
DMP Projet EMSO Ligure Ouest

Plan de gestion de données créé à l'aide de DMP OPIDoR

Créateurs du PGD : Maurice Libes, Dominique Lefevre

Affiliation du créateur principal : CNRS

Modèle du PGD : INRA - Trame générique

Dernière modification du PGD : 09/07/2021

Résumé du projet :

EMSO (European Multidisciplinary SeaFloor Observatory) est un réseau européen d'observatoires sous-marins pour l'environnement. EMSO vise à acquérir des séries temporelles longues (> 12 ans) dans les mers du pourtour européen avec pour un des objectifs principaux : l'étude de l'impact du réchauffement climatique sur les océans entourant l'Europe, l'étude des écosystèmes marins profonds dans une optique de recherche fondamentale mais aussi de gestion durable, en s'intéressant particulièrement aux facteurs anthropogéniques et climatiques, l'étude des processus tectoniques, volcaniques, hydrothermaux et gravitaires et la surveillance des risques naturels associés (séismes, tsunamis, instabilité des pentes) pour les zones côtières à forte densité de population. L'infrastructure de recherche EMSO-France regroupe plusieurs sites d'activité dont les sites ANTARES et MEUST au large de Toulon. Cette infrastructure allie l'ensemble des techniques d'observation du milieu marin ainsi qu'un télescope de nouvelle génération pour l'astronomie des neutrinos. Cet observatoire interdisciplinaire MEUST est proposé conjointement par l'INSU et l'IN2P3, l'Université de la Méditerranée et l'Université du Sud Toulon-Var, en concertation avec l'IFREMER dans le cadre du GIS Oceanomed, du CETSM et des projets européens KM3NeT et l'ERIC EMSO. Ce projet est en phase avec le développement de la Technopôle de la Mer de TPM. Cette action est complémentaire des actions menées dans le projet MISTRALS et le service d'observation et SOERE MOOSE. Cette infrastructure de recherche présente une opportunité unique pour développer de nombreux projets multidisciplinaires en constituant une radiale d'observatoires temps réel côte-large. Elle permet d'avoir les données en temps quasi réel à 2 471 m au large de Toulon et de collecter les données de la ligne de mouillage ALBATROSS situé à 6.64 km de distance grâce à une communication acoustique entre les deux structures. Ce travail a été réalisé en étroite collaboration entre EMSO-France et KM3NeT. Les données ainsi collectées vont permettre un suivi fréquence des propriétés hydrologiques de la colonne d'eau en relation avec la dynamique masses d'eau en Méditerranée Nord Occidentale (Lien avec les sites instrumentés du golfe du Lion et Dyfamed, et le réseau MOOSE). Cela ouvre une nouvelle ère pour l'observation in situ en attendant l'extension de cet observatoire fin 2018 avec la mise en place d'imagerie in situ temps réel, d'un crawler sous marin pour le couplage pélagique/benthique entre autres et des capteurs de nouvelles génération dédiés à la biodiversité, la matière organique dissoute, ...

Chercheur Principal : Dominique Lefevre

Identifiant ORCID : Lefevre

Contact pour les Données : Maurice Libes

Produits de recherche :

1. Ligure-ouest MII : EMSO-Ligure Ouest observatory data (MII) from 2019-08 (Jeu de données)
2. Ligure-ouest ALBATROS : EMSO-Ligure Ouest observatory data (mooring ALBATROSS) from 2019-08 (Jeu de données)

DMP Projet EMSO Ligure Ouest

Informations sur le plan de gestion

Auteur du plan de gestion s'il est différent du Principal Investigator/Researcher

Maurice Libes : IR CNRS OSU Pytheas

- maurice.libes@osupytheas.fr

Affiliation de l'auteur du PGD

OSU Pytheas UMS3470 CNRS

Date de création du PGD

Janvier 2018

Version en cours

v 1.1 Oct 2018

v 1.2 June 2021

Date de la dernière version

5/07/2021

Informations sur le projet

Identifiant de l'appel à projet (call for proposal)

EMSO-Ligure-Ouest-01-2018

Financeur(s) du projet

- CNRS INSU
- IN2P3,
- FEDER – CPER NuMerEnv.

Nom du programme de recherche

EMSO-ERIC (European Multidisciplinary Subsea Observatory) :
réseau européen d'observatoires sous-marins pour l'environnement.

Référence de la convention de financement

EMSO-France

Acronyme du projet

EMSO-Ligure Ouest

Nom du projet de recherche

European Multidisciplinary Seafloor Observatory and water column

Institution leader du projet, coordinateur bénéficiaire (nom, pays)

- CNRS INSU

Autres partenaires (nom, pays, rôle de chacun des partenaires en dehors de l'institution leader)

- IN2P3 – pour l'Infrastructure câblée.
- Université Aix-Marseille (AMU) et l'Université du Sud Toulon-Var : ETP

Unité de rattachement du responsable du projet

Institut Méditerranéen d'Océanographie

- MIO UMR 7294 CNRS

Dates et durée du projet

- 2019 en cours : acquisition continue quotidienne, hormis périodes de pannes matérielles

Présentation succincte des données du projet

Ligure-ouest MII : EMSO-Ligure Ouest observatory data (MII) from 2019-08

Brève présentation des données générées, collectées ou réutilisées :

- **périmètre, type, échelle,**
- **origine : collecte de nouvelles données ou réutilisation de données (conversion / transformation de données ; partage / échange de données ; achat de données),**
- **publications associées**
- **périmètre** : données de capteurs acquises à 2469m de profondeur au large de Toulon
- **type** : Observatoire marin dédié aux sciences environnementales. Collecte des *paramètres physiques, hydrologiques, biogéochimiques*. Données textuelles publiques.
- **origine** : collecte de nouvelles données par 2 dispositifs:
 - un module d'interface instrumenté (MII) et par
 - une ligne de mouillage instrumentée (Albatross : Autonomous Line with a Broad Acoustic Transmission for Research in Oceanography and Sea Sciences).
- **données 2015** : <http://www.seanoe.org/data/00360/47129/>
- données 2019 mooring Albatross : <https://doi.org/10.17882/74513>
- données 2019 MII : <https://doi.org/10.17882/75839>

Ligure-ouest ALBATROSS : EMSO-Ligure Ouest observatory data (mooring ALBATROSS) from 2019-08

Brève présentation des données générées, collectées ou réutilisées :

- **périmètre, type, échelle,**
- **origine : collecte de nouvelles données ou réutilisation de données (conversion / transformation de données ; partage / échange de données ; achat de données),**
- **publications associées**

Question sans réponse.

Droits de propriété intellectuelle

Qui détiendra les droits sur les données et les autres informations créées lors du projet ?

- CNRS INSU
- M.I.O UMR 7294 CNRS
- DT INSU UPS 855 CNRS
- OSU PYTHEAS UMS 3470 CNRS

Du matériel protégé par des droits spécifiques sera t-il utilisé au cours du projet ? Dans ce cas, qui s'occupe des formalités à accomplir, obtient les autorisations d'utilisation et de diffusion éventuelle ...

- Pas de matériel protégé par des droits spécifiques au cours du projet

Confidentialité

Identification des jeux de données confidentielles

- Pas de données confidentielles
- les données EMSO Ligure Ouest sont publiques

Quelles sont les mesures prises et les normes auxquelles il est nécessaire de se conformer pour garantir cette confidentialité ?

Les données EMSO sont accessibles en libre accès sur la plateforme ERDAPP de l'OSU Pytheas

- <https://erddap.osupytheas.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=EMSO>

à ce jour il n'y a pas de restriction de droits sur les données qui sont en accès public

Le cas échéant, comment la confidentialité de données fournies par des personnes sera garantie lorsque les données seront partagées ou rendues disponibles pour une analyse de second niveau ?

- Données publiques,
- pas d'identification nécessaire
- statistiques de traçabilité des accès possibles mais non en vigueur actuellement

Partage des données à l'issue du projet

Y a t'il une obligation de partage (ou à l'inverse une interdiction ou une restriction)

- Obligation de mise à disposition publique des données brutes en temps « réel » puis des données validées dans les 2 ans.
 - <https://erddap.osupytheas.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=EMSO>

Quelles données seront partagées à l'issue du projet ? Si toutes les données ne sont pas disponibles de la même façon, ou en même temps, le préciser

toutes les données des capteurs hébergés sur le MII et la ligne ALBATROSS sont rendues publiques et partagées sur la plateforme logicielle ERDDAP

- <http://erddap.osupytheas.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=emso>
- **MIl capteur CSTAR** : *Instrument_Serial, Reference_Counts, Signal_Counts, Corrected_Signal_raw_counts, Calculated_Beam_c_inverse_meters, Internal_thermistor_counts*
- **MIl capteur OXYGEN** : *Optode_Type, Optode_Serial, Oxygen, Saturation_Oxygen (ml/l), Temperature*
- **MIl capteur MICROCAT** : *Temperature, Conductivity, Sea Water Electrical Conductivity, Pressure*
- **MIl capteur AQUADOPP** : *s.time, X_East, Y_North, Z_up, Amplitude_X, Amplitude_Y, Amplitude_Z, Batterie, SoundSpeed, Heading, Pitch, Roll, Pressure, Temperature In Situ, Speed, Direction*
- **Ligne Instrumentée ALBATROSS : capteur MICROCAT** : *s.stationname, s.latitude, s.longitude, s.time, s.sensor_theoric_depth, s.sensor_serial_number, s.temperature, s.conductivity, s.pressure, s.oxygen, s.depth, s.salinity, s.potential_temperature, s.oxygen_saturation, s.sigma_theta*

Quelles sont les réutilisations potentielles de ces données ?

- recherche scientifique ,
- Surveillance de l'environnement marin profond, impact des changements globaux.

La lecture des données nécessite-t-elle le recours à un logiciel ou un outil spécifique ? Si oui, lequel et comment y accéder ?

Les données brutes acquises par les capteurs in situ sont traitées :

- par différents scripts en Python3 (un script par capteur) pour générer des données Ascii, au format CSV et au format NetCDF,
 - Les fichiers au format CSV sont utilisables par n'importe quel logiciel comme lisant le CSV (Excel, LibreOffice)
 - Les fichiers NetCDF sont davantage standardisés et comportent toutes les métadonnées nécessaires dans l'entete NetCDF
-
- les données NetCDF du serveur sont accessibles par n'importe quel logiciel client supportant le protocole DAP (thredds, Hyrax, Erddap)
-
- Les données CSV et NetCDF sont intégrés dans une plateforme de gestion de données ERDDAP qui permet l'interrogation et le graphage des données
 - <http://erddap.osupytheas.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=EMSO>

Comment les données seront-elles partagées ?

Les données sont partagées via une plateforme de gestion de données ERDDAP qui permet la mise à disposition, l'interrogation et le graphage des données

- <http://erddap.osupytheas.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=emso>

Les données sont partagées avec l'ERIC EMSO européen via un réseau de serveur ERDAP

Les données sont aussi transmises sur la plateforme SEANOE

- <https://www.seanoe.org/data/00633/74513/>

Avec qui ? sous quelle licence ?

- licence publique

A partir de quand ?

- début Janvier 2015
- 1er mouillage 2017
- 2eme mouillage des lignes Mai 2019

Pendant combien de temps ?

- acquisition continue
- Durée de vie de l'observatoire, plus de 10 ans.

Les données seront-elles identifiées par un identifiant pérenne (DOI ou autre) ?

- DOI données 2015: 00360/47129
 - <http://doi.org/10.17882/47129>
- DOI données Albatross 2019
 - <https://doi.org/10.17882/74513>
- DOI données robot MII 2019
 - <https://doi.org/10.17882/75839>

Quel est l'organisme qui se chargera de la demande d'identifiant dans le cas de projets multi-partenaires ?

- Ifremer Brest. Frédéric Merceur,
- <https://www.seanoe.org/recordview>

Description et organisation des données

Ligure-ouest MII : EMSO-Ligure Ouest observatory data (MII) from 2019-08

Quels méthodes et outils sont utilisés pour acquérir et traiter les données ? Précisez les différents formats dans lesquels les données seront disponibles aux différentes phases de la recherche

Méthode et outils

I/ Acquisition des données

les données brutes EMSO Ligure Ouest sont issues des capteurs MII et ALBATROSS situés à -2465 m sont concentrées sur un serveur de l'IN2P3 antbjs.in2p3.fr à Toulon

- emso-mio@antbjs.in2p3.fr:/data/albatross/meust/MICROCAT
- emso-mio@antbjs.in2p3.fr:/data/albatross/meust/CTAR
- emso-mio@antbjs.in2p3.fr:/data/albatross/meust/OXYGEN

II/ Transfert des données

les données brutes des fichiers de capteurs sont transférées quotidiennement et automatiquement par réseau depuis le site de la DT INSU à La Seyne/mer, le matin à 5h sur les disques de l'OSU Pytheas du site de Luminy à Marseille.

Les données brutes des capteurs arrivent dans le répertoire réseau :

- racine_MII="/mnt/EMSO/PACHA/albatross/meust/"
- racine_albatross="/mnt/EMSO/PACHA/albatross/meust/acoustic/data/"

un script "recup_emso_aquadopp.sh" récupère les fichiers bruts et les dépose dans un répertoire de traitement appelé /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_InputFiles/

III/ Traitement des fichiers bruts

Les données brutes sont alors traitées par des programmes Python et rendues accessibles sur la machine

erddap.osupytheas.fr

- decodeEMSO-MII-MCO.py : décode les données du MII des capteurs CSTAR, MICROCAT et OXYGENE
- decodeEMSO-MII-Aquadoppy.py : décode les données du MII du capteur AQUADOPP
- decodeEMSO-Albatross-Aquadoppy.py : décode les données de la ligne ALBATROSS du capteur AQUADOPP
- decodeEMSO-Albatross-Micro.py : décode les données de la ligne ALBATROSS du capteur MICROCAT

les programmes décodent les fichiers de capteurs et transforment les données au format CSV et NETCDF

les données traitées sont déposées dans

II /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutCSV/
total 95

drwxr-xr-x 18 root root 36 Dec 1 2020 Emso_Albatross
drwxr-xr-x 29 root root 38 Jul 9 02:45 Emso_Aquadoppy
drwxr-xr-x 31 root root 39 Jul 9 05:30 Emso_Cstar
drwxr-xr-x 32 root root 40 Jul 9 05:30 Emso_Microcat
drwxr-xr-x 9 root root 9 Apr 20 2020 Emso_Oxygen

II /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutNetCDF/
total 96

drwxr-xr-x 19 root root 19 Dec 1 2020 Emso_Albatross
drwxr-xr-x 2 root root 3 Jul 1 02:15 Emso_Albatross-2019
drwxr-xr-x 29 root root 38 Jul 9 02:45 Emso_Aquadoppy
drwxr-xr-x 31 root root 39 Jul 9 05:30 Emso_Cstar
drwxr-xr-x 2 root root 3 Jul 1 02:15 Emso_MII-2019
drwxr-xr-x 32 root root 40 Jul 9 05:30 Emso_Microcat
drwxr-xr-x 9 root root 9 Apr 20 2020 Emso_Oxygen

IV/ Diffusion et Affichage dans ERDDAP

Les fichiers ascii au format CSV et NETCDF sont directement pris en compte par la plateforme logicielle erddap

- <http://erddap.mio.univ-amu.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=EMSO>

Documentation associée aux données

Les métadonnées sont affichées par la plateforme Erddap qui analyse et permet d'afficher les **métadonnées** pour chaque jeu de données

- **Station de fond MII**
 - capteur AQUADOPP : https://erddap.osupytheas.fr/erddap/tabledap/Emso_Ligure_Ouest_MII_Aquadoppy_NetCDF.graph
 - Capteur CSTAR : https://erddap.osupytheas.fr/erddap/tabledap/Emso_Ligure_Ouest_MII_Cstar_NetCDF.graph
 - Capteur MICROCAT : https://erddap.osupytheas.fr/erddap/tabledap/Emso_Ligure_Ouest_MII_Microcat_NetCDF.graph
 - Capteur OXYGEN : https://erddap.osupytheas.fr/erddap/tabledap/Emso_Ligure_Ouest_MII_Oxygen_NetCDF.graph
- **Ligne instrumentée ALBATROSS**
 - https://erddap.osupytheas.fr/erddap/tabledap/Emso_Ligure_Ouest_Albatross_Microcat_csv.graph
 - https://erddap.osupytheas.fr/erddap/tabledap/Emso_Ligure_Ouest_Albatross_Microcat_NetCDF.graph
- Les données sont automatiquement acquises quotidiennement
- des programmes python3 transforment les données de capteur en fichiers CSV et NetCDF
- les fichiers sont placés en temps réel sur [la plateforme ERDDAP de l'OSU Pytheas](#)

Les métadonnées normalisées du projet EMSO Ligures ouest sont également placées **sur le catalogue geonetwork de l'OSU Pytheas**:

- <https://dataset.osupytheas.fr/geonetwork/srv/fre/catalog.search#/metadata/658d6730-6d00-463d-86ba-fad78301cc9f>

Quels types de métadonnées seront produites pour accompagner les données ? Quels sont les standards et les vocabulaires ou taxonomies qui seront utilisés pour décrire les données ?

- métadonnées ISO19115 et ISO 1939
- Conventions : COARDS, CF-1.6, ACDD-1.3
- InfoURL
- creator_name
- Institution
- lineage
- Summary
- licence
- request for acknowledgement

Comment les métadonnées seront elles produites ?

I/ Les métadonnées des mesures suivent la convention CF-1.6 et sont produites par les programmes python dans l'entete des fichiers NetCDF déposés sur la plateforme erddap

II/ Les métadonnées ISO 19139 sont également produites dans le catalogue geonetwork de l'OSU Pytheas

- <https://dataset.osupytheas.fr/geonetwork/srv/fre/catalog.search#/metadata/658d6730-6d00-463d-86ba-fad78301cc9f>

Comment les fichiers de données sont-ils gérés et organisés au cours du projet : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers...

Les fichiers de données sont stockés sur un disque réseau partagé et monté sur le serveur erddap.osupytheas.fr dans le répertoire

- /mnt/EMSO/EMSO_Data/DATA_MII/micro_csv/
- /mnt/EMSO/EMSO_Data/DATA_MII/oxygene_csv
- /mnt/EMSO/EMSO_Data/DATA_MII/ctar_csv

convention de nommage : les fichiers sont appelés **nomducapteur_AAAMMJJ.csv**

- MICROCAT_20171210.csv
- CTAR_20171209.csv
- OXYGEN_20171208.csv

Après traitement par les programmes Python3, les fichiers *NetCDF* et *CSV* produits intégrés à ERDAPP sont placés dans des répertoires /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutNetCDF/

- root@erddap:~# ll /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutNetCDF/
total 96
drwxr-xr-x 13 root root 65 Jul 15 05:50 Emso_Albatross
drwxr-xr-x 2 root root 3 Jul 10 12:25 Emso_Albatross-2019
drwxr-xr-x 17 root root 32 Jul 15 04:45 Emso_Aquadopp
drwxr-xr-x 20 root root 34 Jul 15 05:30 Emso_Cstar

```
drwxr-xr-x 2 root root 3 Jul 10 12:25 Emso_MII-2019
drwxr-xr-x 20 root root 34 Jul 15 05:30 Emso_Microcat
• drwxr-xr-x 9 root root 9 Apr 20 16:38 Emso_Oxygen
```

Après traitement par les programmes Python3, fichiers *bruts* traités sont déplacés dans des répertoire II
/mnt/EMSO/EMSO_Data/DATA_MII/EMSO_Treated/

```
root@erddap:~# II /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_Treated/
total 167
```

```
drwxr-xr-x 2 root root 1274 Jul 15 05:50 Emso_Albatross
drwxr-xr-x 2 root root 421 Jul 15 09:45 Emso_Aquadop
drwxr-xr-x 2 root root 503 Jul 15 05:30 Emso_Cstar
drwxr-xr-x 2 root root 496 Jul 15 05:30 Emso_Microcat
drwxr-xr-x 2 root root 142 Apr 17 09:41 Emso_Oxygen
```

Quelle est la procédure de contrôle qualité des données ? joindre éventuellement le plan d'assurance qualité

- pas de contrôle qualité des données EMSO MII actuellement.
- le logiciel istSOS <http://istsos.oupytheas.fr> peut permettre de faire du controle qualité et rajouter des codes qualité sur les données

Ligure-oues ALBATROS : EMSO-Ligure Ouest observatory data (mooring ALBATROSS) from 2019-08

Quels méthodes et outils sont utilisés pour acquérir et traiter les données ? Précisez les différents formats dans lesquels les données seront disponibles aux différentes phases de la recherche

Question sans réponse.

Documentation associée aux données

Question sans réponse.

Quels types de métadonnées seront produites pour accompagner les données ? Quels sont les standards et les vocabulaires ou taxonomies qui seront utilisés pour décrire les données ?

Question sans réponse.

Comment les métadonnées seront elles produites ?

Question sans réponse.

Comment les fichiers de données sont-ils gérés et organisés au cours du projet : contrôle des versions, conventions de nommage des fichiers, organisation des fichiers...

Question sans réponse.

Quelle est la procédure de contrôle qualité des données ? joindre éventuellement le plan d'assurance qualité

Question sans réponse.

Stockage et sécurité des données

Stockage : Quels seront les supports utilisés pour les données au cours du projet ?

Les données EMSO sont stockées

- sur le serveur antbjs.in2p3.fr à Toulon
- emso-mio@antbjs.in2p3.fr:/data/albatross/meust

puis transférées, vers les serveurs de l'OSU pytheas à Luminy, gérés par le service informatique de l'OSU Pytheas

- sur le NAS
 - **172.20.9.100:/mnt/md72/MIO/EMSO /mnt/EMSO**
- répertoire de stockage accessible via le serveur **erddap.osupytheas.fr**

Stockage : Quels seront les types de flux empruntés par les données au cours du projet ?

un script shell recup_emso.sh

*transfère quotidiennement les données par le réseau **entre la baie de stockage de Toulon et la baie de stockage de l'OSU Pytheas sur le site de Luminy à Marseille***

- de emso-mio@antbjs.in2p3.fr:/data/albatross/meust
- vers erddap.osupytheas.fr

Sur le serveur

- *erddap.osupytheas.fr*
- dans le répertoire */root/EMSO*
 - le script *recup_emso.sh* va chercher les fichiers de la journée antérieure
 - et les dépose dans le répertoire */mnt/EMSO/EMSO_Data/DATA_MII/EMSO_InputFiles/*

Stockage : Quelle est la volumétrie prévisionnelle ?

volumétrie faible:

- fichiers bruts 377M sur 24 mois

Fichiers bruts de capteurs depuis 2019 (au 5/7/2021)

- **Juillet 2021 :**
 - **\$ du -sh /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_Treated/**
377M /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_Treated/

```
ll /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_Treated/
total 167
drwxr-xr-x 2 root root 1274 Jul 15 05:50 Emsa_Albatross
drwxr-xr-x 2 root root 421 Jul 15 10:45 Emsa_Aquadopp
drwxr-xr-x 2 root root 503 Jul 15 05:30 Emsa_Cstar
drwxr-xr-x 2 root root 496 Jul 15 05:30 Emsa_Microcat
drwxr-xr-x 2 root root 142 Apr 17 09:41 Emsa_Oxygen
```

Fichiers de sortie depuis 2019 sur 24 mois (au 5/7/2021)

- root@erddap:~/EMSO# du -sh /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutCSV/
73M /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutCSV/
- root@erddap:~/EMSO# du -sh /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutNetCDF/
151M /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_OutNetCDF/

Stockage : Où sont hébergées physiquement les données, sur quel type d'hébergement ?

- les données brutes et les données traitées (NetCDF) sont stockées sur un **NAS du Service Informatique de l'OSU Pytheas sur le site de Luminy à Marseille (batiment Oceanomed)**
 - **172.20.9.100:/mnt/md72/MIO/EMSO**
- dans une salle serveur climatisée du MIO UMR 7294 CNRS , batiment Oceanomed sur le campus de Luminy

Stockage : Où sont localisées géographiquement les données ?

- les données traitées sont stockées sur un **NAS**
 - **172.20.9.100:/mnt/md72/MIO/EMSO**

- dans une salle serveur climatisée du service informatique de l'OSU Pytheas :
 - **batiment Oceanomed sur le campus de Luminy , Marseille ; FRANCE**

Sécurité : L'entité hébergeant physiquement les données a-t-elle une politique de sécurité pour son système d'information ?

Les données EMSO Ligure Ouest sont hébergées sur les baies de disques du service informatique de l'OSU Pytheas avec une politique de sécurité

- les données sont sur un NAS sauvegardé dupliqué
- les données sont sauvegardées dupliquées sur un autre media
- l'accès aux données est pour l'instant public

Sécurité - Confidentialité : les données feront-elles l'objet d'échange ou de partage avec de tiers acteurs ?

Sécurité - Confidentialité :

- Données publiques, avec contraintes de traçabilité des accès.
- Les données brutes sont échangées entre le serveur in2p3 à La Seyne et le serveur de l'OSU Pytheas erddap.osupytheas.fr
- Les données traitées sont publiques et mise à disposition, elles font l'objet d'échange avec le portail Seanoe pour les DOI
 - <https://www.seanoe.org/data/00633/74513/>

Sécurité - Confidentialité : comment sont déterminés les droits d'accès aux données pendant les recherches ?

- Données en accès public, avec contraintes de traçabilité des accès.

Sécurité - Confidentialité : De quelle manière l'ensemble des chercheurs partenaires du projet auront-ils accès aux données pendant la recherche ?

les données sont publiques et en accès libre et mises à disposition via la plateforme erddap.osupytheas.fr

- <https://erddap.osupytheas.fr/erddap/search/index.html?page=1&itemsPerPage=1000&searchFor=EMSO>

Sécurité - Intégrité - Tracabilité : Quelles sont les mesures de protection mises en œuvre pour suivre la production et l'analyse des données au cours du projet ?

- un serveur de monitoring XYMON permet de vérifier que les serveurs sont en état de fonctionnement et que les données sont arrivées dans le répertoire
 - <http://rancid.mio.osupytheas.fr/xymon/>
- le serveur n'est pas public, accessible uniquement en interne sur le laboratoire MIO

Archivage et conservation des données après la fin du projet

Quelles sont les données à conserver sur le moyen ou le long terme et quelles sont les données à détruire ?

données à conserver sur le moyen ou le long terme :

a minima les données **brutes** des capteurs MII et ALBATROSS doivent être conservées sur le serveur **erddap.osupytheas.fr** du répertoire

- **du -sh /mnt/EMSO/EMSO_Data/DATA_MII/EMSO_Treated/**
- Ce répertoire contient les fichiers bruts provenant directement des capteurs des instruments MII et ALBATROSS

les données traitées dans les fichiers NetCDF et .csv peuvent être recalculées et reproduites par les programmes Python3 à partir des fichiers bruts

- machine erddap.osupytheas.fr
- répertoire : /root/EMSO
- programmes python:

```
root@erddap:~/EMSO# ls *.py
```

- EMSO_Library.py
- decodeEMSO-MII-Aquadopp.py
- decodeEMSO-Albatross-Aquadopp.py
- decodeEMSO-MII-MCO.py
- decodeEMSO-Albatross-Micro.py functions_Albatross.py
- L'observatoire OSU Pytheas permettra de fournir de la donnée pérenne (plus de 10 ans).

Sur quelle plateforme d'archivage pérenne seront archivées les données à conserver sur le long terme ? Sinon, quelles procédures seront mises en place pour la conservation à long terme ?

les données sont stockées sur un NAS du service informatique de l'OSU Pytheas et sauvegardées régulièrement
l'OSU Pytheas peut conserver ces données sur 10ans

Il n'est pas prévu d'archivage pérenne à ce jour

Quelle est la durée de conservation des données ?

- La vertu de l'observatoire est de fournir de la données pérennes (plus de 10 ans).

Qui sera responsable de la conservation à long terme ? nommer un contact individuel.

- *P.I : Dominique Lefevre dominique.lefevre@mio.osupytheas.fr*
- *data manager : Maurice Libes : maurice.libes@osupytheas.fr*
- *service informatique de l'OSU Pytheas ; sip@osupytheas.fr*
- *DT INSU*

Quel sera le volume de ces données ?

- *quelques centaines de mégaoctet de fichiers bruts et traités en Juillet 2021 (depuis 2019)*

```
$ du -sh /mnt/EMSO/EMSO_Data/DATA_MII/
624M /mnt/EMSO/EMSO_Data/DATA_MII/
```

soient environ 4 a 5Go sur 10 ans

Quelles garanties de financements couvriront les coûts associés à la conservation à long terme ?

- *probleme non traité à a voir avec EMSO France*
- *voir si les données nécessitent d'etre archivées de manière pérenne aupres du CINES par exemple*