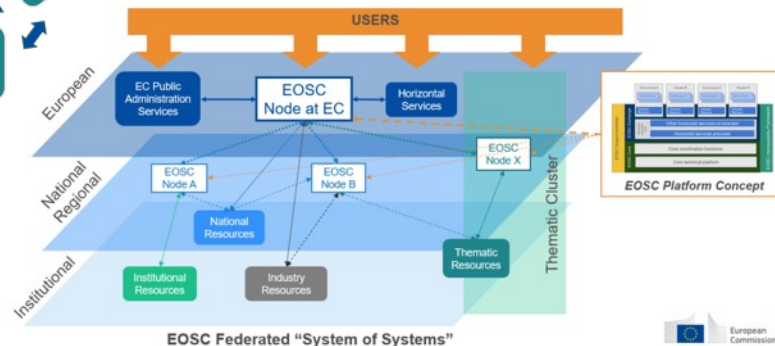
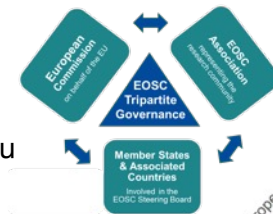
**FAIR EASE** (début le 01/09/2022, 5M€, 26 partenaires, coord : Data Terra/CNRS A. Rizzo, IRD). Objectif : développer et exploiter des services distribués et intégrés d'observation et de modélisation du système Terre, de l'environnement et de la biodiversité, avec les communautés d'utilisateurs et les infrastructures de recherche ESFRI dans un cadre EOSC. 143 PM/1,2M€

**FAIR IMPACT** (début le 01/06/2022, 10M€, Coord.: DANS). But : réaliser un FAIR EOSC « EOSC de données et de services FAIR », en soutenant la mise en œuvre des pratiques FAIR dans les communautés scientifiques et les résultats de la recherche aux niveaux européen, nationaux et internationaux. 14PM

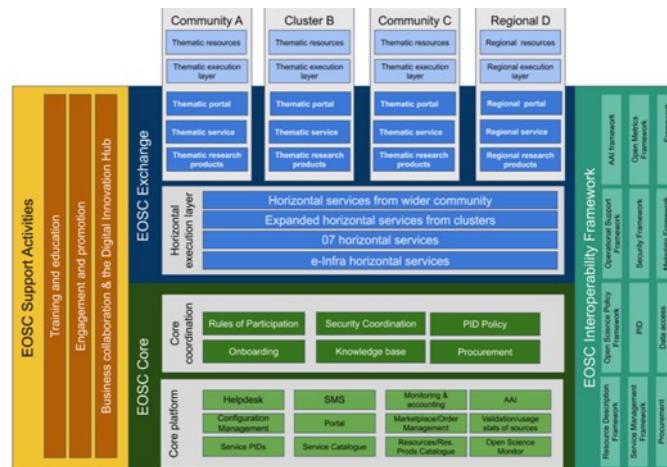
**GEO INQUIRE** (début le 01/10/2022, 15M€, Coord. : GFZ). Accès aux observations, produits de données et services permettant la surveillance et la simulation des processus dynamiques dans la géosphère à des niveaux de détail et de précision sans précédents. 8PM

**DT GEO** (début le 01/09/2022, 10M€). Déployer un prototype de jumeau numérique sur les extrêmes géophysiques, composé de « Digital Twin Components » (DTC) interdépendants traitant des géorisques liés aux tremblements de terre (naturels ou d'origine humaine), aux volcans et aux tsunamis déclenchés par des tremblements de terre ou des glissements de terrain. 4PM  
BioDT, InterTwin

**ENVRI-HUB Next** (début le 01/10/2023, 5M€, Coord. : EGI). Construction du hub de données/méta données multidomains (Data Terra porteur des offres de services aux données des IRs thématiques (Océan, Atmosphère, Surfaces continentales, ..), WP coordonné par EPOS Eur (CNR)



EOSC infrastructure node at the European level





# Projet FAIR-EASE

<https://fairease.eu/>



**Objectif:** « The first interdomain digital architecture for integrated use of environmental data »

**WP2** FAIR-EASE Interdisciplinary **Data Discovery and Access Service**

**WP3&4** FAIR-EASE Earth **Analytical Lab and Data Lake**

**WP5**

**Thematic use-cases:**

- Coastal Water Dynamics
- Earth Critical Zone
- **Volcano Space Observatory**
- Ocean Bio-geochemical Observation
- Marine Omics Observation

Projet porté par **Data Terra**  
(PI: **Alessandro RIZZO**)



**DATA TERRA**

5 PM ForM@Ter  
20 PM AERIS  
2 ans CDD

**Démarrage :** Septembre 2022  
**Durée :** 4 ans

## Travaux en cours:

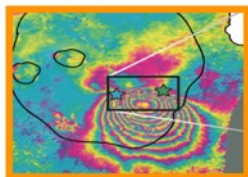
- à l'échelle du projet:
  - ▶ remontée des besoins des use-cases vers les WP techniques.
  - ▶ définition des données utilisées et des stratégies techniques des experts / thématiciens
  - ▶ définition du concept de « Data Lake » dans le contexte du projet (le projet ne finance pas l'infra)
- à l'échelle du use-case « Volcans »:
  - ▶ AERIS: consolidation du portail « VolcPlume » pour l'exploration/analyse des observations des panaches volcaniques (sol/satellite)
  - ▶ ForM@Ter: définition de la stratégie de traitement InSAR (problématique de l'accès rapide aux données Sentinel-1 via PEPS ; investigation en cours via le DataLabs du CNES).

## Volcano Space Observatory

### Volcanic deformation

- modes of transport and storage of magma
- mass balance

### Sentinel-1 – InSAR



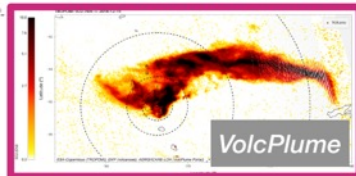
10 km

pixel size = 1–100 m

### Volcanic emissions

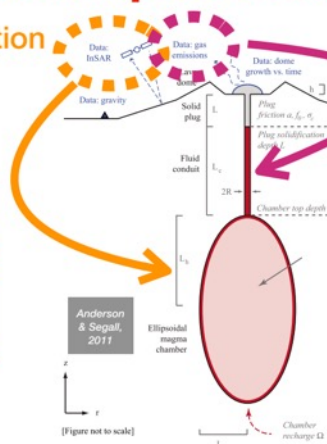
- gas flux
- mass balance
- composition, depth
- atmospheric transport and pollution

### Sentinel-5P TROPOMI – SO2



1000 km

pixel size = 5–10 km



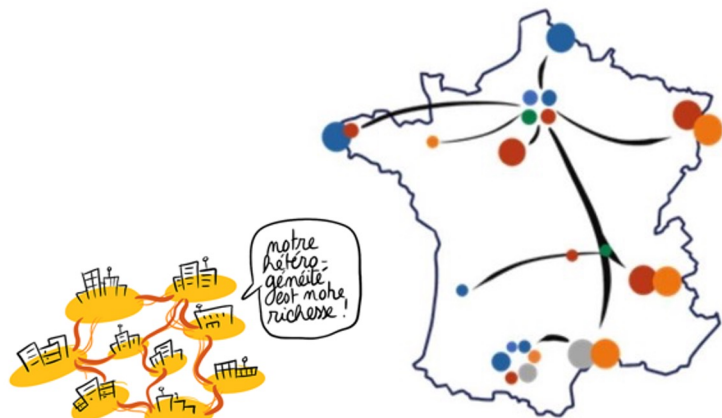
[Figure not to scale]

PI:  
**Marie BOICHU**  
(LOA)



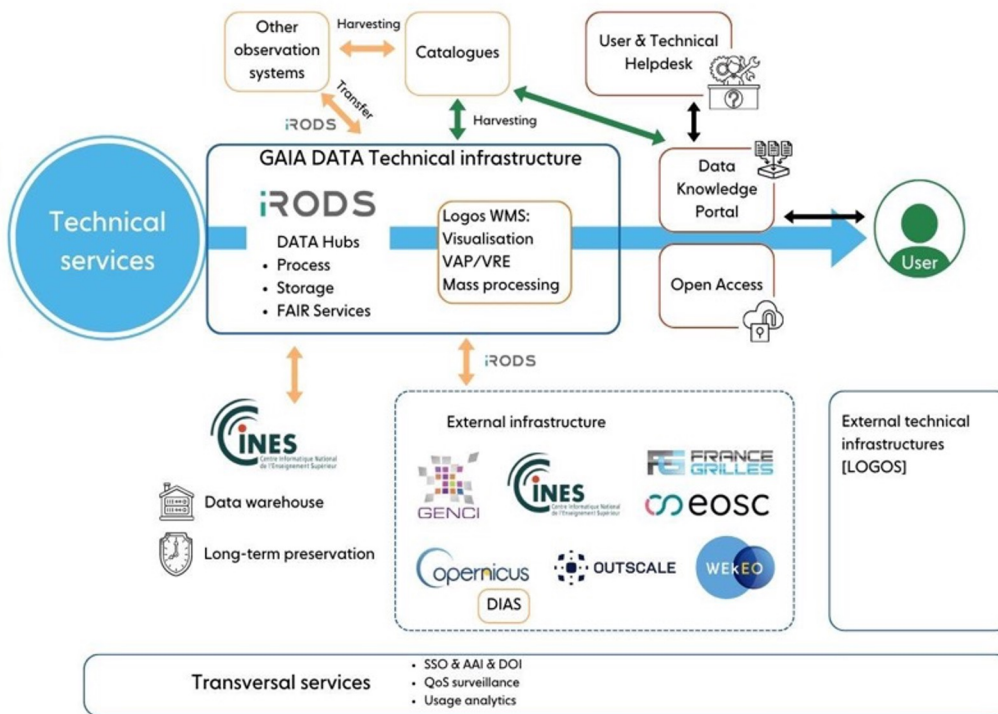
Université de Lille





8 CENTRES de CALCUL déjà SPECIALISÉS

8 sites principaux  
30 sites existants





# INFRASTRUCTURE GAIA DATA

Grille de données et de services : 8 principaux centres en réseau



En relation avec des projets connexes

## Projets Equipex+ ou PIA4 infra

- FITS
- MesoNet
- Clusster

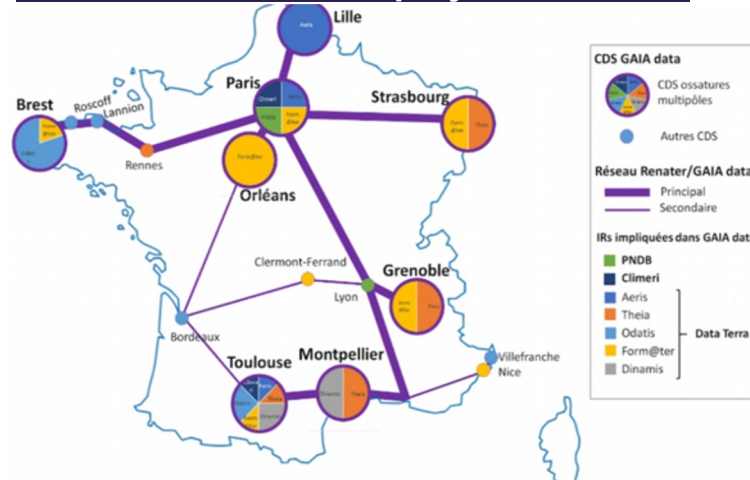
## Projets Equipex+ ou PEPR thématiques

- Obs4Clim
- TerraForma
- Marmor
- OneWater

## Projets H2020 – Horizon Europe

- IS-ENES
- PHIDIAS
- EOSC-Pillar
- FAIR EASE
- FAIR IMPACT

## Projets CPER en région



Intégré dans le paysage international / Européen



- Mise en place d'un réseau dédié haut-débit et sécurisé
- Déploiement d'une grille de données (système iRODS AC) / S3 pour permettre un accès distant aux données et le transfert rapide et automatique de grands ensembles de données d'un centre vers un autre
- Interopérabilité des traitements entre les 8 centres de Gaia Data, avec les centres HPC en France et avec les clouds commerciaux (GAIA-X - DIAS)





## 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 **FAIRisation**



## 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



# Portails GAIA data

Accès unique aux  
données

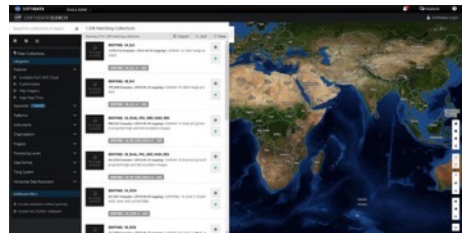
Multiples domaines

Services à la demande

vocabulaires et ontologies  
pour sélectionner les jeux de  
données

Portails thématiques,  
pôles, domaines ou  
projets

Visualisation (WebMap)  
Accéder / explorer l'ensemble  
des données disponibles



« WebMap utilisateur »  
Afficher un ensemble de jeux de  
données défini et configuré par  
l'utilisateur

Entrée par un domaine (ex. : Ocean) :  
Pointeur vers page thème « Océan »  
=> Informations sur les missions  
spatiales et données associées

- WebMap ne présentant que les données relatives à l'Océan
- Ensemble de services liés à l'Océan



**DATA**TERRA



[contact@data-terra.org](mailto:contact@data-terra.org)  
+33 (0)4 67 54 87 08

[www.data-terra.org](http://www.data-terra.org)