



DATA  
TERRA



ODATIS

# Data Terra et le pôle ODATIS Une Infrastructure de Recherche pour les données marines

Sabine Schmidt<sup>1</sup>, Gilbert Maudire<sup>2</sup>, Cécile Nys<sup>2</sup>, Joël Sudre<sup>3</sup>, Valérie Harscoat<sup>2</sup>,  
Caroline Mercier<sup>4</sup>, Gérald Dibarboure<sup>5</sup>, Frédéric Hyunh<sup>3</sup>

<sup>1</sup> CNRS, OASU, UMR5805 EPOC, Université de Bordeaux, Pessac, France

<sup>2</sup> Ifremer, Centre de Bretagne, Plouzané, France

<sup>3</sup> UMS CPST, Coordination Pôles de données et de services pour le Système Terre, Montpellier, France

<sup>4</sup> Akka, Toulouse, France.

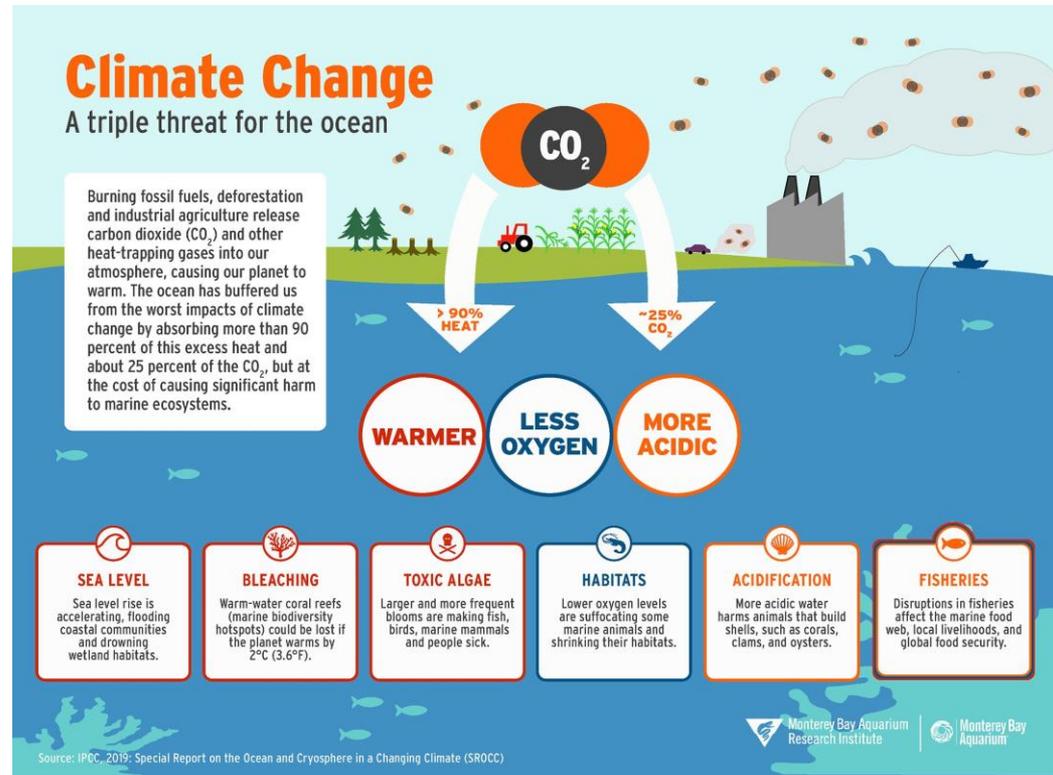
<sup>5</sup> CNES, Toulouse, France.



# L'anthropocène

Depuis la révolution industrielle, l'empreinte des activités humaines sur l'environnement mondial s'est accrue.

Les conséquences actuelles et attendues du changement global sur l'océan sont multiples.



# Un besoin crucial de mieux comprendre pour prévoir les impacts du changement global

Des observations sont nécessaires à tous les stades du processus scientifique :  
description, compréhension, modélisation et prévision

Considérant que:

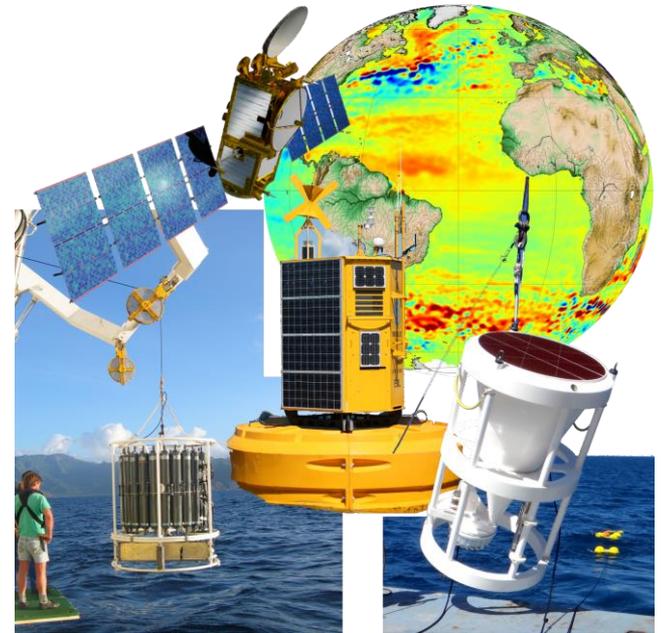
- l'acquisition des données marines est difficile et onéreuse :  
Nécessité d'accéder à des sites distants et d'utiliser de nombreux moyens techniques (flotte océanographique, sites instrumentés, gliders, bouées Argo, ...)
- sans archivage, > 30% des données sont perdues ou inutilisables 10 ans après leur acquisition (source: Ifremer).

➡ La préservation des données marines est un enjeu majeur,

# Les défis de l'accès aux données marines

Augmentation du nombre des observations marines (*in situ* et par télédétection) au cours des dernières décennies.

Pour tirer le meilleur parti de ce flux de données au profit de la connaissance et de la société, les centres de données doivent respecter des principes communs.

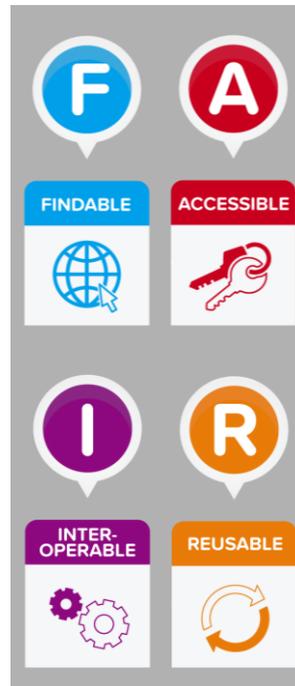


# Les défis de l'accès aux données marines

## Défi 1 : la qualité des données (principes FAIR)

Les métadonnées et les données doivent être faciles à trouver et à (ré)utiliser (décrire vos données, appliquer des identificateurs persistants)

Les données doivent être intégrables à d'autres données (format ouvert, vocabulaire cohérent, normes de métadonnées)



Envisager ce qui sera partagé et comment on peut y accéder

Les données doivent être réutilisables, avec des métadonnées bien décrites et une licence appropriée

# Les défis de l'accès aux données marines



## Défi 2: certification des dépôts de données

La Research Data Alliance (RDA) fournit un cadre commun pour la mise en œuvre et la maintenance des dépôts numériques.

the CoreTrustSeal requirements

La certification est importante pour garantir: :

- ✓ la fiabilité et la durabilité des dépôt de données,
- ✓ l'archivage et le partage à long terme des données pour les utilisateurs et les financeurs.



[www.coretrustseal.org](http://www.coretrustseal.org)

# Les défis de l'accès aux données marines

## Défi 3: Amener les producteurs à partager leur données

L'ICSU (International Council for Science) promouvait déjà en 2011

*« un accès complet et ouvert aux données scientifiques, en particulier lorsque la recherche est financée par des fonds publics ».*

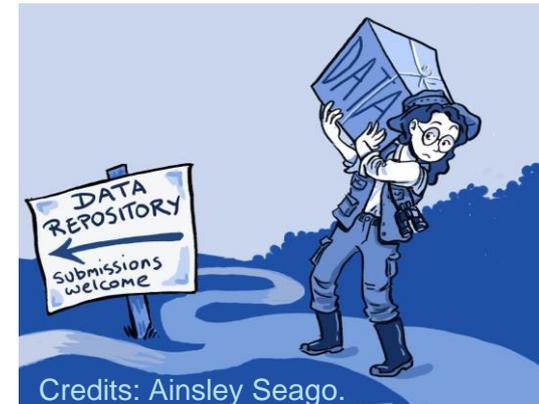
Les chercheurs peuvent être réticents à partager leurs données **en raison des coûts réels et/ou perçus**

Sentiment de - perte de contrôle sur les données,  
- contraintes sans bénéfice,

Manque de - ressources informatiques/humaines  
- formation.

**Les dépôts de données doivent rendre le partage de données**

**plus faciles pour le producteur  
pas seulement pour l'utilisateur.**

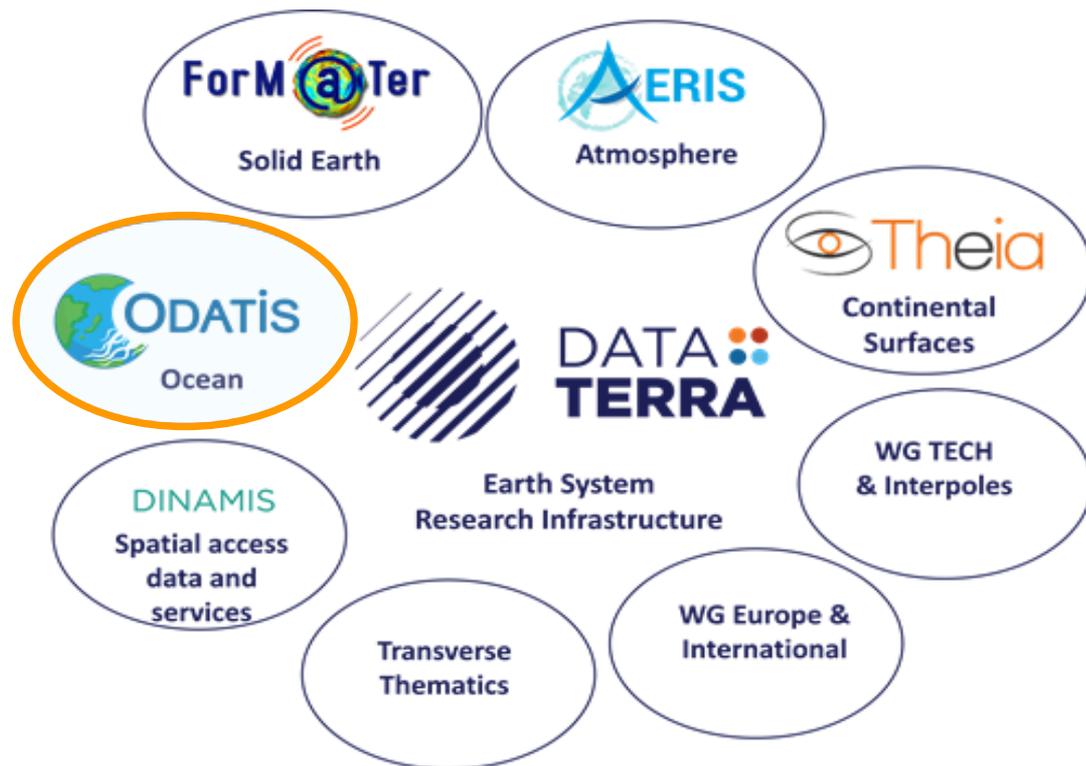


Credits: Ainsley Seago.



# L'IR Data Terra

Infrastructure de recherche « *pôle de données et services pour le Système Terre* » :  
*de la donnée à la connaissance du Système Terre*



Composée de:

- 4 pôles thématiques
- groupes de travail
- services transverses

# ODATIS: Ocean DATA Information and Services

[www.odatis-ocean.fr](http://www.odatis-ocean.fr)

**Mission** Préserver, gérer et diffuser les données et produits d'observations des océans

## Partenaires



Demande en cours  
d'instruction



## Source des Données

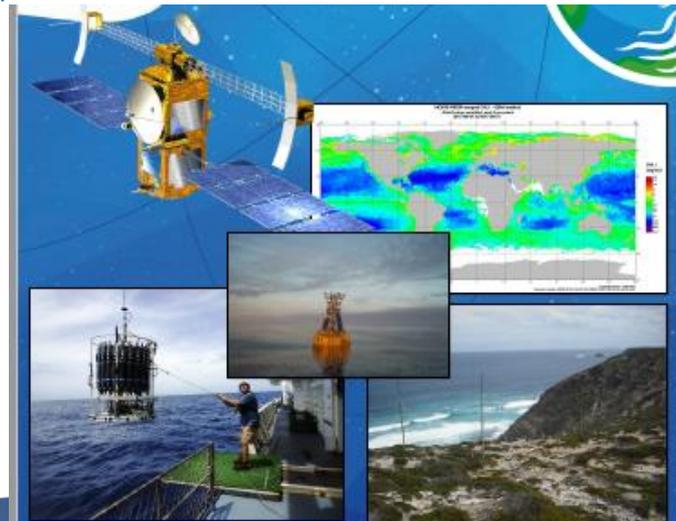
missions satellites «marines»

*altimètre, radiomètre, diffusiomètre, optique, ...*

observations *in situ*

SNO/SO côtiers (IR ILICO, régionaux) et hauturiers (Ohis, Argo, SSS ...)

TGIR Flotte Océanographique, LEFE-CYBER, ...



Merigéo – webconférence – Novembre 2020



# Feuille de route d'ODATIS

- Offrir une vision globale des observations (*in situ*; satellite) et de leurs produits;
- Faciliter l'accès, via un portail unique, aux données, entièrement décrites et qualifiées en accord avec les normes en vigueur (FAIR, TRUST) ;
  - ce qui implique d'assurer
    - la conservation à long terme des jeux de données;
    - l'interopérabilité des jeux de données (espace; temps; disciplines);
- Promouvoir des utilisations combinées de données de nature (*in situ*; satellite) ou d'origine différentes (réseaux opérationnels/expériences scientifiques);
- Offrir la possibilité d'explorer, d'extraire et d'analyser des données en proposant des outils et des ressources informatiques.

# Organisation d'ODATIS

## STRATEGIC LEVEL

### Inter-institution Steering Committee



### Data Terra Executive Board

Defines a common strategy for all the thematic clusters to provide an unified access to data (interoperability, vocabulary), products, softwares, tools and services.

### Scientific Council

Experts on marine and coastal sciences.  
Represent the user scientific community, express the needs and recommend scientific consortia.

## EXECUTIVE BOARD

Management team



Representative of each DSC

Director  
Technical director  
Scientific director  
Assistants

## Data and Service Centres (DSCs)

Catalog: web & client: CNES  
server & catalog: Ifremer

Data:  
Satellite (CDS-SAT)  
CERSAT, AVISO  
In situ (CDS-in situ)  
Coriolis, SHOM, SISMER  
OMP, OASU, IMEV. IMEV, SBR

## Technical Workshop

Proposes technical guidelines (interoperability, services, ...), Organizes practical sessions to test technics and exchange experiences among CDSs.

## Scientific Consortia

Propose and instruct new thematic data compilation (ex: EOVS) or new product on the request of the scientific community.

## OPERATIONNAL LEVEL



# Les Centres de Données et Services

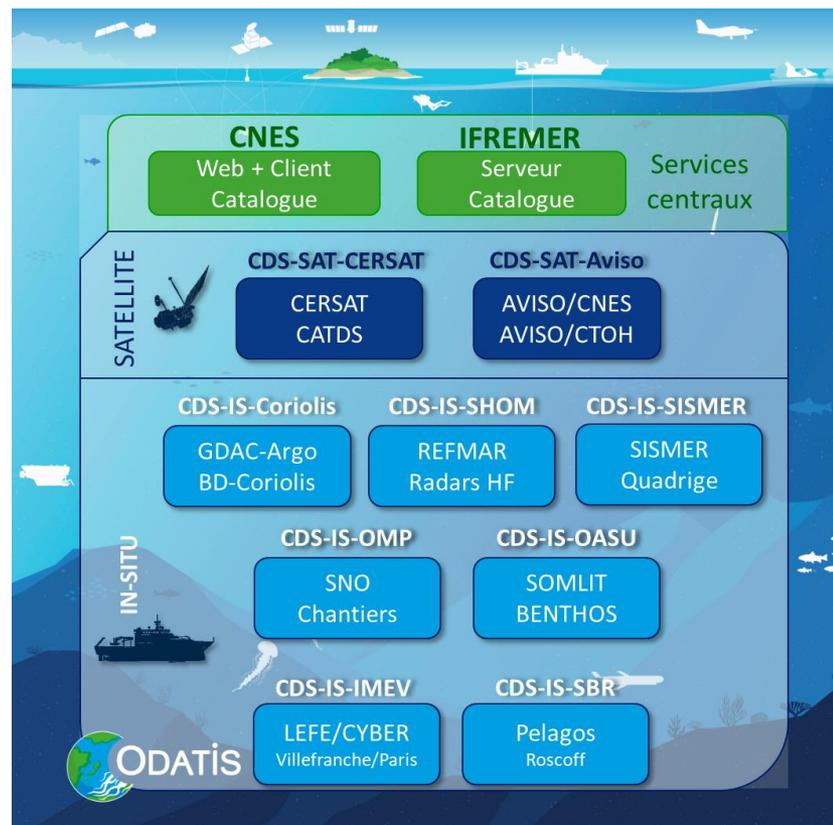
En charge de:

- gestion des données:  
bancairisation, pérennisation, diffusion, ...
- élaboration de produits au quotidien au plus proche des utilisateurs;

Soutenus par les tutelles du pôle

Dans le cadre d'un mandat (type de données)  
et d'un cahier des charges

Mis en relation par des modules communs :  
Portail web et Catalogue des données



# Le portail ODATIS

www.odatis-ocean.fr



LANGUAGES | CONTACT |

QUI SOMMES-NOUS? ▾ DONNÉES ET SERVICES ▾ ACTIVITÉS ▾ ACTUALITÉS ▾ RESSOURCES ▾

Accès au catalogue

**Rechercher**

Rechercher... CATALOGUE CARTE **1** 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Résultats 1 à 20 sur 128 (20 par page)

**Carte d'habitats physiques des fonds marins en France métropolitaine (échelle 1 / 1 000 000) - Version 2011**

Habitats physiques dans la typologie Eunis résultant du croisement de trois variables physiques relatives à la profondeur de 200 mètres, hauteur du substrat, nature biologique, version de croisement au fond. Cotes de France métropolitaine. Rapport Réseau National Arrière, Réseau National, Réseau National. Source: Agence des Aires Marines Protégées, Réseau.

**Carte d'habitats physiques des fonds marins en France métropolitaine (échelle 1 / 300 000) - Version 2011**

Habitats physiques dans la typologie Eunis résultant du croisement de trois variables physiques relatives à la profondeur de 200 mètres, nature du substrat, nature biologique, version de croisement au fond. Cotes de France métropolitaine. Rapport Réseau National Arrière, Réseau National, Réseau National. Source: Agence des Aires Marines Protégées, Réseau.

**Lieux d'observation et de surveillance du réseau REPHY**

Le REPHY est un réseau national couvrant le littoral métropolitain au sein des Aires Marines Protégées. L'objectif est de fournir l'ensemble des données physico-chimiques et biologiques nécessaires à la connaissance des écosystèmes marins et à la surveillance de leur état de conservation.

**Lieux d'observation et de surveillance du réseau ROCCH**

Depuis 2008, le Réseau d'Observation de la Coastal Ocean (ROCCH) est géré par le Centre National d'Observation de la Mer (CNOM) afin de garantir la qualité des données et de faciliter l'accès à ces données. Le ROCCH agit en complémentarité avec les réseaux nationaux, communautaires et internationaux de surveillance côtière. L'actuel état de ce réseau de...

**Décrire**

Données de Marées du SHOM - Tide Database

Un système dénommé « Tide Database (TDB) » a été mis en place au SHOM pour archiver, traiter et rendre accessible les mesures in situ du niveau de la mer et les métadonnées associées, telles que les références verticales, les repères de marée, les instruments, les photographies, etc.

Accès aux données

Dat(e) : 2010-01-01T00:00:00

Contact(s) : VOINSON Guillaume ( SHOM )

Source : Producteurs de données : SHOM et ses partenaires (LEOS, IGN, DHO, Ports autonomes, SPC)

Généologie : Juillet 2010 : La TDB remplace les anciens outils à la cellule HDC. Janvier 2012 : Mise en service de TDB au SHOM

Les tâches réalisées par les différents intervenants dans la base de données comprennent l'archivage et la saisie des métadonnées, le traitement et la validation, puis le contrôle et l'approbation. Ces travaux sont réalisés par les entités du SHOM impliquées au travers d'un processus métier (workflow), qui assure que seule la donnée validée et approuvée est disponible à l'ensemble des utilisateurs.

Contraintes : Limitation d'utilisation : Modes(s) de tarification : Coût de la mise à disposition

Autres contraintes : Accès restreint aux utilisateurs SHOM. Les partenaires du SHOM peuvent toutefois demander un export de certaines mesures à titre gracieux.

Informations géographiques : Type de données : Vecteur Échelle : 1:50000  
Système de coordonnées : WGS 84 (EPSG:4326)  
coordonnées :

Ressources associées : Est une source pour Météorologie

**Visualiser**

Fond de carte : Sextant

Rechercher un lieu

Filter

RF sea\_surface.height.above.sea.level

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Tempo : 1984-2017

Échelle : 0.291 (= 0.372m)

Effacer

Styles :  habitude

WGS84 4 456 3 703

**Visualiser**

Fond de carte : Sextant

Filter

Activités Numériques

Campagnes océanographiques

Campagnes ADOP de coque

Autre

**LEVENES** **ERREUR** **SOURCES**

Campagnes ADOP de coque

Role des tonnes

WGS84 209 180 51 289

**Télécharger**

Email

Prénom

Nom

Organisation

Utilisation

**Télécharger**

Annuler  Télécharger

DATA TERRA

ODATIS



# Le portail ODATIS

www.odatis-ocean.fr



LANGUAGES | CONTACT |

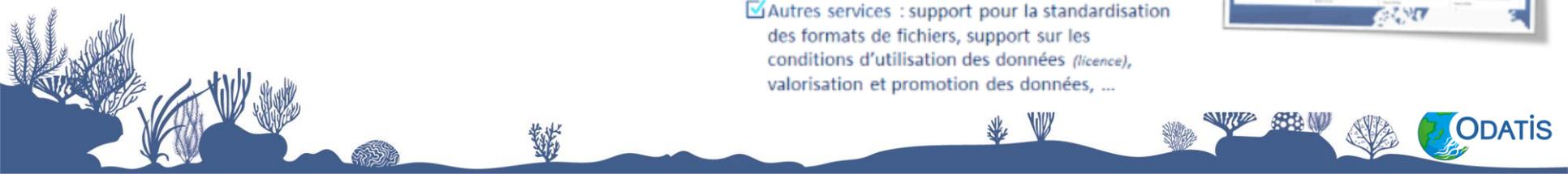
QUI SOMMES-NOUS? ▾ DONNÉES ET SERVICES ▾ ACTIVITÉS ▾ ACTUALITÉS ▾ RESSOURCES ▾

Accueil



Accès aux données >  
Services aux producteurs de données

## Services aux producteurs



# Le portail ODATIS

www.odatis-ocean.fr



LANGUAGES | CONTACT |

QUI SOMMES-NOUS? ▾ DONNÉES ET SERVICES ▾ ACTIVITÉS ▾ ACTUALITÉS ▾ RESSOURCES ▾

Accueil



Accès aux données >  
Services aux producteurs de données

**Services aux producteurs**

spécificité ODATIS → SEANOE

<https://www.seanoe.org/>

dépôt de données sous licences Creative Commons

Merigéo – webconférence – Novembre 2020



# Publish your marine data

SEANOE offers you the possibility of publishing free of charge your scientific data in the field of the marine sciences.

**Each data set published by SEANOE has a DOI.** Your data may well be cited in scientific articles in a reliable and sustainable way.

SEANOE offers a fast responding service : if your dataset is well described<sup>1</sup>, **you should get a DOI within 24 hours.**

Possibilité d'embargo (2 ans max. ) pour restreindre l'accès aux données, par exemple d'une publication en cours d'évaluation, par exemple.

# Les ateliers techniques

réunissent les partenaires autour de présentations et tables rondes pour:

- définir les orientations techniques du pôle;
- réaliser des mises en pratique sur la base d'exercices d'application de logiciels de visualisation de la donnée.

Tous les ateliers sont mis en ligne

Merigéo – webconf



## Atelier Technique #7 : Gestion de données biologiques marines

**ORGANISATION** : Joël SUDRE (LEGOS/UMR5566), Cécile NYS (IFREMER IRSI/SISMER).

**DATE** : mardi 22 septembre (9h-18h) et mercredi 23 septembre (9h-16h30) 2020.

**LIEU** : Visioconférence GoToMeeting / Salle de visioconférence du bâtiment Jules Barrois à l'IMEV (ex-OOV). Suite à votre inscription, les informations de connexion ont été transmises par mail.



Agenda

Compte-rendu

Liste des participants

- Présentation du **Pôle de données Océan** – ODATIS – *Gilbert MAUDIRE*
- Présentations des **initiatives dans la gestion des données biologiques marines** (20min/présentation + 10min de questions) :
  - **PNDB, de la gestion de données à son analyse, présentation des outils et services** – *Yvan LE BRAS*
  - **EMBRC** (France) Données biologiques marines dans le Centre National de Ressources Biologiques Marines - *Gemma GIMENEZ-PAPIOL* - *présentation accessible exclusivement à destination des participants pendant l'atelier.*
  - **SAR/SIMM** : **référentiels utilisés dans le cadre du Système d'Information sur le Milieu Marin** – *Clémence RABEVOLO*
  - **ECOTAXA & Zooplancton** - *Jean-Olivier IRISSON* - *présentation accessible sur authentification; veuillez renseigner vos identifiants dans la marge en haut à droite de cette page pour voir apparaître le titre et le lien de cette présentation à la suite de cette liste à puces.*
  - **SeaDataCloud/EMODNet Bio/PhytOBS** : **Mise en place des chaînes de traitement à Roscoff pour la publication des données des suivis à long terme** – *Mark HOEBEKE*
  - **SNO BENTHOBS** - *Vincent BOUCHET / Nicolas DESROY*
  - **Quadrige** - *Anaud ROUILLY*
  - **APA & SI MORSE** - **Le projet MORSE, vers un nouveau portail de suivi des échantillons biologiques** - *Sylvie VAN ISEGHEM*
  - **Imagerie** - *Catherine BORREMANS*
  - **Cytométrie** - *Gérald GREGORI, Felipe ARTIGAS, Maurice LIBES, Marc Sourisseau, Melilotus Thyssen*
    - Qu'est-ce que la cytométrie en flux (diapos 1-10),
    - Cytométrie en flux automatisée et accessibilité des données pico-nano-microplancton (diapos 11-30),
    - Avancement chaîne de traitement de cytométrie en flux (diapos 31-45).
- Ateliers de discussions :
  - Atelier de discussion « **Référentiels** »
  - Atelier de discussion « **Format de données** »
  - Atelier de discussion « **Outils** »
  - Atelier de discussion « **Imagerie** »
- Autres points & Préparation du prochain Atelier Technique

# Les Consortiums d'Expertises Scientifiques (CES)

Objectif: promouvoir ou/et valoriser des méthodes de traitement et des produits innovants d'observation spatiale, aéroportée ou in-situ de l'océan et de ses interfaces (atmosphère, littoral et sous-sol sous-marin) avec les autres pôles de données,

Regroupent les experts nationaux du domaine (inter-organismes, inter-laboratoires)

Dans le cadre d'une démarche projet : calendrier, mode de financement, livrables (durée 2 ans)

## Plusieurs catégories:

CES de prospection (financement des missions)

CES techniques:

- définition de produits, d'algorithmes
- élaboration de prototypes de traitement (financement de la réalisation du prototype)
- implémentation (ex d'un collègue qui aurait développé un algorithme)

# CES Cytométrie en flux

## CES Cytométrie en flux

Le Consortium d'Expertise Scientifique (CES) Cytométrie en flux a été mis en place début 2020 afin de fédérer les acteurs scientifiques au niveau national, voire international, autour de la cytométrie en flux (CMF), technologie très largement utilisée de nos jours pour caractériser la structure des assemblages de **micro-organismes planctoniques**, aussi bien en milieu marin qu'en eau douce. La CMF permet de **classer ces micro-organismes** en différents groupes fonctionnels, définis chacun par des cellules aux propriétés optiques similaires, et d'en **déterminer les abondances**. En règle générale, ces groupes cytométriques appartiennent au pico-, nano- ou microphytoplancton, ainsi qu'aux virus, procaryotes et nanoflagellés hétérotrophes. Certains instruments sont également capables d'associer une prise de photo à l'analyse individuelle des cellules, offrant la possibilité d'une identification taxonomique supplémentaire.

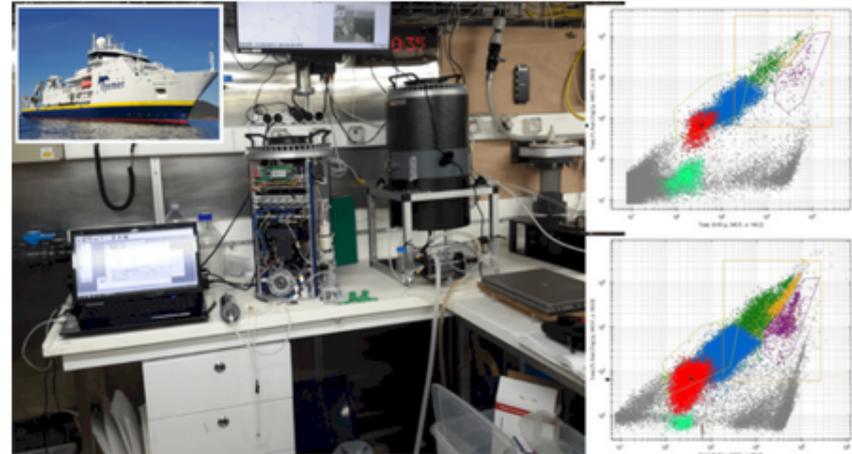
**Contexte** : Le **phytoplancton** inféodé à la couche euphotique de l'Océan est responsable par la photosynthèse de la moitié de la production primaire annuelle globale. Le **bactérioplancton hétérotrophe** présent dans toute la colonne d'eau et la partie superficielle des sédiments est le principal responsable de la minéralisation de la matière organique et de la régénération des sels nutritifs indispensables au développement du phytoplancton en surface. L'ensemble de ces micro-organismes affecte les cycles du carbone, de l'azote, du phosphore et d'autres éléments clés. Indispensables au maintien de l'écosystème, ils sont toutefois très sensibles au changement de leur environnement, aussi bien dans l'océan côtier qu'en océan ouvert. Leur étude est donc fondamentale dans le contexte du changement global et de l'impact des activités anthropiques.

**Objectif** : L'objectif de ce CES, piloté par Gérald Grégori (MIO, CNRS), est de générer une initiative communautaire pour bancariser, partager et diffuser à l'échelle nationale et internationale les données de cytométrie en flux environnementales. Afin que ces données jouent leur rôle de levier pour la science et l'innovation, il est indispensable qu'elles soient facilement accessibles et réutilisables selon les principes FAIR. Il est également indispensable que ces données soient qualifiées et pour cela répondent aux protocoles d'analyse standardisés reconnus par la communauté.

Ce CES comporte trois groupes de travail (WP) :

- **WP Vocabulaire** : animé par Melilotus Thyssen (MIO, CNRS) et David Pecqueur (OOB, CNRS),
- **WP Chaîne de traitement des données de cytométrie** : animé par Maurice Libes (O.S.U Pytheas, CNRS) et Marc Sourisseau (IFREMER),
- **WP Bonnes pratiques** : animé par Felipe Artigas (LOG, ULCO), Christophe Lambert (LEMAR, CNRS)

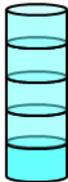
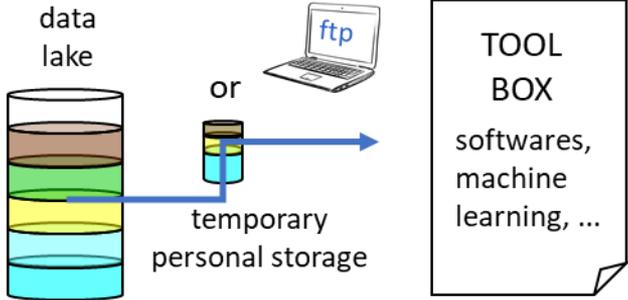
Le CES aura aussi en charge d'identifier et promouvoir ces protocoles en réalisant une veille technologique grâce à des échanges d'informations avec la communauté. Il pourra également identifier les verrous ainsi que les outils susceptibles d'améliorer la procédure de génération ou de visualisation de ces données, et leur archivage sur des bases de données interopérables.



*Cytomètres en flux en cours d'installation à bord du R/V « Pourquoi Pas ? » lors de la campagne océanographique Peacetime, en 2017 en mer Méditerranée et cytogrammes (à droite). Crédits M. Thyssen (MIO).*

# Le besoin de développer une hiérarchisation des centres de données

La mise en œuvre d'une telle e-infrastructure nécessite d'optimiser les ressources humaines et informatiques, sur la base de trois types de centres de données:

SERVICE	data repository	production	on demand
Involved structures			 
USER	data repositories doi, licences reporting on data use	combination of different marine dataset ( <i>in situ</i> / satellite) from the same thematic or area.	data analyses and interpretation cross analyses of different data from all Earth compartments
BACK OFFICE	<p>Data Assembling Centres:</p> <ul style="list-style-type: none"> <li> Close to the producer</li> <li> Common catalogue and vocabulary servers</li> <li> Long-term archive</li> </ul>	<p>Data &amp; Service Centres:</p> <ul style="list-style-type: none"> <li> National data hub</li> <li>Aggregates large collections at the national minimum level</li> </ul>	<p>Virtual Research Environments (VRE)</p>  <p>data lake</p> <p>or</p> <p>temporary personal storage</p> <p>TOOL BOX</p> <p>softwares, machine learning, ...</p>

Dépôt du projet Equipex PIA3 Gaia Data porté par l'IR Data Terra



**DATA**  
**TERRA**



**ODATIS**

**Merci pour votre attention**

