



DATA
TERRA



ODATIS

Gestion FAIR des données bioinformatiques en biologie marine

Atelier technique Données bioinformatiques de diversité



11/03/2024

contact@odatis-ocean.fr | www.odatis-ocean.fr

F
Findable



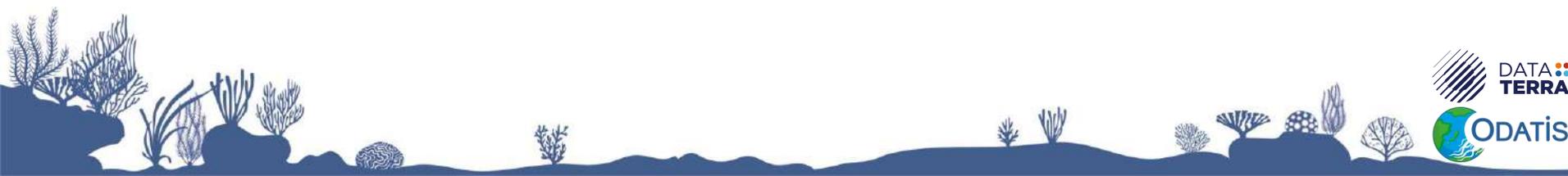
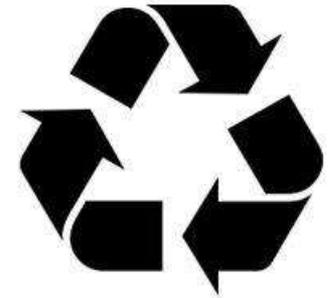
A
Accessible



I
Interoperable

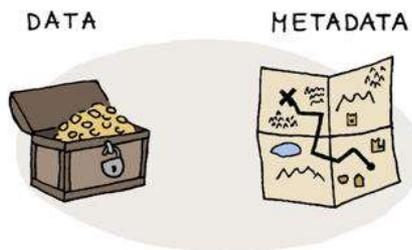


R
Reusable



Findable

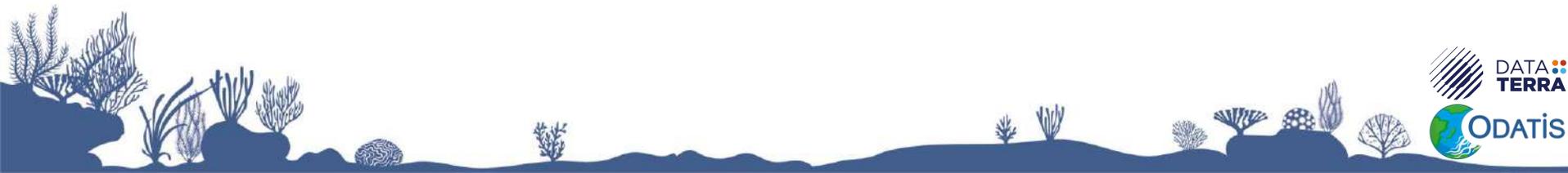
- Les métadonnées se voient attribuer un identifiant **unique** et **persistant** à l'échelle mondiale.



Dataedo /cartoon

Piotr@Dataedo

- Les données sont décrites avec des métadonnées **riches** (+ liens vers publications, données rattachées, etc)

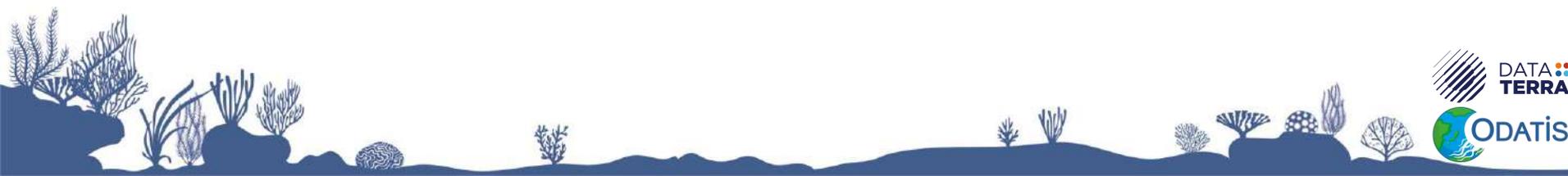


Findable

- Les métadonnées se voient attribuer un identifiant **unique** et **persistant** à l'échelle mondiale.

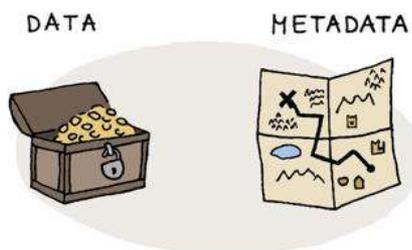


A screenshot of the SEANOE (SEA SCIENTIFIC OPEN ACCESS NETWORK ENVIRONMENT) search results page. The page features a search bar at the top with the query "phytoplankton" and a "Search everywhere" dropdown. Below the search bar, there are filter options for "Reset filters", "54 result(s)", and "Order by Relevance". The left sidebar contains filters for "PUBLICATION YEAR" (2024, 2023, 2022, 2021), "DISCIPLINE" (Biological oceanography, Chemical oceanography, Environment, Physical oceanography), "ACCESSIBILITY" (Open access, On demand), and "LICENSE" (CC-BY, CC-BY-NC, CC-BY-NC-SA, CC-BY-SA). The main content area displays three search results. The first result is titled "Effect of temperature on the unimodal size scaling of phytoplankton growth" by Fernández-González Cristina and Merañón Emilio, dated 2020-11-12. The second result is titled "Realised niche and suitability index of toxic phytoplankton species. Dataset" by Guillot Carles, Chappelle Annie, and Bechar Cedric, dated 2020-09-03. The third result is titled "PHYTOBS dataset - French National Service of Observation for Phytoplankton in coastal waters" and is dated 2023. Each result includes a brief abstract and a small thumbnail image.



Findable

- Les métadonnées se voient attribuer un identifiant **unique** et **persistant** à l'échelle mondiale.



Dataedo /cartoon

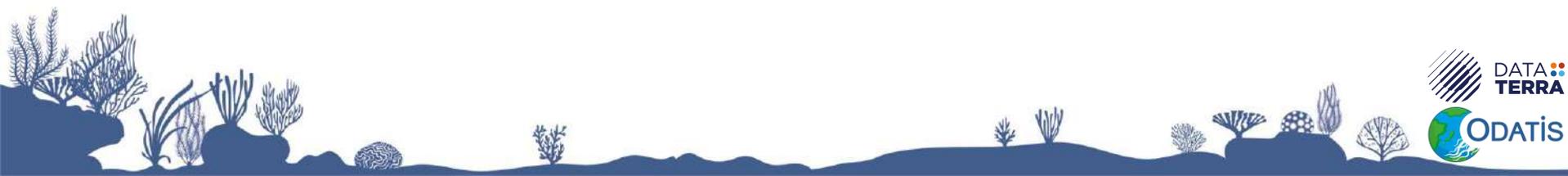
Piotr@Dataedo

- Les données sont décrites avec des métadonnées **riches** (+ liens vers publications, données rattachées, etc)



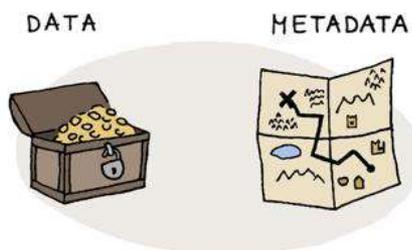
Piotr@Dataedo

- Les métadonnées sont enregistrées ou indexées dans une ressource **consultable**.



Findable

- Les métadonnées se voient attribuer un identifiant **unique** et **persistant** à l'échelle mondiale.



Dataedo /cartoon

Piotr@Dataedo

- Les données sont décrites avec des métadonnées **riches** (+ liens vers publications, données rattachées, etc)

- Les métadonnées sont enregistrées ou indexées dans une ressource **consultable**.



Piotr@Dataedo

Pourquoi pas?

Valoriser ses données dans un **data paper**

Findable

- Utiliser des vocabulaires contrôlés « machine readable »
 - Méditerranée => <http://marineregions.org/mrgid/1905>



Marineregions.org
Towards a standard for geographical marine names

About Gazetteer Maritime Boundaries Sources Statistics Downloads

Marine Gazetteer Placedetails

Search

Browse

About

Tutorial

Webservices

Login

MrgID: <http://marineregions.org/mrgid/1905>

Status: Proposed standard

Name	Language	Name source
Mediterranean Sea	English	(1953). Limits of oceans and seas, 3rd edition, IHO Special Publication, 23. International Hydrographic Organization (IHO): Monaco, 38 pp. (look up in IHO)

PlaceType: IHO Sea Area

Latitude: 38° 1' 25" N (38.02360535°)

Longitude: 15° 5' 32.3" E (15.09230423°)

Projection: 2012629 meter

Min. Lat: 30° 15' 50" N (30.2639°)

Min. Long: 0° 1' 57.6" W (-0.0327°)

Max. Lat: 43° 47' 0" N (43.7833°)

Max. Long: 36° 13' 2.2" E (36.2173°)

Source: (1953). Limits of oceans and seas, 3rd edition, IHO Special Publication, 23. International Hydrographic Organization (IHO): Monaco, 38 pp. (look up in [IHO](#))

Notes: [General information \(en\)](#): Based on IHO 23-3rd: Limits of Oceans and Seas, Special Publication 23, 3rd Edition 1953, published by the International Hydrographic Organization

Relation: Part of [Mediterranean Sea Area \(General Sea Area\)](#) ([View hierarchy](#))

Map

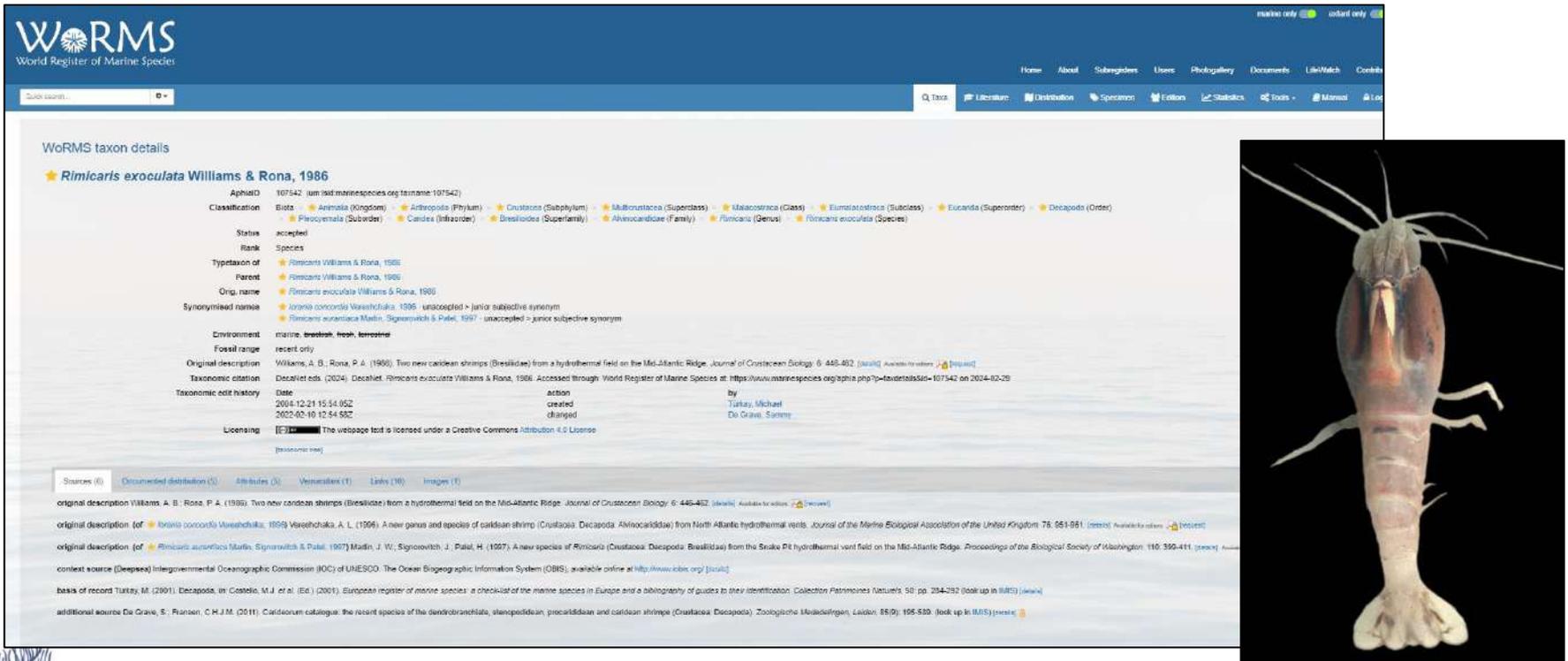
lat: 40.8 lon: -8.03

Download Layer: [MarineRegionsIHO](#) - format: [GML3](#) [Download](#)

[Shapefile](#) [Download](#) or [view the complete IHO Sea Area shapefile](#)

Findable

- Utiliser des vocabulaires contrôlés « machine readable »
 - Méditerranée => <http://marineregions.org/mrgid/1905>
 - *Rimicaris exoculata* => <https://www.marinespecies.org/aphia.php?p=taxdetails&id=107542>



The image shows a screenshot of the WoRMS (World Register of Marine Species) website. The page displays the taxon details for *Rimicaris exoculata* Williams & Rona, 1986. The details include classification (Aphelioda, Biotis, Arthropoda, Crustacea, Malacostraca, Eumalacostraca, Eucarida, Decapoda), status (accepted), rank (Species), type taxon of (Williams & Rona, 1986), parent (Williams & Rona, 1986), original name (Rimicaris exoculata Williams & Rona, 1986), synonymised names (Isonia concordia Vasechchka, 1989; Rimicaris aurantiaca Martin, Signeuvilch & Patel, 1997), environment (marine, benthic, fresh, boreotem), fossil range (recent only), original description (Williams, A. B., Rona, P. A. (1986). Two new caridean shrimps (Biotidae) from a hydrothermal field on the Mid-Atlantic Ridge. Journal of Crustacean Biology, 6: 445-462.), taxonomic citation (Decalmet, (2024) Decalmet, Rimicaris exoculata Williams & Rona, 1986. Accessed through World Register of Marine Species at: https://www.marinespecies.org/aphia.php?p=taxdetails&id=107542 on 2024-02-29), taxonomic edit history (Date: 2004-12-21 15:54:05Z, action: created, by: Turley, Michael; 2022-02-10 12:54:58Z, action: changed, by: De Grave, Sammy), and licensing (CC BY). A photograph of the shrimp is shown on the right side of the page.

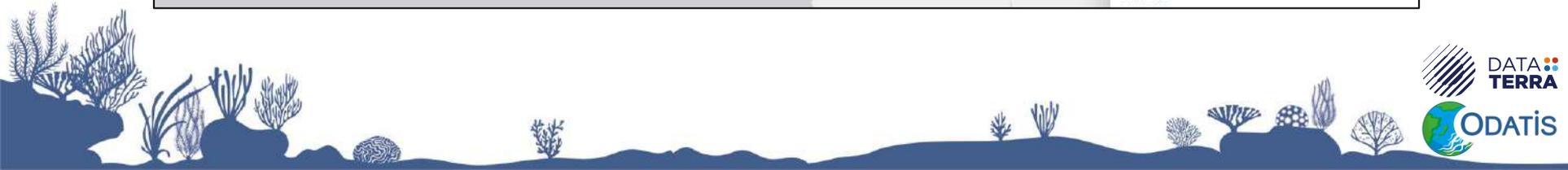


Findable

- *Rimicaris exoculata* => <https://www.marinespecies.org/aphia.php?p=taxdetails&id=107542>

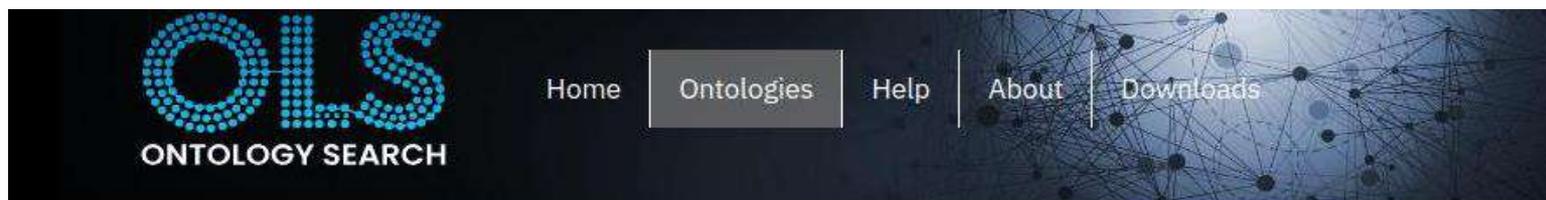
The screenshot displays the MSIS interface. On the left, a map of the Atlantic Ocean shows several colored squares representing data points. On the right, a sidebar contains the following sections:

- Create layer** 214 records Save
- Scientific name**
Enter scientific name
- | Scientific name | Taxon ID |
|--|----------|
| <i>Rimicaris exoculata</i> Williams & Rona, 1985 | 107542 |
- Datasets**
- Nodes**
- Institutes**
- Areas**
Enter area name
- | Area | ID |
|--------------------|----|
| No areas selected. | |
- Publisher countries**
- Geometry**
- Time range**
- Depth range**
- Extensions**
- Options**
- Styling**
- Color scale**



Findable

- Utiliser des vocabulaires contrôlés « machine readable »
 - Méditerranée => <http://marineregions.org/mrgid/1905>
 - *Rimicaris exoculata* => <https://www.marinespecies.org/aphia.php?p=taxdetails&id=107542>
 - Ecosystème méditerranéen => http://purl.obolibrary.org/obo/ENVO_01000047



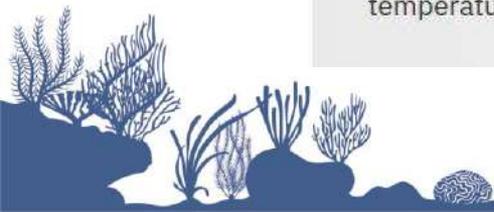
Ontologies > ENVO > Classes > ENVO:01000047  Copy



mediterranean sea biome

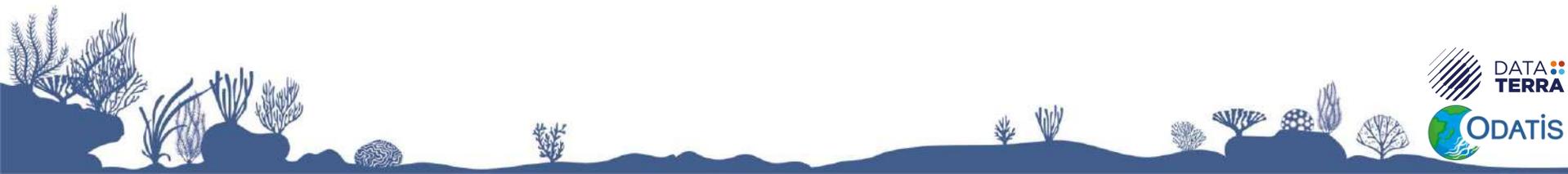
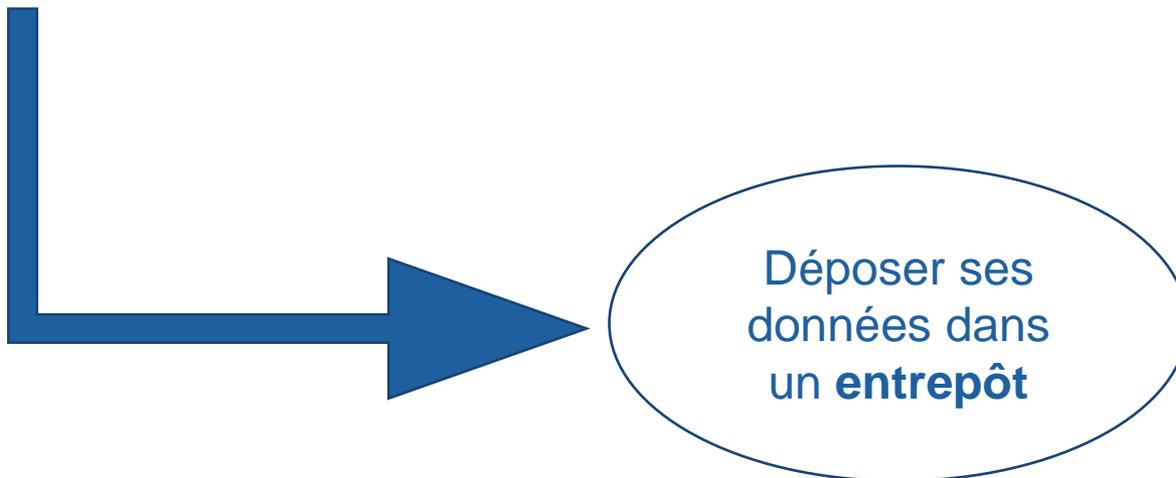
 http://purl.obolibrary.org/obo/ENVO_01000047  Copy

The mediterranean sea biome comprises mostly enclosed seas that have limited exchange of deep water with outer oceans and where the water circulation is dominated by salinity and temperature differences rather than winds. 



Accessible

- Les métadonnées doivent être accessibles même si les données ne sont **pas ou plus disponibles**
- Les métadonnées et données doivent être accessibles via des **protocoles d'accès standardisés**



Accessible

- Où déposer ses données?

- Entrepôt générique



- Entrepôt thématique international



- Entrepôt institutionnel et thématique national



Accessible

- Comment choisir un entrepôt?

- Entrepôt faisant consensus parmi les chercheurs du domaine



- Recommandations de journaux scientifiques

- Bonnes pratiques des collègues

- Registres d'entrepôts



- Entrepôt digne de confiance : certification, reconnu par la communauté



Interoperable

- Utiliser des **standards** pour les données et les métadonnées : Darwin Core, EML, ISO 19115, ...

Darwin core

- DNA derived data extension
- Pour les occurrences obtenues à partir d'échantillonnage d'ADN

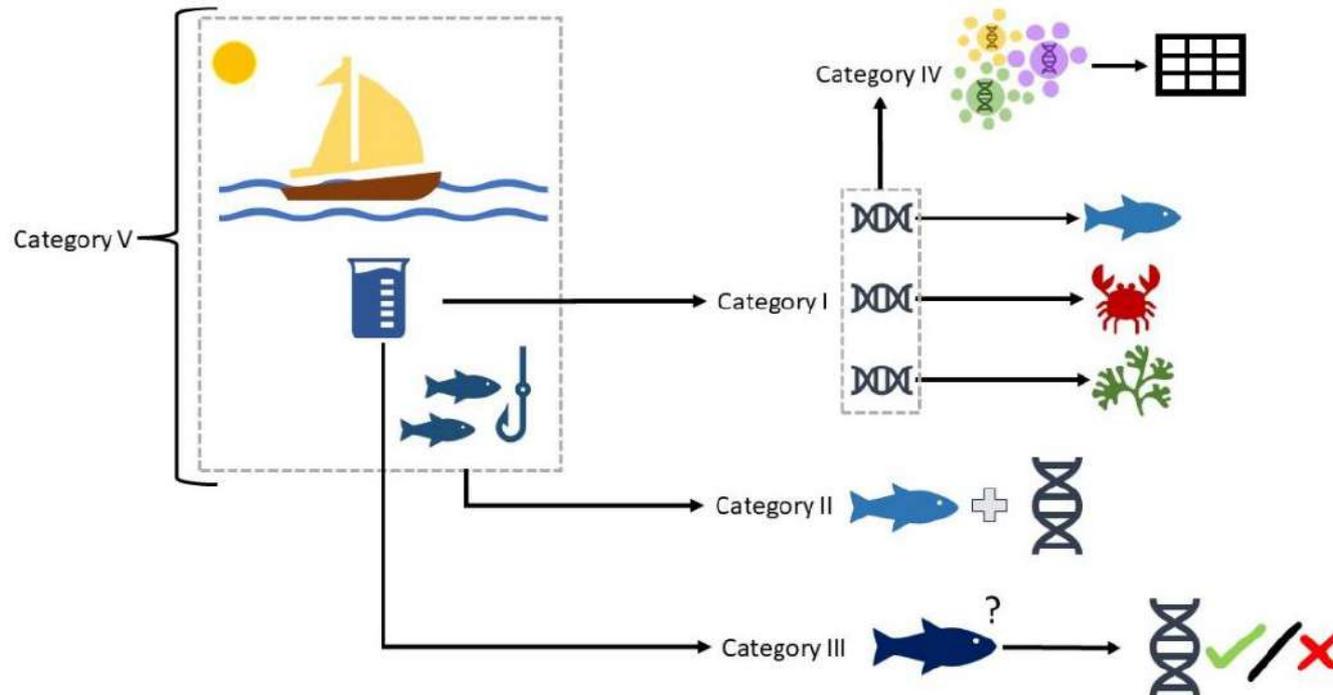


Figure 6 in Publishing DNA-derived data guide (Abarenkov et al. 2023 <https://doi.org/10.35035/doc-vf1a-nr22>).

Pour en savoir plus :

https://rs.gbif.org/extension/gbif/1.0/dna_derived_data_2021-07-05.xml

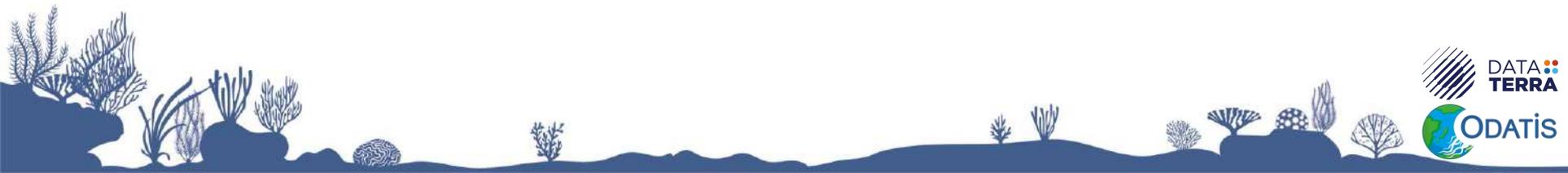
<https://docs.gbif-uat.org/publishing-dna-derived-data/en/>

Interoperable



- Utiliser des formats standards pour les **valeurs** : dates, unités, ...
 - Créer des **liens** entre fiches de métadonnées

Title	Sequencing raw data from genome of hydrothermal shrimp <i>R. exoculata</i>
DOI	http://dx.doi.org/10.12770/e2ae7d37-cee5-4a00-b9c2-96e00099ca00
Auteurs	Cambon Marie-Anne (ORCID id 0000-0002-4076-0472) Cormier Alexandre (ORCID id 0000-0002-7775-8413) Aube Johanne (ORCID id 0000-0001-9677-2950) Cueff-Gauchard Valerie (ORCID id 0000-0001-6969-642X)
Institut	IFREMER (EDMO 752)
Taxonomie	Rimicaris exoculata (WORMS 107542)
Plongée	HER2-PL2049-07
Licence	CC-BY-NC-ND



Interoperable



- Utiliser des formats standards pour les **valeurs** : dates, unités
- Créer des **liens** entre fiches de métadonnées

Title	Sequencing raw data from genome of hydrothermal shrimp <i>R. exoculata</i>
DOI	http://dx.doi.org/10.12770/e2ae7d37-cee5-4a00-b9c2-96e00099ca00
Auteurs	Cambon Marie-Anne (ORCID id 0000-0002-4076-0472) Cormier Alexandre (ORCID id 0000-0002-7775-8413) Aube Johanne (ORCID id 0000-0001-9677-2950) Cueff-Gauchard Valerie (ORCID id 0000-0001-6969-642X)
Institut	IFREMER (EDMO 752)
Taxonomie	Rimicaris exoculata (WORMS 107542)
Plongée	HER2-PL2049-07
Licence	CC-BY-NC-ND

Activities

- > Employment (1)
- Works (50 of 119)

Items per page: 50 Page 1 of 3

Distinct development trajectories and symbiosis modes in vent shrimps

Evolution
2024-03 | Journal article
DOI: [10.1093/evolut/tpad217](#)
SOURCE-WORK-ID: 97798
CONTRIBUTORS: Pierre Methou; Guéganton Marion; Copley Jonathan T; Watanabe Hiromi Kayama; Florence Pradillon; Marie-Anne Cambon-Bonavita; Chong Chen

SeaDataNet FOR EUROPEAN INFRASTRUCTURES FOR OCEAN & MARINE DATA MANAGEMENT

EUROPEAN DIRECTORY OF MARINE ORGANISATIONS (EDMO)

ORGANISATION DETAILS

DETAILS

Name: Ifremer Head Office (ifremer)
 Postal name: Ifremer Siège
 Address: 1825 route de Sables-Moutiers CS 10070
 Zipcode: 29100
 City: PLOUZANES
 Country: France
 Phone: +3302 98 22 00 00
 Fax: +3302 98 22 45 45
 Email: edmo@ifremer.fr
 Centre Website: <http://edmo.ifremer.fr/>

ORGANISATION PROFILE

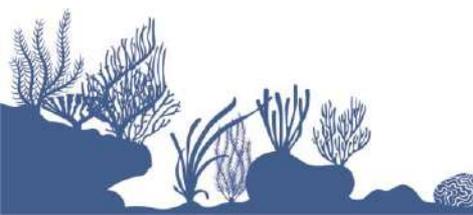
Through studies and expert assessments, Ifremer contributes to knowledge about the oceans and to resources the monitoring of marine and coastal zones and the sustainable development of maritime activities. For these goals, it designs and operates observational, experimental and monitoring tools and facilities. Ifremer manages the French ocean research fleet for the whole

GET DIRECTIONS



Sources (6) Documented distribution (5) Attributes (5) Vernaculars (1) Links (10)

- ★ To Barcode of Life (91 barcodes)
- ★ To Biodiversity Heritage Library (1 publication) (from synonym ★ *Rimicaris aurantiaca* Martin, Signorovitch & Pa...
- ★ To Biodiversity Heritage Library (5 publications)
- ★ To Biological Information System for Marine Life (BISMAL)
- ★ To European Nucleotide Archive (ENA)
- ★ To GenBank (340 nucleotides; 227 proteins)
- ★ To PESI
- ★ To USNM Invertebrate Zoology Arthropoda Collection (21 records)
- ★ To USNM Invertebrate Zoology Arthropoda Collection (5 records) (from synonym ★ *Rimicaris aurantiaca* Ma...
- ★ To USNM Invertebrate Zoology Arthropoda Collection (Holotype USNM 228443)



Interoperable



- Utiliser des formats standards pour les **valeurs** : dates, unités, ...
 - Créer des **liens** entre fiches de métadonnées

Title	Sequencing raw data from genome of hydrothermal shrimp <i>R. exoculata</i>
DOI	http://dx.doi.org/10.12770/e2ae7d37-cee5-4a00-b9c2-96e00099ca00
Auteurs	Cambon Marie-Anne (ORCID id 0000-0002-4076-0472) Cormier Alexandre (ORCID id 0000-0002-7775-8413) Aube Johanne (ORCID id 0000-0001-9677-2950) Cueff-Gauchard Valerie (ORCID id 0000-0001-6969-642X)
Institut	IFREMER (EDMO 752)
Taxonomie	Rimicaris exoculata (WORMS 107542)
Plongée	HER2-PL2049-07
Licence	CC-BY-NC-ND

HERMINE2 / Plongée : 2049-7

Plongée : 2049-7

DESCRIPTION DE LA PLONGÉE

Engin	Nautilie
Début de plongée	31/07/2022 - 11:24:05
Fin de plongée	31/07/2022 - 19:35:07

RAPPORT CHRONOLOGIQUE DE LA PLONGÉE

Date	Localité	Latitude	Longitude	Profondeur(m)	Commentaire
2022-07-31 15:42:42.0	-	23.3691227	-44.9509243	3497	HER2-PL07-ASPI-02 début 15:42:42 aspiration de Rimicaris exoculata adultes population 2
2022-07-31 16:17:26.0	-	23.3687582	-44.9519445	3467	null

Interoperable



- Utiliser des formats standards pour les **valeurs** : dates, unités, ...
 - Créer des **liens** entre fiches de métadonnées

Liste / Campagne n°18001851

HERMINE2

Type Campagne océanographique

Navire [Pourquoi pas ?](#)

Propriétaire navire Ifremer

Dates 08/07/2022 - 24/08/2022

Chef(s) de mission [PELLETER Ewan](#) , [CATHALOT Cécile](#) 

GEO-OCEAN - UMR 6538
IFREMER - Technopôle Brest-Iroise
IUEM - Rue Dumont d'Urville
29280 Plouzané
<https://www.geo-ocean.fr/>

DOI [10.17600/18001851](#)



Citer cette campagne



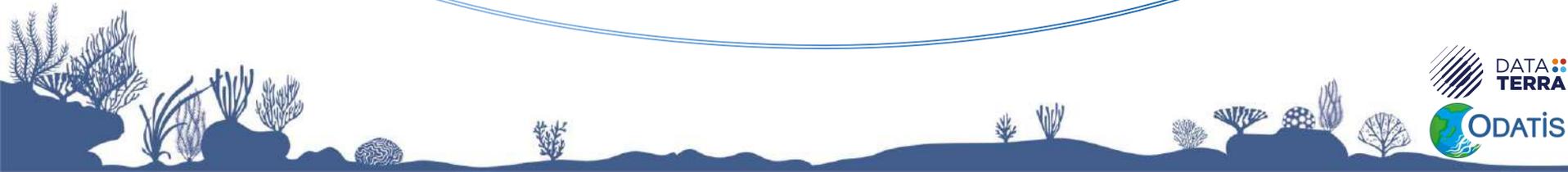
PELLETER Ewan, CATHALOT

HERMINE2 / Plongée : 2049-7

Plongée : 2049-7

DESCRIPTION DE LA PLONGÉE

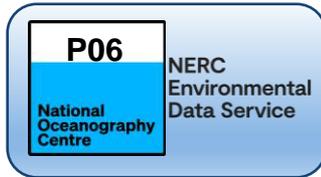
Engin	Nautilie
Début de plongée	31/07/2022 - 11:24:05
Fin de plongée	31/07/2022 - 19:35:07



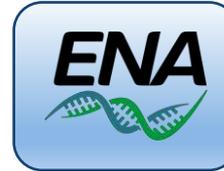
Interoperable

- Utiliser des **vocabulaires contrôlés FAIR** dans les métadonnées ET les données :

Unités



Jeux de données



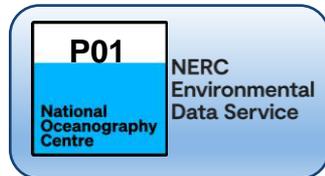
Systèmes biologiques



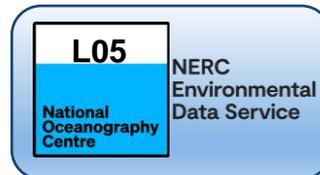
Taxonomie



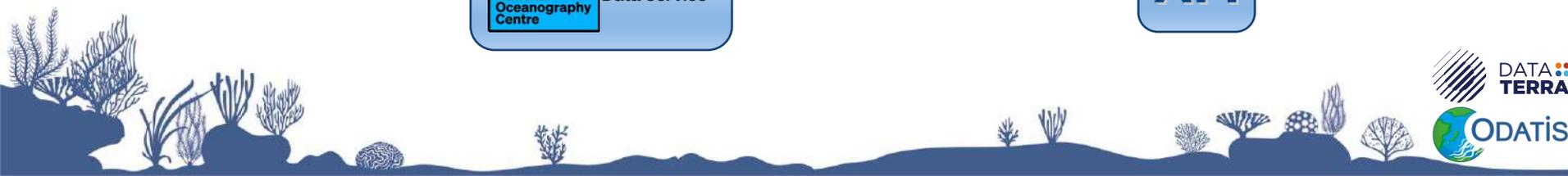
Paramètres



Instruments



Analyses



Interoperable

- Où trouver ces vocabulaires contrôlés?



- Répertoires de vocabulaires :
 - [BioPortal](#)
 - [EarthPortal](#)
 - [EMBL-EBI Ontology Lookup Service](#)
- [NERC Vocabulary Server](#)
- Pour les outils et analyses liées : [bio.tools](#)



Reusable

- Renseigner de la manière la plus compétè et exhaustive possible les métadonnées



Se mettre à la place du réutilisateur :
de quelles informations à t'il besoin?



Reusable

- Renseigner de la manière la plus compétè et exhaustive possible les métadonnées
- Expliciter la provenance des données diffusées

Où?

Comment?

Qui?

Quand?

Pourquoi?

Quoi?



Reusable

Comment?

- Pour des **données de séquençage** :
 - Sélection du fragment (PCR, RT-PCR...)
 - Stratégie de séquençage (WGS, Amplicon, Hi-C...)
 - Données pairées ou non
 - Quelle technologie (short reads / long reads)
 - Quel instrument (MiSeq, Minlon...)
 - Les primers utilisés

- Pour des **données analysées** :
 - Quels outils
 - Utilisation d'un workflow
 - Quelles versions
 - Quels paramètres



Reusable

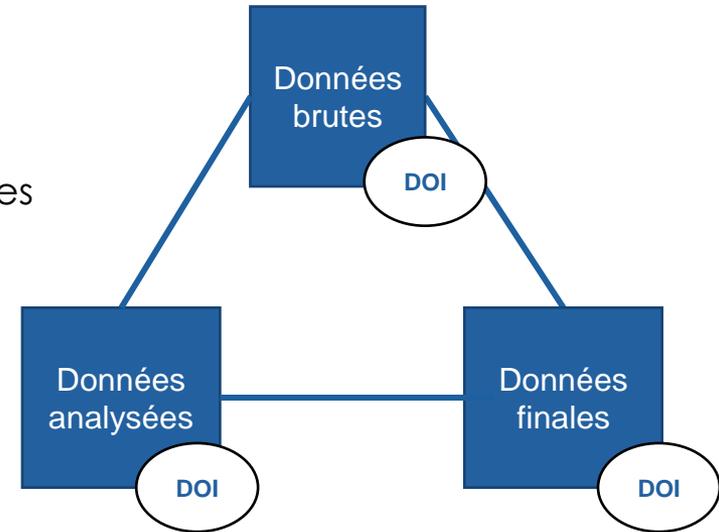
- Pour le **design expérimental** :
 - Quelles conditions ont été testées
 - Age des individus
 - Ploidie
- Pour **des données de séquençage** :
 - Quelle région a été séquencée
- Pour des **données analysées** :
 - Quelles étaient les données brutes

Quoi?



Reusable

- Expliciter la provenance des données diffusées
 - Diffuser et lier les données brutes, analysées et finales (décrites par des métadonnées différentes)



- Décrire l'organisation des données et les différents champs dans une norme dédiée (ISO 19110, EML) ou dans un .txt
- Ajouter aux métadonnées des fichiers complémentaires ou des liens vers des ressources externes sur le matériel utilisé, les scripts/logiciels, le protocole, ...

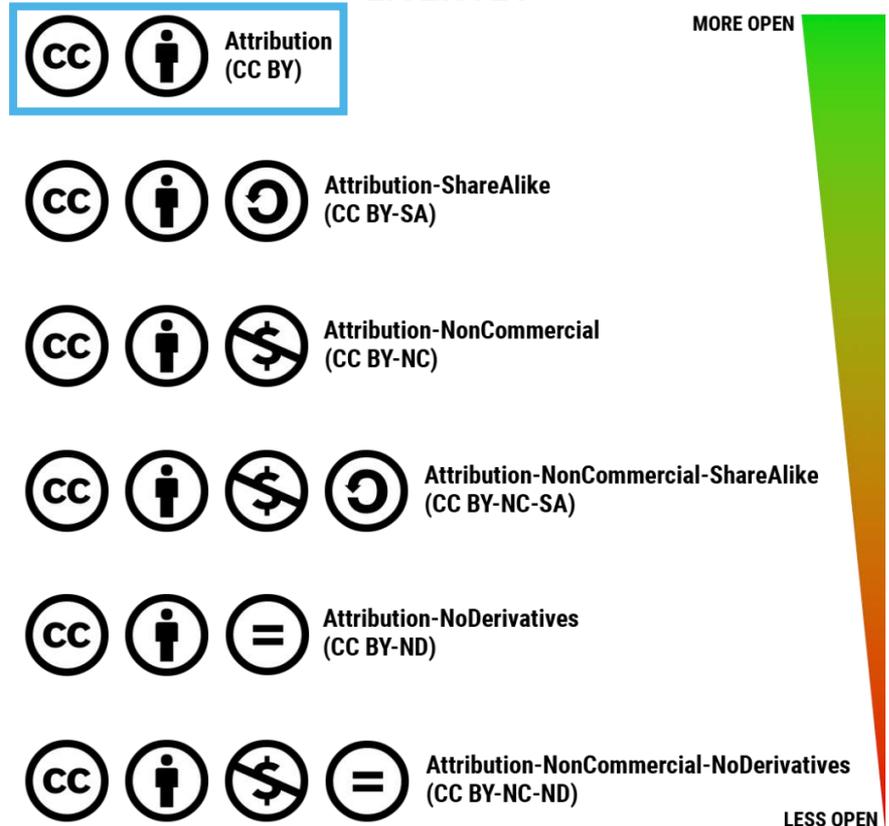


Reusable

- Indiquer la licence sur les données => pas de licence : Est-ce que les données sont ouvertes?



LICENCES



<https://au.creativecommons.net/>



As open as possible, as closed as necessary!

<https://creativecommons.org/licenses/by/4.0/>



DATA
TERRA



ODATIS

Merci pour votre attention

Laura Leroi laura.eroi@ifremer.fr

Clémence Cotten clemence.cotten@ifremer.fr



11/03/2024

contact@odatis-ocean.fr | www.odatis-ocean.fr