

FAIRisation des logiciels marins : un atout pour les données géographiques de demain

Erwan BODERE, Léo BUVRY-LAGADEC, Jérôme DETOC (IRSI/ISI)
Clémence COTTEN, Steven PIEL (IRSI/SISMER)

Nantes, 19/03/2025

Plan

- 1) Les principes FAIR appliqués aux logiciels
- 2) Retour d'expérience suite à la FAIRisation de 2 logiciels
- 3) Evaluation FAIR
- 4) Résumé et perspectives



Les principes FAIR appliqués aux logiciels

Logiciel

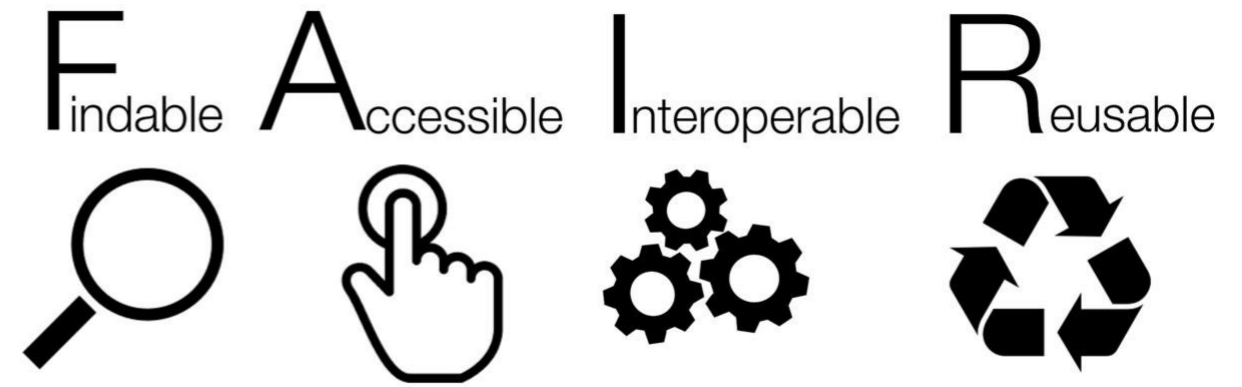
« Ensemble des programmes, procédés et règles, et éventuellement de la documentation, relatifs au fonctionnement d'un ensemble de traitement de données » (Journal Officiel du 17 janvier 1982)

Terme générique utilisé pour désigner les applications, les scripts et les programmes qui s'exécutent sur un dispositif électronique.

Dans tous les cas : ce sont des objets vivants et complexes !



Logiciel « FAIR »



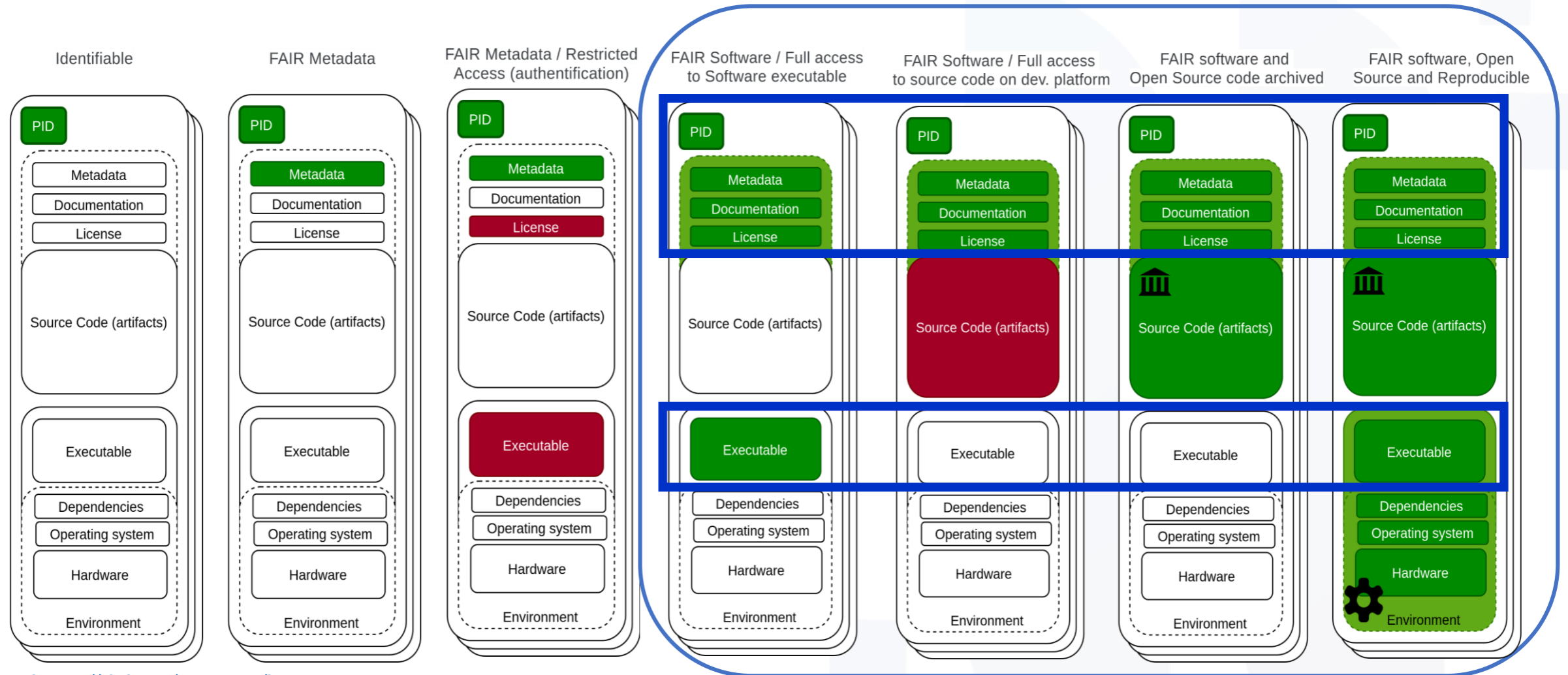
<< Les principes FAIR ont été initialement pensés pour les données de la recherche. Cependant, dans la plupart des recherches qui produisent des données, les logiciels et les codes sont des résultats de la recherche aussi importants que les données. Il est donc important d'adapter les principes FAIR aux logiciels dans l'objectif d'améliorer leur partage et leur réutilisation >>

Institut Pasteur, Comment rendre son logiciel FAIR ? 15 mars 2023

<https://openscience.pasteur.fr/2023/03/15/comment-rendre-son-logiciel-fair/>



Des logiciels plus ou moins « FAIR »



Points indispensables pour un logiciel « FAIR »

- Identifiant pérenne (PID)
- Indexation et référencement
- Métadonnées
- Documentation
- Licence ouverte et claire
- Code source
- Exécutable

Retour d'expérience suite à la FAIRisation de 2 logiciels

Projet FAIR-IMPACT



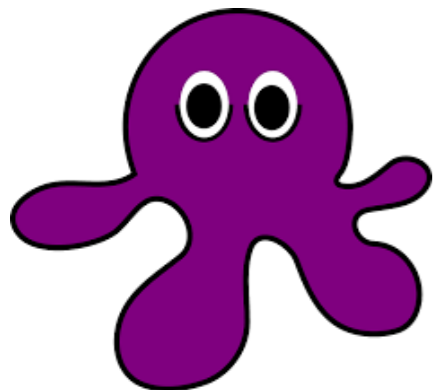
Appel à candidatures en avril 2024 relatif à la FAIRisation des logiciels de recherche (3 mois : de juin à août 2024)

> 2 logiciels retenus :

- OCTOPUS
- SCOOP3-ARGO



Logiciels OCTOPUS et SCOOP3-ARGO



OCTOPUS : Outil de vérification, de division et de conversion multiformat pour les formats SeaDataNet. C'est un outil unique et ergonomique pour de nombreuses conversions de formats.



Infrastructure pan-européenne pour la gestion des données marines et océanographiques



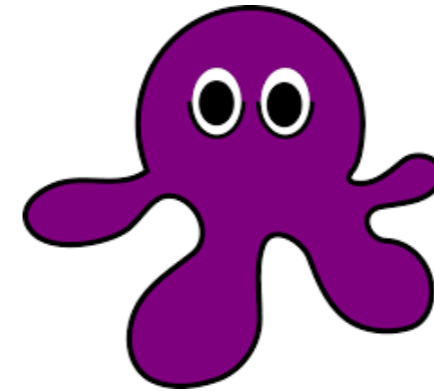
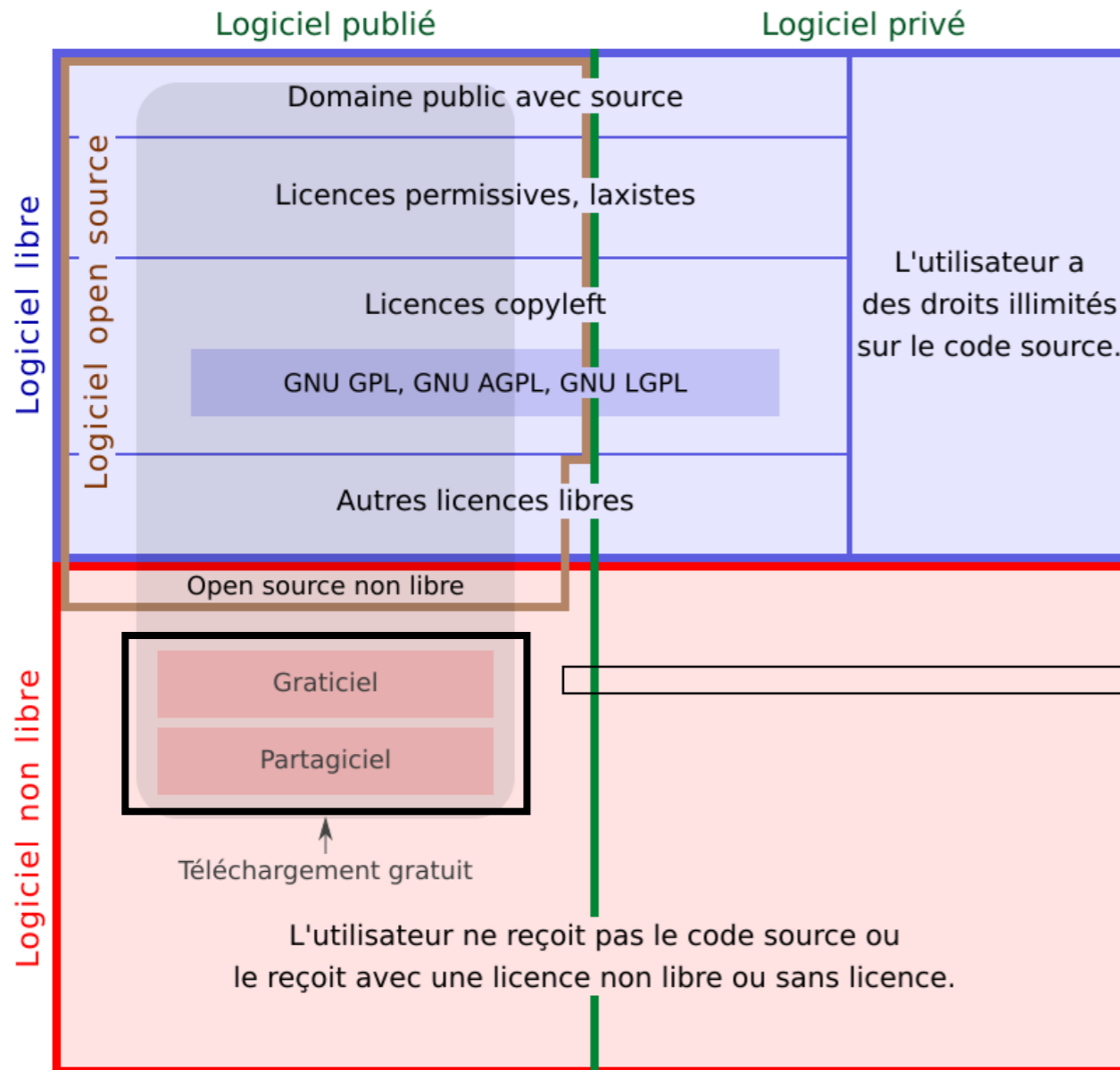
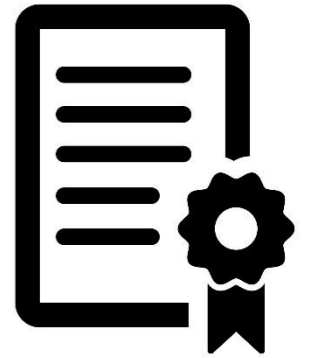
SCOOP3-ARGO : Outil facilitant la visualisation et le contrôle de la qualité des fichiers de données netCDF des flotteurs Argo



Infrastructure de recherche européenne contribuant au programme Argo (international)

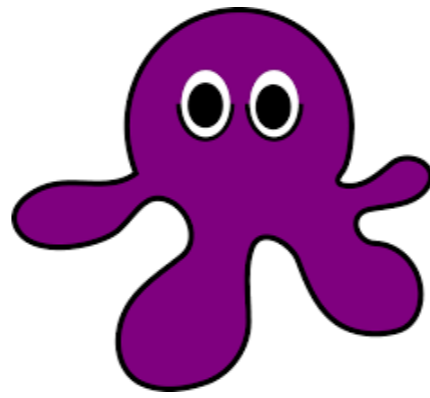


Etat initial (avant FAIRisation)



- Téléchargements libres et gratuits
- Codes sources non accessible (gitlab privé)
- Licences :
 - OCTOPUS : licence SeaDataNet 1.0 spécifique et décrite sur le [site SeaDataNet](#) (usages commerciaux et militaires non autorisés)
 - SCOOP3-ARGO : licence CC-BY

Zoom sur OCTOPUS



- Dépôt gitlab privé (gitlab.ifremer.fr)
- [Page de téléchargement et de présentation sur le portail SeaDataNet](#)
- Licence spécifique (SeaDataNet licence 1.0) interdisant les réutilisations commerciales et militaires
- Manuel utilisateur en anglais et à jour





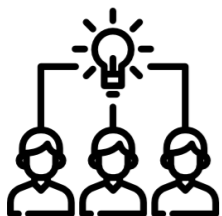
Focus initial sur 4 (sous-)métriques



FRSM-04-F2 (F2-1, F2-2, F2-3)

Software requires descriptive metadata to support indexing, search and discoverability.

métadonnées



FRSM-06-F2 (F2-1, F2-2, F2-3)

Software should make it easy to recognize and credit all contributors.

métadonnées



FRSM-13-R1 (R-1, R-2)

Software is made more reusable by providing suitable machine-actionable information on dependencies, build and configuration.

métadonnées



FRSM-15-R1.1 (R1-1-1, R1-1-2, R1-1-3)

Clear software licensing enables reuse.

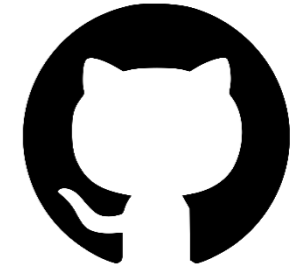
licence



Principales actions liées à la FAIRisation (1/2)



- Création d'une **organisation** et d'un **entrepôt** (github.com)



<https://github.com/seadatanet/octopus/>

- Ajouts des fichiers suivants :

- **README.md** contenant différentes parties


- **CITATION.cff** créé via <https://citation-file-format.github.io/>

- **LICENCE.txt** (après choix d'une licence ouverte)



Software Package Data eXchange



- Format de fichier > documenter informations sur les licences de logiciels
- Groupe de travail qui le rédige, composé de > 20 organisations différentes, sous les auspices de la fondation Linux
- Devenu en août 2021 une norme  : ISO/CEI 5962:2021
- Dernière liste : version 3.26.0 (30/12/2024) comprenant :
 - identifiant court normalisé + nom complet
 - URI (URL permanente) > texte de la licence
 - indication si validé par FSF et OSI



- Choix de la licence 
Free as in Freedom



Principales actions liées à la FAIRisation (2/2)



- **Référencement sur Software Heritage**

https://archive.softwareheritage.org/browse/origin/directory/?origin_url=https://github.com/seadatanet/octopus

Initiative mondiale à but non lucratif (lancée par l'INRIA) qui vise à collecter, préserver et partager l'ensemble du patrimoine logiciel mondial



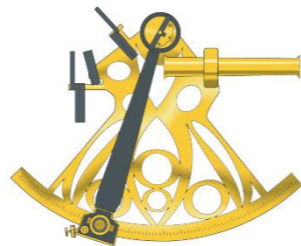
Archive universelle contenant tout type de logiciel libre ou open source

- **Création d'un fichier `codemeta.json`**

<https://codemeta.github.io/codemeta-generator/>

schéma de métadonnées minimal avec indexation par Software Heritage

- **Création d'un PID sur Sextant via Datacite**



DOI 10.12770/097e42fc-1aeb-4683-a781-35d92fa307f0





OCTOPUS Software (multiformat splitter & converter tool for SeaDataNet format)

Afficher Retour

Octopus is a multiformat splitter & converter tool for SeaDataNet format. It is a unique and ergonomic tool for many format conversions. The following conversions are possible with Octopus:

- Non SDN MedAtlas to SDN MedAtlas
- Non SDN MedAtlas to SDN ODV
- Non SDN MedAtlas to SDN NetCDF
- SDN MedAtlas to SDN ODV
- SDN MedAtlas to SDN NetCDF
- SDN ODV to SDN NetCDF



Accès aux données



DOI 10.12770/097e42fc-1aeb-4683-a781-35d92fa307f0

Date de la donnée 17-10-2024 (Révision)
2015-07 (Publication)

Auteur(s) Piel Steven (IFREMER) ID
Bregent Sophie
Crouzille Sébastien
Fachero Vincent (IFREMER)
Fichaut Michele (IFREMER) ID
Gatti Julie (IFREMER) ID
Le Hingrat Françoise (IFREMER)
Le Lonquer Leo
Mahoudo Pierre
Pertuisot Cecile (IFREMER) ID
Tosello Vanessa (IFREMER) ID

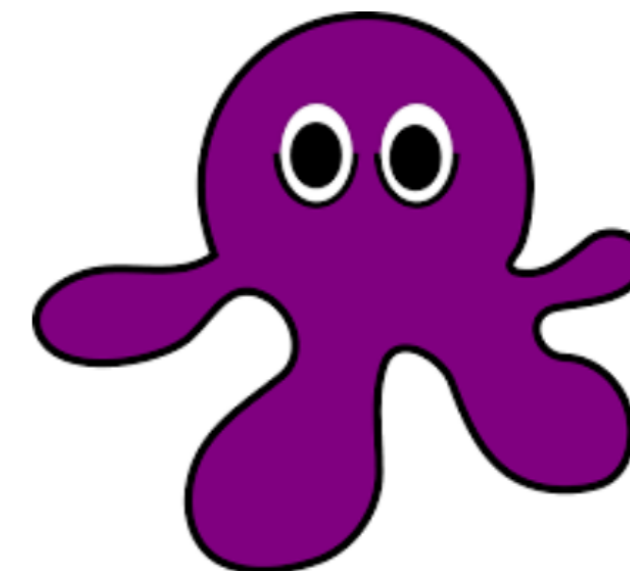
Contact(s) (helpdesk)
(ifremer, Scientific Information Systems for the sea)

Source SeaDataNet

Généalogie This softward was developed by the SeaDataNet project
Fundrs : Ifremer, European Commission under the Horizon 2020 (SeaDataNet projects: SeaDataNet, SeaDataNet 2, SeaDataCloud)

Contraintes Limitation d'utilisation : GNU Lesser General Public License v3.0 only : <https://spdx.org/licenses/LGPL-3.0-only.html>
Contraintes d'utilisation : Licence

Distributions



Available conversions

input output	MedAtlas	ODV	ODV variants	netCDF (CF)
MedAtlas non SDN	✓	✓	✗	✓
MedAtlas SDN	✓	✓	✗	✓
ODV SDN	✗	✓	✗	✓
ODV variants	✗	✗	✓	✗
netCDF SDN (Cfpoint)	✗	✓	✗	✓
MGDv81	✗	✓	✗	✗



OCTOPUS Software (multiformat splitter & converter tool for SeaDataNet format)

Afficher Retour

Octopus is a multiformat splitter & converter tool for SeaDataNet format. It is a unique and ergonomic tool. The following conversions are possible with Octopus:

- Non SDN MedAtlas to SDN MedAtlas
- Non SDN MedAtlas to SDN ODV
- Non SDN MedAtlas to SDN NetCDF
- SDN MedAtlas to SDN ODV
- SDN MedAtlas to SDN NetCDF
- SDN ODV to SDN NetCDF

doi PID sous la forme d'un DOI

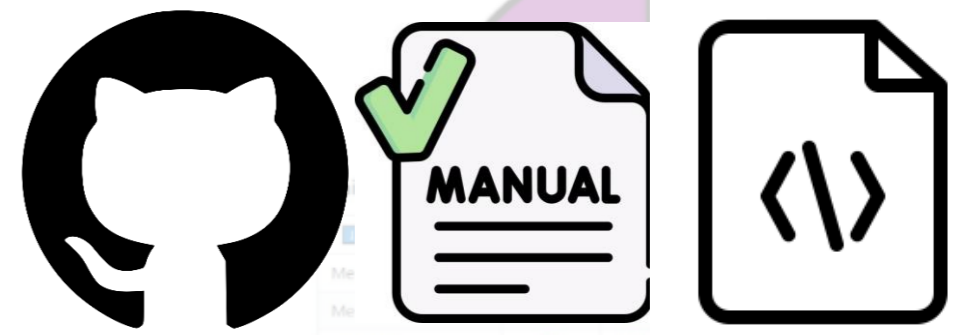
  Télécharger

DOI 10.12770/097e42fc-1aeb-4683-a781-35d92fa307f0

Page de téléchargement des fichiers compilés sur le site SeaDataNet

- GitHub Octopus
- User manual
- Dépôt Software heritage

Dépôt Github (public)
Manuel utilisateur (en anglais)
Dépôt du code source sur Software heritage



Contact(s) (helpdesk) (ifremer, Scientific Information Systems for the sea)

Source SeaDataNet

Généalogie This software was developed by the SeaDataNet project
Funders : Ifremer, European Commission under the Horizon 2020 (SeaDataNet projects: SeaDataNet, SeaDataNet 2, SeaDataCloud)

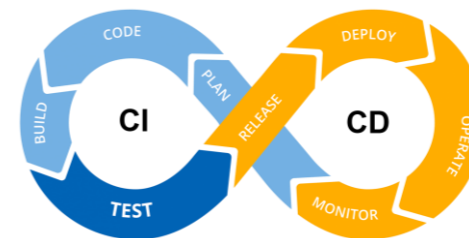
Contraintes Limitation d'utilisation : GNU Lesser General Public License v3.0 only : https://spdx.org/licenses/LGPL-3.0-only.html

Contraintes d'utilisation : Licence




Dépôts synchronisés



Mise en place d'une synchronisation régulière des dépôts
méthode « push » de Gitlab vers Github (chaîne CI-CD existante)



Mirroring repositories

Set up your project to automatically push and/or pull changes to/from another repository. Branches, tags, and commits will be synced automatically. [How do I mirror repositories?](#)

Collapse

Repository	Direction	Last update attempt	Last successful update	
https://*****.*****@github.com/seadatanet/octopus	Push	1 week ago	1 week ago	 



Evaluation FAIR



Evaluation FAIR pour un logiciel

Outil expérimental (scraper) basé sur F-UJI adapté aux logiciels

Summary:

6.67%		Score earned:	Fair level:
	Findable:	0 of 20	incomplete
	Accessible:	0 of 2	incomplete
	Interoperable:	0 of 7	incomplete
Reusable:	3 of 16	initial	

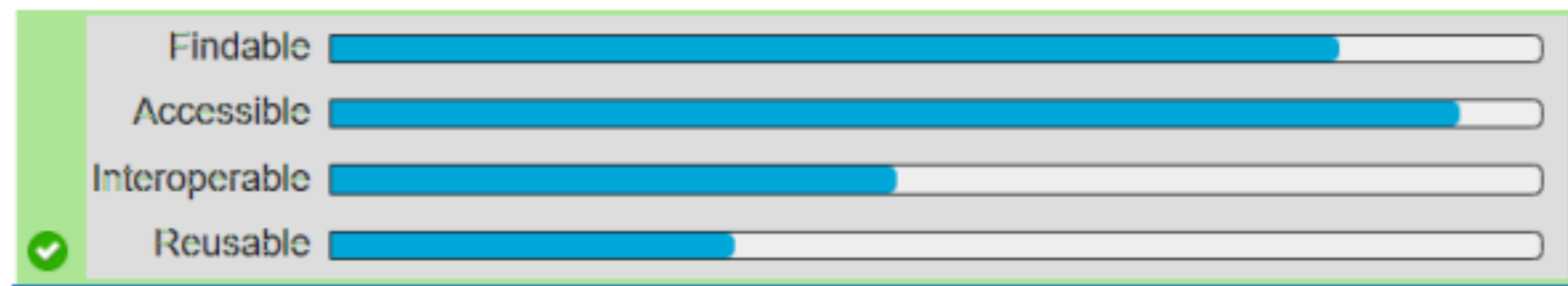


Auto-évaluation FAIR : <https://fairsoftwarechecklist.net/v0.2/>

Balise visible sur github




<https://fairsoftwarechecklist.net/v0.2?f=21&a=32112&i=31101&r=021>



Résumé et perspectives

Points clés à retenir

- Préalable indispensable : **dépôt public**
- Complétion de la **documentation**
 - > readme.md, contributing.md, build, changelog.md...
 - > captures d'écran, jeux de données test...
 - > utilisation de vocabulaires contrôlés : codes EDMO/ROR, ORCID...
- **Métadonnées** rajoutées : codemeta.json, pom.xml, licence.
- **Citation** des logiciels améliorée 
- Licence clairement indiquée (licence.txt) et choix d'une **licence ouverte**
- Archivage à long terme du **code source** sur Software Heritage
- Création d'un **identifiant pérenne** (PID)



Perspectives

- Elargissement courant 2025 à d'autres logiciels de l'infrastructure SeaDataNet



- Bonnes pratiques FAIR à généraliser à tous types d'applications, scripts, algorithmes

Logiciels FAIR : un atout pour les données !



- **Interopérabilité** favorisent des formats standardisés pour les données
- **Traçabilité** meilleure documentation des flux de données et de métadonnées
- **Réutilisabilité** garantissent un accès pérenne aux données traitées
- **Accessibilité** facilite l'accès et le partage des données via des interfaces adaptées
- **Standardisation** encourage l'usage de protocoles et standards FAIR pour les données
- **Automatisation** permet l'intégration dans des workflows automatisés et reproductibles
- **Transparence** facilite la compréhension des données via le code ouvert
- **Ethique** renforce la gestion responsable et ouverte des données

Merci pour
votre attention !

