



SeaDataNet

Outils SeaDataNet utilisant les référentiels vocabulaires

Michèle Fichaut (IFREMER)

Atelier technique ODATIS, Marseille, 8 octobre 2019
sdn-userdesk@seadatanet.org – www.seadatanet.org

Contexte

- Dans le cadre de l'infrastructure SeaDataNet qui met en commun les données de plus de 110 centres de données pan-européens → pour l'interopérabilité, importance de définir des formats de transports de données communs, reposant sur les mêmes référentiels de vocabulaires

Formats (1)

- 3 formats majeurs ont été définis
 - 2 formats ASCII : ODV SeaDataNet et MedAtlas SeaDataNet
 - 1 format binaire : NetCDF SeaDataNet (Cfpoint), compatible CF
- Permettent de gérer des données de :
 - profils verticaux (dans la colonne d'eau ou dans le sédiment) : carottes sédimentaires, CTD, XBT-MBT, prélèvements hydrologiques, profileurs Argo, ...
 - Séries temporelles mesurées au point fixe: courantomètres, thermistances, pièges à sédiment...
 - Trajectoires: thermosalinographe, bathymétrie, ADCP....

Formats (2)

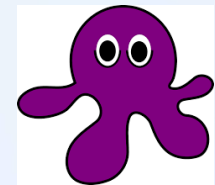
- 3 variantes ODV ont été également définies pour gérer les données de
 - Biologie
 - Microplastiques
 - Cytométrie en flux
- 1 variante NetCDF pour gérer les données de HF-radar
- <https://www.seadatanet.org/Standards/Data-Transport-Formats>

Point commun des formats SeaDataNet

- Contiennent des informations spécifiques pour SeaDataNet
 - LOCAL_CDI_ID : identifiant unique de la données pour un centre de données
 - EDMO_Code : identifiant unique du centre de données donnant accès aux données dans SeaDataNet
 - Utilisation des vocabulaires SeaDataNet pour les codes paramètres (P01), les unités (P06), les capteurs (L22)
 - Possibilité de faire référence aux catalogues SDN des campagnes → lien vers le CSR de la campagne où ont été collectées les données

Outils pour gérer ces formats

- Nemo: outil de formatage SeaDataNet
 - pour générer des fichiers au formats SeaDataNet à partir de fichiers en entrée ASCII : issus des capteurs ou propre à un laboratoire ou à une personne.
- Octopus: convertisseur SeaDataNet
 - Pour générer des formats SeaDataNet, à partir d'un format SeaDataNet



NEMO [version courante 1.6.7]



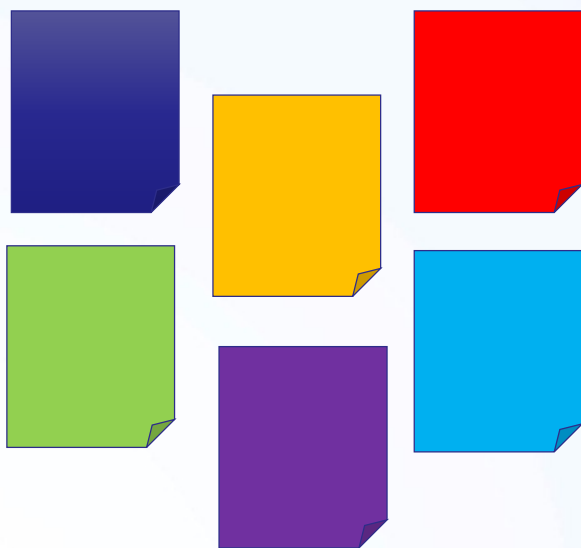
- Téléchargeable sur le site SeaDataNet
<https://www.seadatanet.org/Software/NEMO>
- Ecrit en **Java** (Version ≥ 1.7)
- Bilingue (Français, Anglais)
- Disponible pour plusieurs environnements :
 - Microsoft : Windows
 - Unix - Solaris
 - Linux
- 2 modes : interactif et batch
- Utilise la librairie Java1.8 (IFREMER) sur les vocabulaires **SeaDataNet**
 - Pour mettre à jour les listes de vocabulaires SeaDataNet
 - *NEMO a besoin d'une connexion internet pour avoir des listes à jour.*
 - *Mais NEMO peut-être utilisé sans connexion dès que les listes sont à jour.*
 - *works offline once the lists are up-to-date*

NEMO

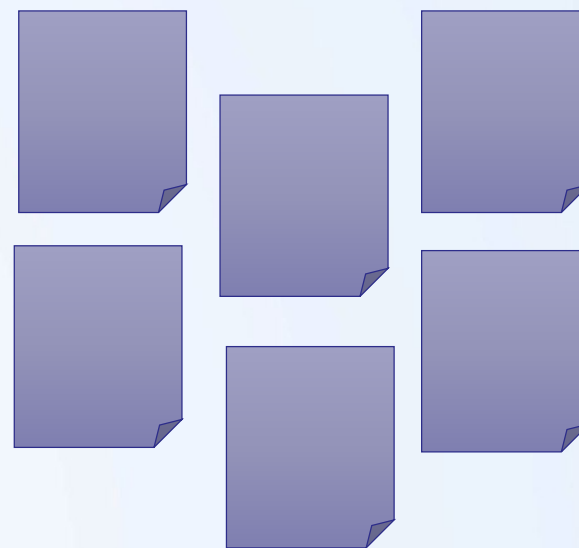
- Mise aux formats SDN de fichiers ASCII



Fichiers quelconques ASCII colonnes



Fichiers aux formats SDN



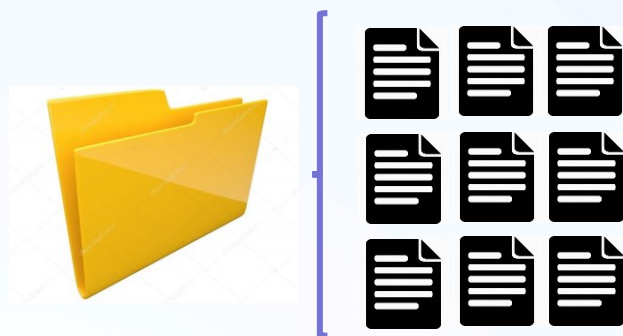
ODV, NetCDF ou MedAtlas

NEMO, fonctions principales (1)

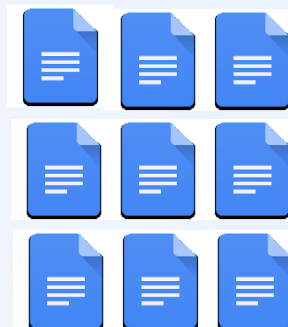


- NEMO est capable de convertir

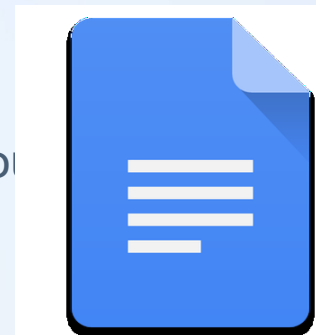
Un repertoire avec des fichiers ASCII homogènes



Fichiers au format SDN

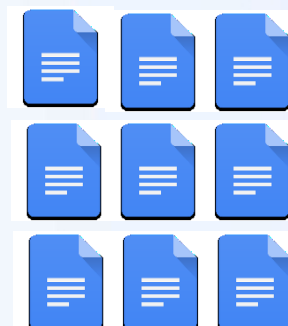


ou

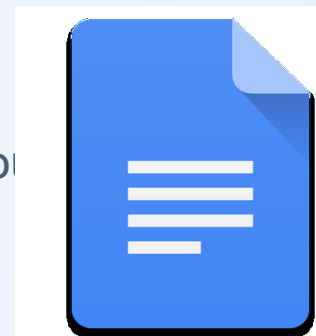


Ou un fichier contenant plusieurs “stations” homogènes

```
Station 1, date, time, lat, lon
Param1 param2 param3 param4
value1 value1 value1 value1
value2 value2 value2 value2
...
Station 2, date, time, lat, lon
Param1 param2 param3 param4
value1 value1 value1 value1
```



ou



NEMO, fonctions principales (2)

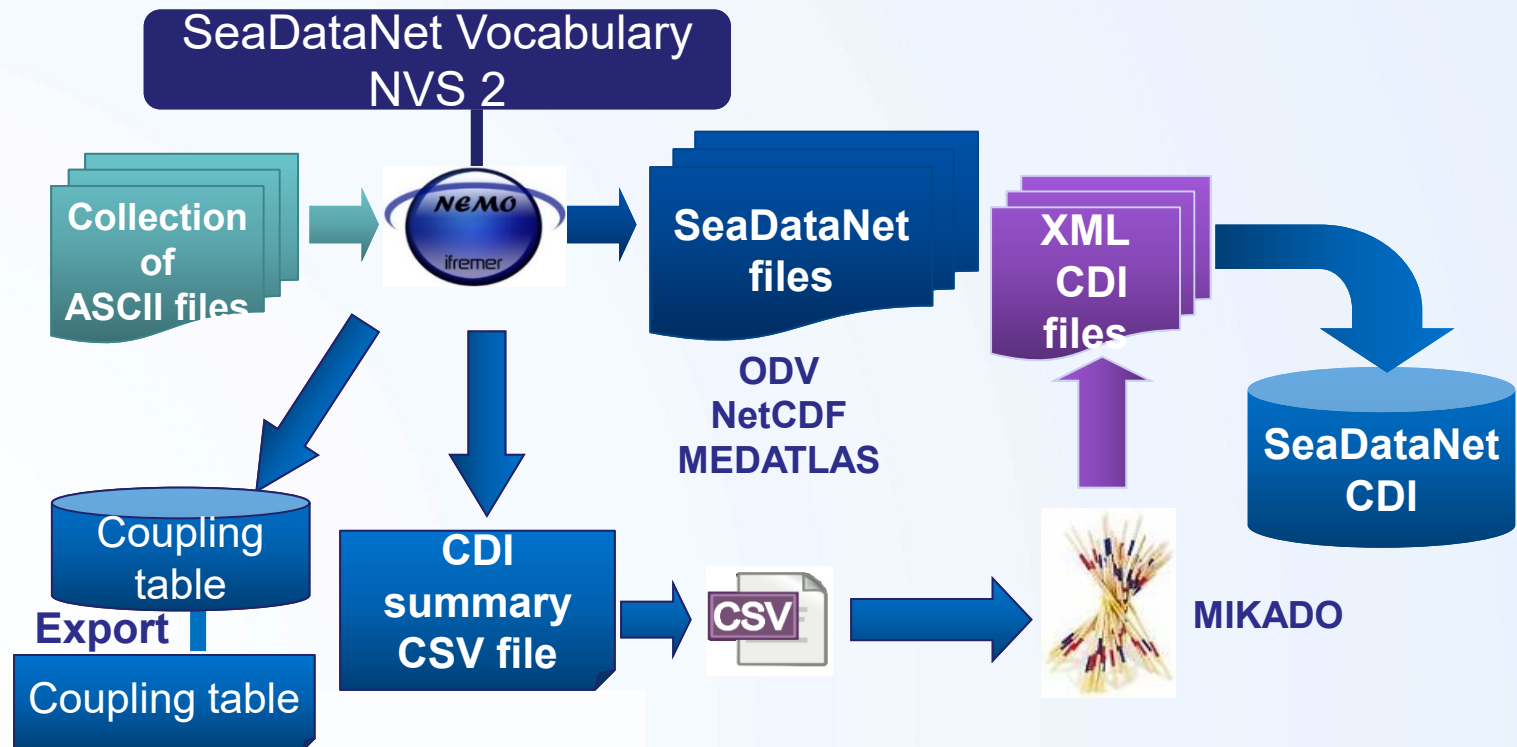


- Les fichiers ASCII de NEMO peuvent être:
 - **Des profils verticaux:** Paramètre de référence PROFONDEUR ou PRESSION, Croissant, même lat-lon pour tout le profil
 - **Des séries temporelles:** Paramètre de référence DATE-HEURE Croissant, même lat-lon pour toute la série
 - **Des trajectories:** Paramètre de référence PROFONDEUR ou PRESSION, lat-lon change sur toutes les lignes de données



NEMO, fonctions principales (3)

- NEMO est intégré dans l'infrastructure SDN et interagit avec d'autres outils comme MIKADO le générateur de descriptions des données au format XML ISO-19139





Principes de NEMO

- NEMO est capable de lire beaucoup de formats ASCII et de les convertir aux formats SeaDataNet
- Les utilisateurs de NEMO décrivent leurs fichiers ASCII en entrée → pour que NEMO soit capable de trouver les informations nécessaires aux formats SeaDataNet
- **Prérequis obligatoire:** l'ensemble des fichiers en entrée de NEMO doit être homogène, les informations doivent
 - **Être localisées à la même position:** même ligne, même position sur la ligne ou même colonne (si format CSV)
 - **Être au même format**
 - *Par exemple la température se trouve toujours:*
 - *Dans la deuxième colonne des paramètres mesurés (CSV),*
 - *du 10ème au 14ème caractère sur la ligne de données (format TXT sans séparateur)*

Fichiers qui ne peuvent pas être convertis avec NEMO

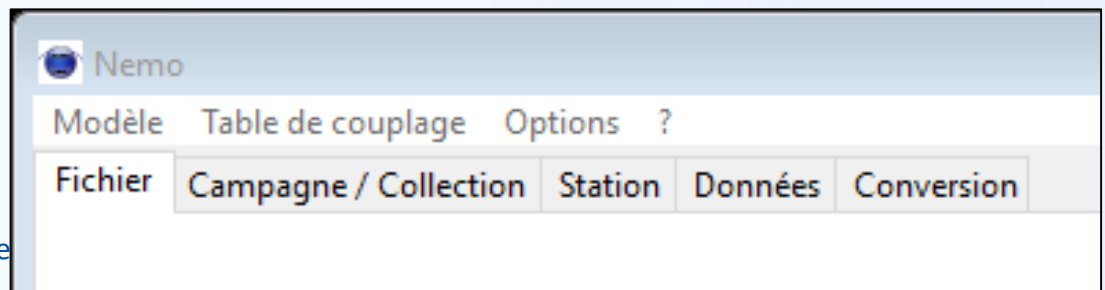


- Formats binaires : EXCEL, WORD ...
 - Les fichiers doivent d'abord être convertis au format texte
- Les fichiers ne respectant pas les prérequis de NEMO
 - Même format, même position
- Les fichiers qui ne sont pas des
 - **Profils verticaux** (avec Prof. ou Pres. comme référence verticale)
 - **Time series** (avec date/heure comme référence)
 - **Trajectories** (avec depth, lat-lon comme référence)



Principes de NEMO

- 1 étape de configuration de NEMO
- 5 étapes pour convertir les fichiers
 1. Description du type de fichier
 2. Métadonnées sur la collection ou sur la campagne de collecte
 3. Métadonnées de “Station”
 4. Description des paramètres mesurés
 5. Conversion





Etape 1 - Description du type de fichier (1)

Modèle Table de couplage Options ?

Fichier Campagne / Collection Station Données Conversion

Numéro station;commentaire station:date station:heure station:latitude;lon
 N0001;RAS;16/01/2009;22:55;N49 041.61;O000 048.87;1:50.22842437;1:1.94000
 N0002;Chlorophylle_R2_perdue;17/01/2009;02:29;N49 036.90;O000 000.38;1:30
 N0003;RAS;17/01/2009;04:28;N49 046.22;E000 004.15;1:35.72762308;1:0.93999
 N0004;RAS;17/01/2009;07:25;N49 050.33;E000 009.48;1:30.4652134;1:1.41999
 N0005;RAS;17/01/2009;11:26;N50 006.33;E000 032.94;1:35.50028544;1:1.87000
 N0006;RAS;17/01/2009;15:56;N50 011.66;E001 010.46;1:29.20029348;1:1.5;1.8
 N0007;RAS;17/01/2009;18:07;N50 017.54;E001 015.66;1:32.96240385;1:0.61000
 N0008;RAS;17/01/2009;20:02;N50 027.87;E001 017.30;1:31.35534483;1:0.37000
 N0009;RAS;17/01/2009;22:52;N50 046.26;E001 029.02;1:35.21355769;1:1.40000
 N0010;RAS;18/01/2009;04:38;N50 036.76;E001 030.26;1:29.77583333;1:1.04999
 N0011;RAS;18/01/2009;07:15;N50 032.33;E001 030.05;1:21.50399194;1:0.37999
 N0012;RAS;18/01/2009;13:13;N50 057.08;E001 040.35;1:31.0925;1:0.15999999
 N0013;RAS;18/01/2009;17:25;N50 059.20;E001 006.25;1:29.85693549;1:3.39000
 N0015;RAS;18/01/2009;21:05;N50 042.32;E000 036.95;1:44.36518519;1:0.49000

Paramètres d'entrée

Traiter une campagne
☒ fichier campagne
☐ dossier campagne

Traiter une collection
☐ Fichiers regroupés par campagne
☐ Fichiers non regroupés par campagne
☐ Fichier multi-stations

Parcourir

☒ Séparateur CSV
☐ Tabulation ☒ Point-virgule ☐ Virgule
☐ Espace ☐ Autre:

Type de fichier
☒ Profil Vertical
☐ Série temporelle
☐ Trajectoire

Variante
☒ physico-chimique
☐ biologique
☐ cytométrie en flux
☐ microplastiques

Paramètres de conversion

☐ Medatlas
☒ ODV
☐ NetCDF

☐ Un fichier par station
☒ Un fichier pour l'ensemble des stations

☐ Trier les données au sein des stations par paramètre de référence croissant (Pression ou temps)

Journal Infos Aide

- Fichier avec plusieurs stations?
- Plusieurs fichiers mono-stations
- Fichier(s) avec séparateurs?
 - Quel séparateur
- Type de fichier?
 - Profil vertical
 - Série temporelle
 - Trajectoire
- Variante?
 - Physico/chimie
 - Biologie
 - Cytométrie
 - Microplastiques
- Format de sortie



Etape 1 - Description du type de fichier (2)

Description de fichier

+

- ▶ Entête de fichier *
- ▶ Entête Station *
- ▶ Fin de Station *
- ▶ Indicateur de Fin de Données *
- ▶ Sélection des colonnes de données

- Organisation du fichier
 - Y-a-t il une entête fichier ?
 - Y-a-t-il une entête station ?
 - Y-a-t-il des lignes entre 2 stations consécutives? (si fichier multi-stations)
 - Comment reconnait-on la fin d'une station? (si fichier multi-stations)
 - Si fichiers CSV y-a-t-il des titres de colonnes?

Etape 2 - Description de la campagne/collection (1)



Description de Campagne / Collection

Valider étape

Initialisation serveur

Initialisation XML

Reset

+

▶ Référence *

▶ Localisation

▶ Source des données

▶ Centre d'archivage

▶ Type de données

▶ Commentaires éventuels

- Renseigner les métadonnées sur la campagne / Collection
 1. Référence
 2. Identifiant
 3. Nom
 4. Navire
- Eventuels commentaires importants à garder dans les fichiers

Etape 2 - Description de la campagne/collection (2)

4 : Code Navire

2: Identifiant local
campagne/collection

<pre>//<sdn_reference xlink:href="http://seadatanet.maris2.nl/v_cdi_v3/print_xml.asp?edmo=486&identifieur=FI35200653001_00001_H10" xlink:role="isObservedBy" xlink:type="SDN:L23::CSR"/> //<sdn_reference xlink:href="http://vocab.nerc.ac.uk/collection/C17/current/06M2" xlink:role="isObservedBy" xlink:type="SDN:L23::NVS2CON"/> //<sdn_reference xlink:href="http://seadatanet.maris2.nl/v_cdi_v3/print_xml.asp?edmo=486&identifieur=FI35200653001_00001_H10" xlink:role="isDescribedBy" xlink:type="SDN:L23::CSR"/> //<sdn_reference xlink:href="http://seadatanet.maris2.nl/v_cdi_v3/print_xml.asp?edmo=486&identifieur=FI35200653001_00002_H10" xlink:role="isDescribedBy" xlink:type="SDN:L23::CSR"/> //SDN_parameter_mapping //<subject>SDN:LOCAL:DEPHPR01</subject><object>SDN:P01::DEPHPR01</object><units>SDN:P06::ULAA</units> //<subject>SDN:LOCAL:Pressure</subject><object>SDN:P01::PRESPR01</object><units>SDN:P06::UPDB</units><instrument>SDN:L22::TOOL0409</instrument> //<subject>SDN:LOCAL:Temperature</subject><object>SDN:P01::TEMPPR01</object><units>SDN:P06::UPAA</units><instrument>SDN:L22::TOOL0409</instrument> //<subject>SDN:LOCAL:Salinity</subject><object>SDN:P01::PSLTZZ01</object><units>SDN:P06::UUUU</units><instrument>SDN:L22::TOOL0409</instrument> //<subject>SDN:LOCAL:Oxygen</subject><object>SDN:P01::DOXMZZXX</object><units>SDN:P06::KGUM</units><instrument>SDN:L22::TOOL0036</instrument> //</pre>															
Cruise	Station	Type	yyyy-mm-dd	Longitude [d]	Latitude [d]	LOCAL_CDI_ID	EDMO_code	Bot. Depth [m]	DEPHPR01 [m]	QV:S Pressure [hPa]	QV:S Temperature [°C]	QV:S Salinity [g/kg]	QV:S Oxygen [ml/l]	QV:S Oxygen [ml/l]	QV:S Oxygen [ml/l]
DVIDE 3		1 C	2006-05-24T1	-10.7	38.433333	FI35200653001_00001_H10	486	4915	1	1	1	1	16.512	1	1
									2	1	2	1	16.512	1	1
									3	1	3	1	16.512	1	1
									4	1	4	1	16.512	1	1
									5	1	5	1	16.512	1	1
									6	1	6	1	16.512	1	1

3: Nom
campagne/collection

1: Reference
campagne/collection



Etape 3 - Description de station

Description de Station

Valider étape

Reset

+

▸ Code station

▸ Type de données

▸ Historique Acquisition

▸ Commentaire

▸ Echantillon de surface

▸ Décalage/TU

▸ Heure

▸ Date

▸ Latitude

▸ Longitude

▸ Bathymétrie

▸ QC Flags

- Métadonnées des stations - où lire les infos suivantes dans les fichiers:
 - Code station
 - Date/heure de la station
 - Latitude
 - Longitude
 - Bathymétrie
 -

Etape 4 - Description des données



NEMO - [Fichier C:\Test logiciels\NEMO\Recette\FAE48390\IBTS2009.csv]===== [Modèle modélIBTS_2009_170.xml]

Modèle Table de couplage Options ?

[Fichier] [Campagne / Collection] [Station] Données Conversion

PH	PHOS	QC PHOS	NTRA	QC NTRA	NTRI	QC NTRI	SLCA
1.940000057	1	9.199999809	1	0.50999999	1	10.94999981	
1.099999905	1	11.14999961	1	0.75	1	14.88000011	
0.939999998	1	16.18000031	1	1.36000014	1	9.970000167	
1.419999957	1	11.07999991	1	1.35000014	1	14.35000038	
1.870000005	1	7.309999943	1	1.039999961	1	16.5	
1.5	1	8.340000153	1	0.980000019	1	18.94000053	
0.610000014	1	6.539999961	1	0.49000001	1	4.130000019	
0.370000005	1	6.53000011	1	0.519999981	1	3.809999943	
1.400000095	1	11.64000034	1	0.689999998	1	3.930000067	
1.049999951	1	8.719999541	1	0.560000001	1	3.99000001	
0.379999995	1	9	1	0.610000005	1	3.700000048	
0.159999996	1	11.89000034	1	0.409999996	1	3.960000038	
3.390000105	1	7.71999979	1	0.589999974	1	9.119999541	
0.400000001	1	4.380000114	1	0.100000001	1	2.180000061	

< >

T..	* CODE	LABEL	* UNITÉ	CONVER...	TEST	* DÉB...	* FIN	* FORMAT	VAL. ...	TEST ...	TEST ..
<input type="checkbox"/>	ADEPZZ01 - D...	Depth	Metres	x*1		7	7	%3.1f	NA		
<input type="checkbox"/>	PHOSZZXX - ...	Phosphate	Millimoles p...	x*1		10	10	%6.3f	NA		
<input type="checkbox"/>	NTRAZZXX - ...	Nitrate	Millimoles p...	x*1		12	12	%6.3f	NA		
<input type="checkbox"/>	NTRIZZXX - C...	Nitrite	Millimoles p...	x*1		14	14	%6.3f	NA		

< >

Journal Infos Aide

Description de Données

Valider étape

Reset

+

Listes de paramètres

☐ P09

☐ P01 via P09

☒ P01 via P02

Mesuré

☒ dans la colonne d'eau

☐ sous le fond

☐ Au-dessus de la colonne d'eau

Références verticales

profondeur colonne d'eau

ADEPZZ01

profondeur sous le fond

pression

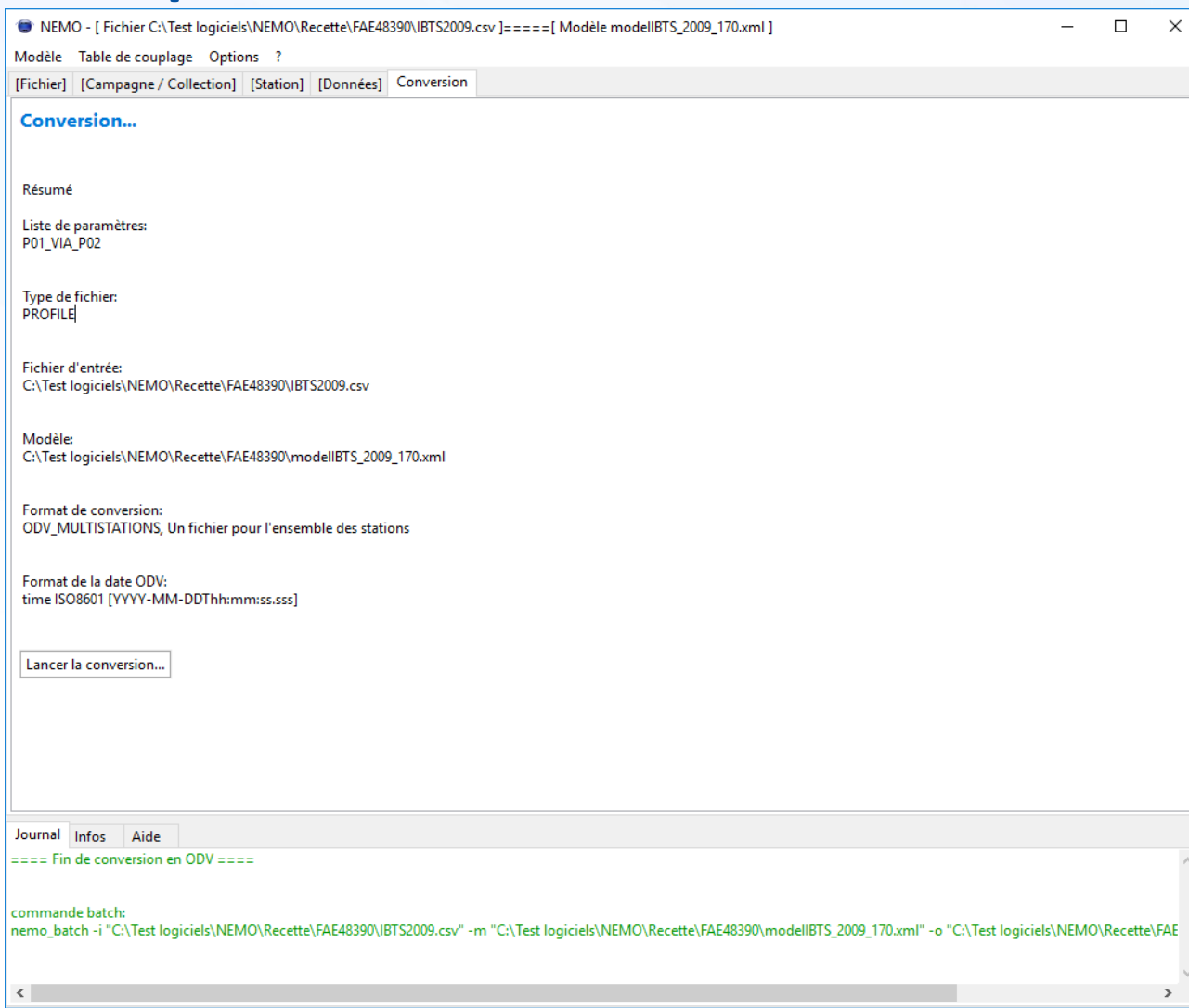
Hauteur

fall rate

sélectionner

- Où lire les paramètres mesurés?
- Quels formats en sortie?
- Quelles sont les valeurs par défaut en entrée?
- Quels codes P01?
- Quelles unités?
- Quel libellé de la colonne ODV correspondante?

Etape 5 - Conversion



- Si OK, message vert et commande Batch qui peut être récupérée
- Si KO, message rouge



NEMO – Sauvegarde de modèles

- Il est possible de sauvegarder la description faite dans NEMO (Menu >Modèle >Enregistrer) pour réutilisation ultérieure en mode interactif ou en mode batch.
- Les modèles sont enregistrés au format XML



NEMO – informations disponibles

- Manuel utilisateur disponible sur le site SeaDataNet
<https://www.seadatanet.org/Software/NEMO>

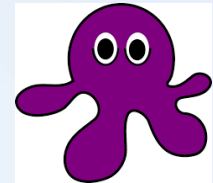
Avec beaucoup de détails sur l'utilisation du logiciel,
et beaucoup de captures d'écran

- Sur cette page aussi:
 - Zip d'exemples de fichiers et des modèles NEMO correspondants
 - Une présentation de NEMO (avec vidéo)
 - Une page FAQ



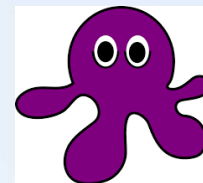
NEMO – Prochaine version

- 1.7.0, livraison prévue le mois prochain
- Nouveautés:
 - Support for les formats ODV biologie, cytométrie en flux, microplastiques
 - Possibilité d'avoir plusieurs fois le même P01 dans la liste des paramètres mesurés:
 - Pour un même paramètre mesuré par 2 capteurs différents
 - Possibilité de mettre les listes de vocabulaires à jour en mode batch
 - Gestion des séparateurs dans les entêtes de stations



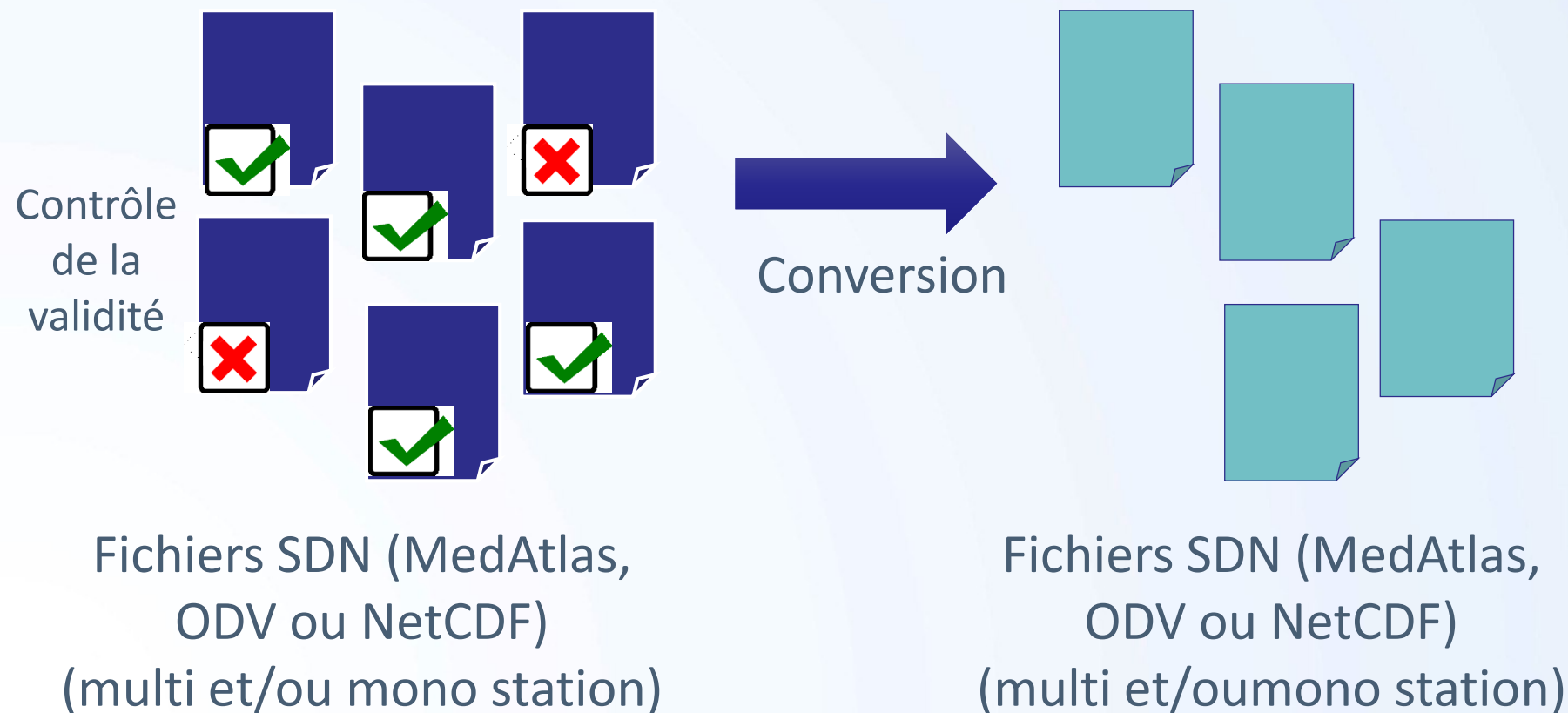
OCTOPUS [version courante 1.5.1]

- Téléchargeable sur le site SeaDataNet
<https://www.seadatanet.org/Software/OCTOPUS>
- Ecrit en **Java** (Version ≥ 1.7)
- Bilingue (Français, Anglais)
- Disponible pour plusieurs environnements :
 - Microsoft : Windows
 - Unix - Solaris
 - Linux
 - Mac (en cours)
- 2 modes : interactif et **batch**
- Utilise les services web sur les vocabulaires **SeaDataNet**

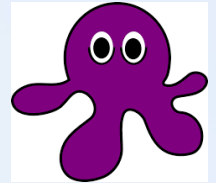


OCTOPUS – Fonctions principales

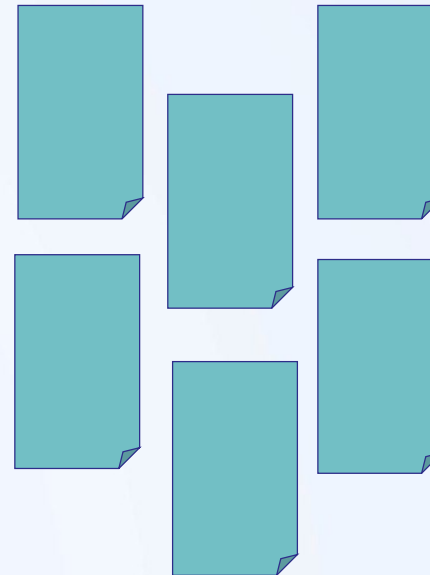
- Pour contrôler et convertir des fichiers SDN



OCTOPUS – Fonctions additionnelles (1)



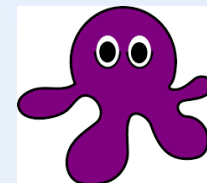
- Pour diviser un fichier multi-stations en n fichiers mono-station



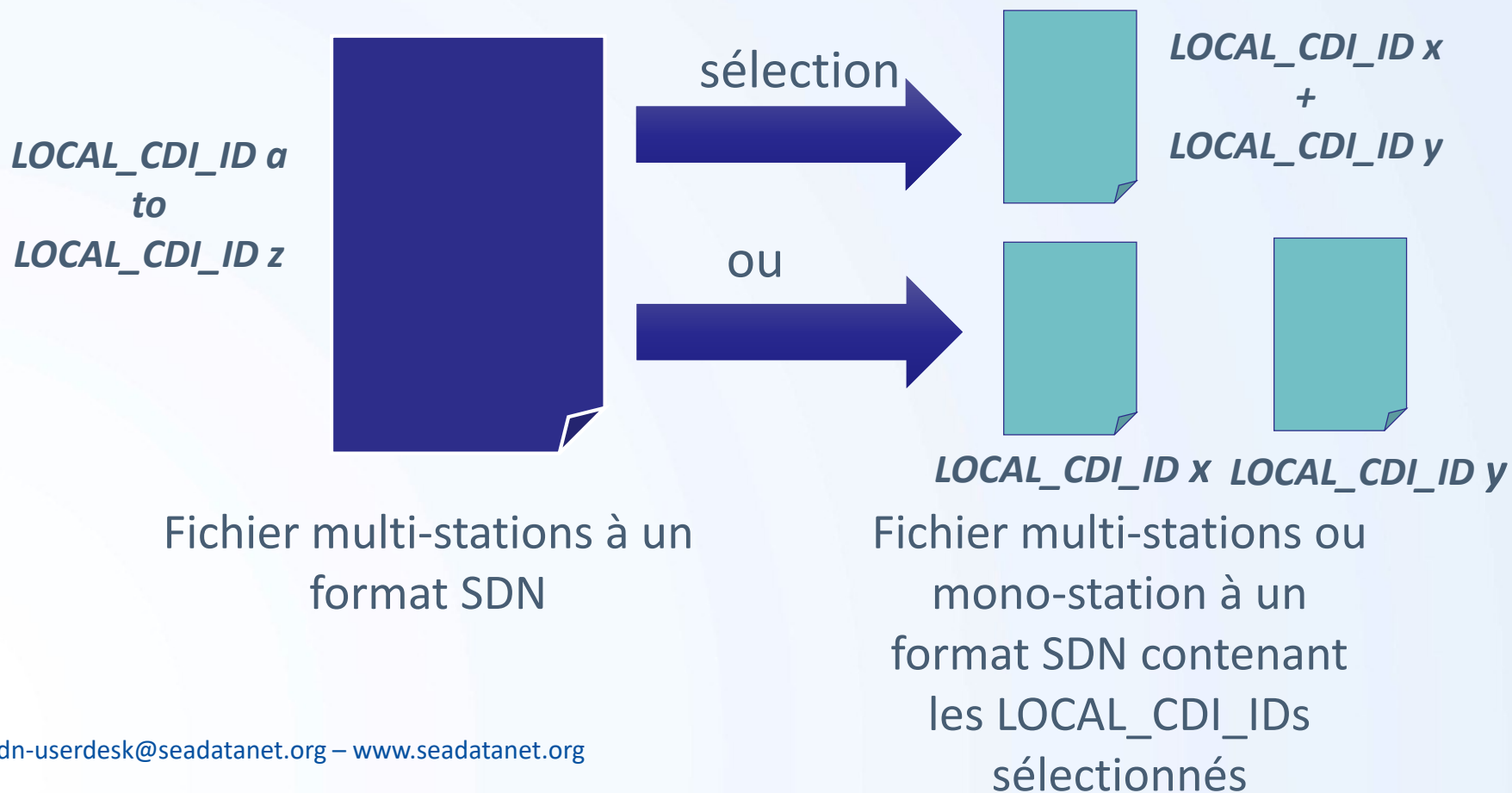
Fichier multi-stations
à un format SDN

Fichiers mono-station à un
format SDN

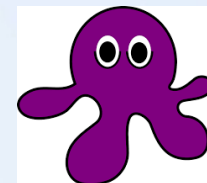
OCTOPUS – Fonctions additionnelles (2)



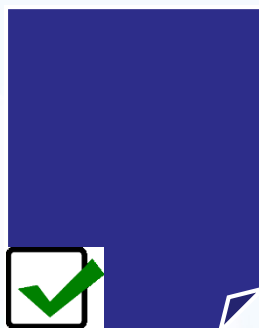
- Sélection de stations de fichiers multi-station



OCTOPUS – Fonctions additionnelles (3)



- Conversion de fichier MGD au format ODV SDN
 - Magnétisme, Gravimétrie et profondeur

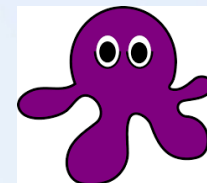


ASCII MGD
V81 et V98

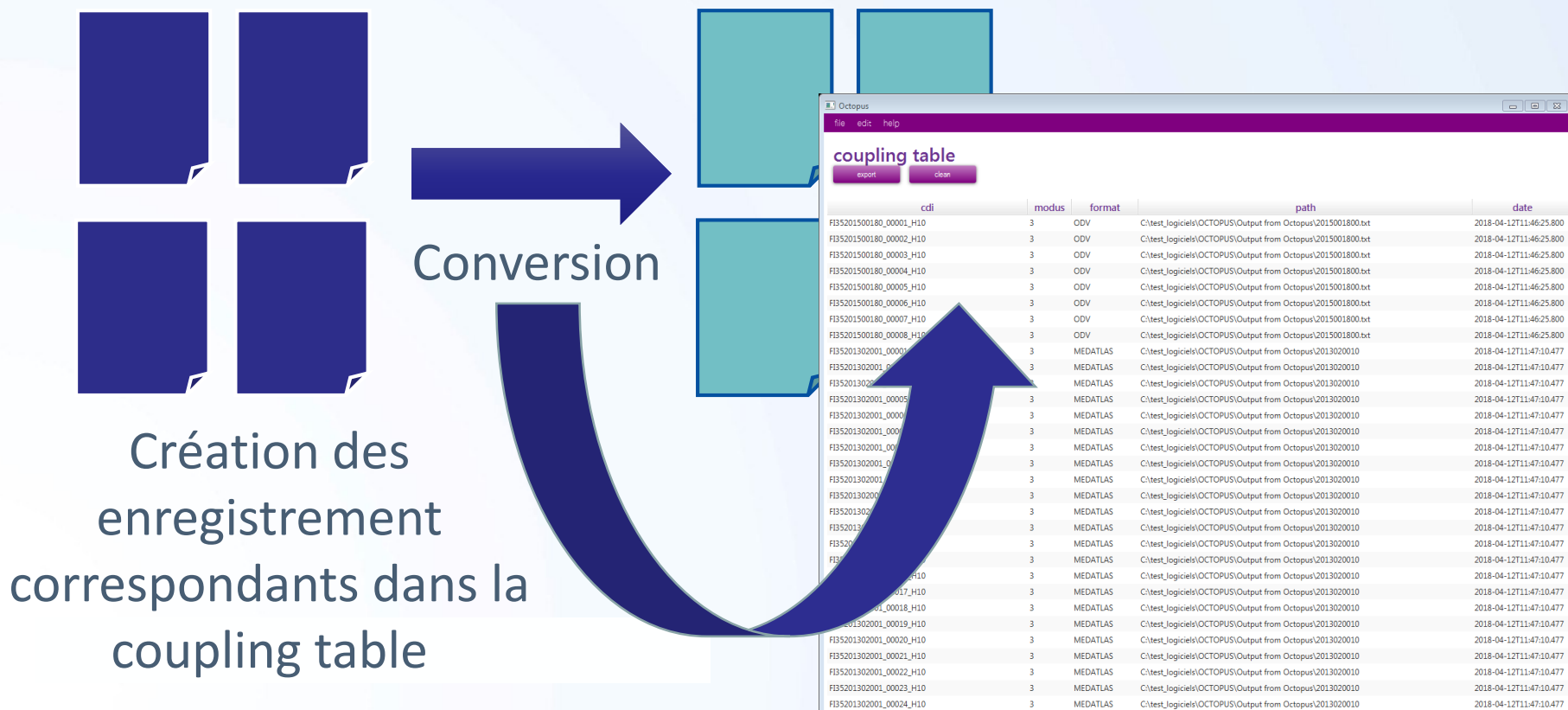


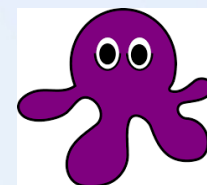
Fichier MGD au
format **ODV** SDN

OCTOPUS – Fonctions additionnelles (4)



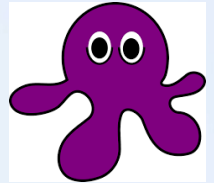
- Pendant la conversion Octopus peut générer un coupling table, pour SDN,





OCTOPUS - Conversions possibles

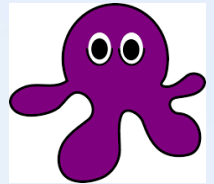
output→ input ↓	MedSDN	ODV	ODV variants	netCDF - CFPoint
Med non SDN	✓	✓	✗	✓
Med SDN	✓	✓	✗	✓
ODV SDN	✗	✓	✗	✓
ODV variants	✗	✗	✓	✗
netCDF-CFPoint	✗	✓	✗	✓
MGDv81	✗	✓	✗	✗
MGDv98	✗	✓	✗	✗



Octopus – Autre fonction utile

- La conversion d'un format SeaDataNet vers le même format SeaDataNet peut également servir à corriger certaines erreurs dans les fichiers
 - Tabulations inutiles
 - Suppression des colonnes de paramètres vides
 - Remplacement des codes BODC obsolètes par les nouveaux codes (pour les fichiers anciens)
 - Rajout de métadonnées → lien vers la description XML du LOCAL_CDI_ID

Octopus - informations disponibles



- Manuel utilisateur disponible sur le site SeaDataNet
<https://www.seadatanet.org/Software/OCTOPUS>
Avec beaucoup de détails sur l'utilisation du logiciel,
et beaucoup de captures d'écran
- Sur cette page aussi:
 - Une présentation d'Octopus(avec vidéo)
 - Une page FAQ
 - Détails des tests effectués lors du contrôle de la validité des formats ODV et NetCDF SeaDataNet
 - Aide à comprendre les erreurs dans les fichiers

Questions

