



DATA
TERRA



ODATIS

Comment la sémantique combinée à l'IA peut améliorer la découverte de données dans les catalogues?

Perspectives d'utilisation dans catalogue ODATIS

Clémence Cotten, Christelle Pierkot, Erwann Quimbert, Erwan Bodéré

Atelier IA ODATIS – 27 janvier 2026



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

Gestion des métadonnées : besoins pour l'amélioration de la découverte dans un catalogue

• Contexte

The image displays two side-by-side screenshots of data catalogs, likely from the FAIRsharing platform, demonstrating how semantic search can enhance discovery.

Screenshot 1 (Left): This catalog page shows a search results grid for "CATALOGUE". A red box highlights the left sidebar, which contains a "Variables essentielles Odatis" section. This section lists various types of variables such as "Aquatique de l'océan" (33), "Activités humaines" (55), "Aérosols" (20), "Bathymétrie" (202), and "Type de jeu de données" (10). Below this is a "Platesformes satellites" section listing "CDS-AVISO" (10), "CDS-CERAT" (10), "CDS-CORIOLIS" (10), "CDS-CASU" (10), "CDS-CBPF" (20), "Institutions de recherche" (10), "IR-AQSO" (10), "IR-EVIO-France" (10), "IR-PDF" (10), "IR-LCD" (10), and "Services natureur observation" (10).

Screenshot 2 (Right): This catalog page shows a search results grid for "Dernières mises à jour". A red box highlights the left sidebar, which contains a "Types de ressources" section. This section lists categories like "Application" (1), "Logiciel" (10), "Support & assistance" (2), "Cycle de vie de la donnée", "Langage de développement", "Variables essentielles Odatis", "Centres de données", "Infrastructures de recherche", and "Services nationaux observation".

Comment la sémantique combinée à l'IA peut améliorer la découverte de données dans les catalogues?

| 2

Gestion des métadonnées : besoins pour l'amélioration de la découverte dans un catalogue

- 1^{ère} idée : aide au déposant
 - Suggérer des mots-clés des thésaurus Odatis au déposant à partir du contenu du titre et du résumé d'un jeu de données ou ressource (comme implémenté dans EaSy Data)
- 2^{ème} idée : aide à l'administration
 - Outil dans l'administration du catalogue permettant de demander à l'IA d'ajouter des mots-clés sur un lot de fiches de manière automatique ou semi-automatique (vérification humaine avant validation)
- 3^{ème} idée : vers une nouvelle manière de penser le catalogue
 - Facettes réduites à des informations « de base » (période temporelle, zone géographique, ...)
 - Recherche textuelle améliorée par l'IA et chatbot en complément du catalogue

Comment la sémantique combinée à l'IA peut améliorer la découverte de données dans les catalogues?

| 3

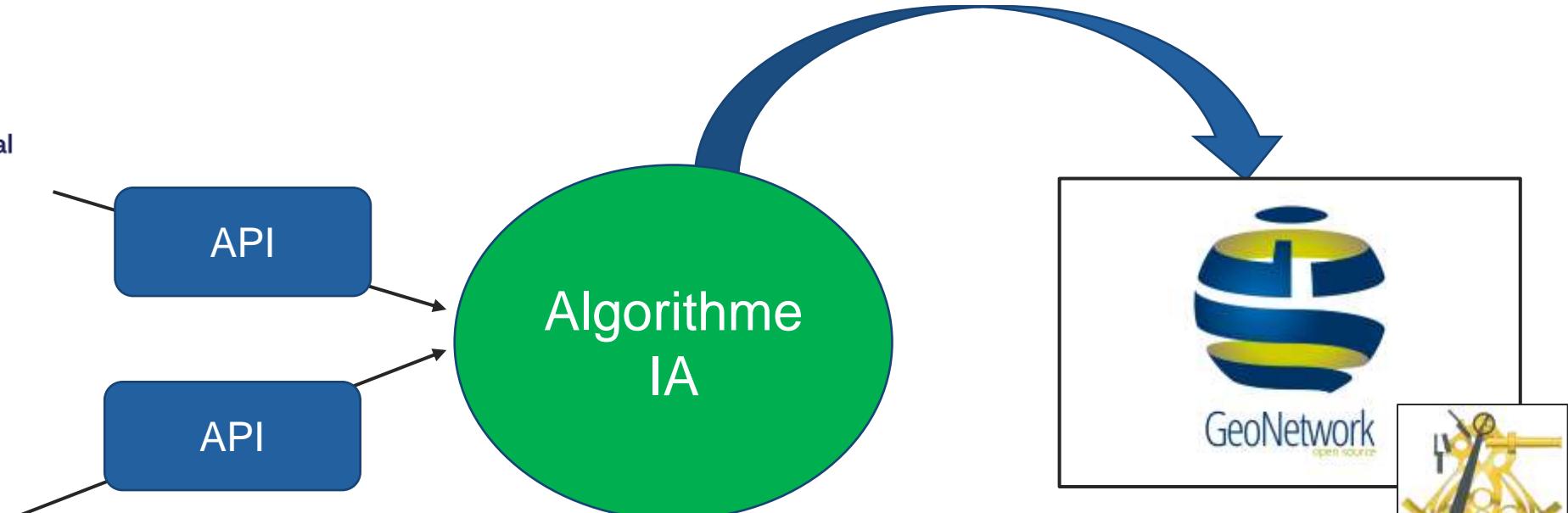
Gestion des métadonnées : besoins pour l'amélioration de la découverte dans un catalogue

- Lien sémantique et IA



NERC
Environmental
Data Service

Thésaurus Odatis



Comment la sémantique combinée à l'IA peut améliorer la découverte de données dans les catalogues?

| 4

Gestion des métadonnées : besoins pour l'amélioration de la découverte dans un catalogue

- Quelques tests effectués avec chatGPT : résultats probants

Acoustic dataset recorded during a technical trials campaign on the R/V L'Atalante, including numerous marine mammal sound emissions

DATE 2025-07-16
AUTHORS [Le Gall Yves](#)¹, [Ponchart Marie](#)¹, [Menut Eric](#)¹
AFFILIATIONS ¹ IFREMER
DOI [10.17882/107662](#)
PUBLISHER [SEANOE](#)

This dataset presents raw acoustic data in WAV format (sampling frequency : 256 kHz) recorded onboard R/V L'Atalante on March 9, 2025, between 12:30 and 16:30 UTC, off the West coast of Guadeloupe (Cruise ESSTECH-AT-2025-1: ship radiated noise measurement).

Ambient noise data recorded on an acoustic recorder submerged 40m below the surface include numerous sound emissions from marine mammals (sperm whales observed visually in the area, other odontocetes and probably distant humpback whales).

DISCIPLINES Physical oceanography
PARAMETERS [Acoustic noise in the water column](#)
DEVICES [hydrophones](#)
KEYWORDS Underwater acoustics, Ambient noise, Marine mammals, Sperm whales, Odontocetes, Mysticetes, Clicks, Whistles, Vocalizations
LOCATION 16.161921N, 16.037895S, -61.806335E, -62.023315W
LICENCE

Devices

Data were collected during ESSTECH-AT-2025-1 sea-trials using an ICLitsen HF-L digital hydrophone (usable frequency band : 1Hz - 200 kHz). The selected sampling frequency was 256 kHz four our acquisitions. This hydrophone was mounted on a drifting raft and deployed at a depth of 40 m.

The screenshot shows a dataset page with the following sections:

- Download Data**: Includes a map of the study area and download links for various file formats (TXT, RIS, XLS, RTF, BIBTEX).
- Termes OD1 recommandés**: A table listing recommended terms with their codes, labels, and justifications.

Code OD1	Label	Justification
ACOUST	Ocean Acoustics	Cœur du jeu de données : enregistrements sonores sous-marins, bruit ambiant, propagation acoustique
MTBM	Marine turtles, birds, mammals	Couvre les cétacés enregistrés (cachalots, odontocètes, mysticètes)

5

Gestion des métadonnées : besoins pour l'amélioration de la découverte dans un catalogue

- Quelques tests effectués avec chatGPT : résultats probants

Morpho-bathymétrie du substratum rocheux de la baie de Douarnenez (G. Jouet, 2005), raster

Morphobathymétrie du substratum rocheux de la baie de Douarnenez (G. Jouet, 2005), raster

Détails de la demande correspondant à la "Carte morpho-bathymétrique du substratum rocheux de la baie de Douarnenez", éditée en 2005 dans l'Atlas thématique de l'environnement marin de la baie de Douarnenez. Campagne géologique. À l'origine du jeu de données :

- AUGUSTE Claude, HERMANN Alain (2005) : CADMOU05r01a. RV Thémis, rapport interne 13 / 2005/01023. Références bibliographiques
- AUGUSTE C., HERMANN A., HANROT D., LELLOUCHE A., LE RISCH R., ROUET J., JOUET G., HERTIN S., DEBRUNY M., DROGO M., BERNARD E., MAILLARD X. (2005) : Atlas thématique de l'environnement marin de la baie de Douarnenez (Finistère). 10 cartes, échelle 1:20 000 et 1:50 000. Un inventaire géologique. TIRAGE 1000 exemplaires.
- Jouet G., AUGUSTE C., HERMANN A., Le Risch R., Rouet J. (2005) : La baie d'Iroise : un patrimoine hydrographique minéral en baie de Douarnenez (Finistère, France). Comptes Rendus Géoscience 325, 467-484.

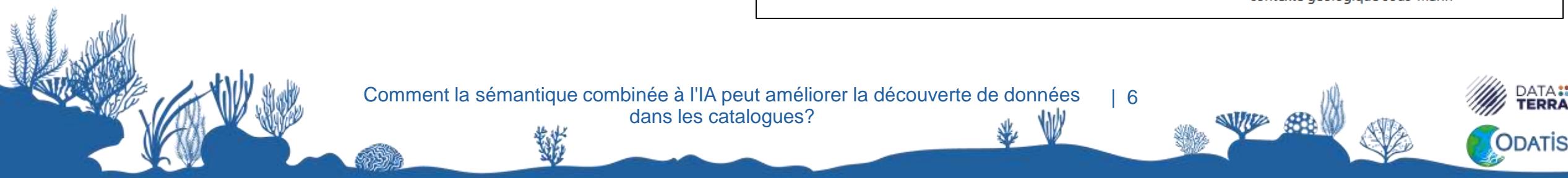
Date de la demande : 01-01-2020 (Chercher)
Auteur(s) : JOUET Gérard (France) ; VAIZE Jean-Pierre (France)
Géodonneur : Waternat (France) ; Waternat (France)
Source : Waternat
Contraintes :
Licences d'utilisation : CC-BY (Creative Commons - attribut)
Licences d'utilisation : Accès aux données à la demande giscom@finenat.fr
Contraintes d'utilisation : Licence
Autres contraintes : Pas de restriction d'accès public
Autres contraintes : Chiffre : 202878 - Substratum rocheux de la baie de Douarnenez (Finistère), thème
Informations géopgraphiques :
Détailles :
Visualiser :
Télécharger : Substratum_BaieDouarnenez_Jouet_2005_Substratum rocheux de la baie de Douarnenez

Acès aux données :



Termes OD1 recommandés

Code OD1	Libellé	Justification
BATHY	Bathymetry	Correspond exactement à la nature des données : modèle numérique du fond marin
GEOL	Geology	Géomorphologie, substratum rocheux = contexte géologique sous-marin



Comment la sémantique combinée à l'IA peut améliorer la découverte de données dans les catalogues? | 6

Gestion des métadonnées : besoins pour l'amélioration de la découverte dans un catalogue

• Quelques tests effectués avec chatGPT : résultats probants

Nutrients and phytoplankton pigment concentrations at the Porticciolo station (Corsica Channel, Corsica)

DATE 2025-05-12
TEMPORAL EXTENT 2019-03-25 – 2024-11-28
AUTHORS Goffart Anne¹
AFFILIATIONS 1. Oceanology, FOCUS Research Unit, University of Liège, Belgium
DOI [10.17882/106439](https://doi.org/10.17882/106439)
PUBLISHER SEANOE

This dataset contains time series (2019 – 2024) of monthly nutrient and HPLC phytoplankton pigment concentrations measured at the Porticciolo station, a coastal station situated in the Corsica Channel (Northern Tyrrhenian Sea) over the 5 m isobath. The analyzed nutrients (in uM) are nitrate (NO_3^-), nitrite (NO_2^-), ammonia (NH_4^+) and silicate (Si(OH)_4^-).
The measured pigments (in ng L^{-1}) include chlorophyll a (chl a), divinyl chlorophyll a (DVchl a), total chlorophyll a (Tchl a), peridinin (Peri), 19'-butanoyloxyfucoxanthin (Buta), fucoxanthin (Fuco), neoxanthin (Neo), prasinoxanthin (Prasino), violaxanthin (Viola), 19'-hexanoyloxyfucoxanthin (19'HF), alloxanthin (Allo), zeaxanthin (Zea), lutein (Lutein) and total chlorophyll b (Tchl b). Sampling and analysis methods are described in Goffart et al. (2015). These data contribute to characterize seasonal patterns of phytoplankton biomass and composition in a region where very limited information on phytoplankton dynamics is available. The data are also used to calibrate and validate ocean color remote sensing algorithms and to infer phytoplankton functional types using machine learning approaches.

DISCIPLINES Biological oceanography
PARAMETERS Ammonium and ammonia concentration parameters in water bodies, Carotenoid and flavonoid pigment concentrations in water bodies, Chlorophyll pigment concentrations in water bodies, Nitrate concentration parameters in the water column, Nitrite concentration parameters in the water column,

Download data
Download metadata
TXT RIS XLS RTF BIBTEX

Correspondance avec OD1

Code OD1	Libellé	Justification
NUTS	Nutrients	Englobe NO_3^- , NO_2^- , NH_4^+ , Si(OH)_4^- mesurés dans la colonne d'eau
PIGM	Pigments	Mesures HPLC de chlorophylles, caroténoïdes, autres pigments
PHYTO	Phytoplankton	Données sur la composition, biomasse, types fonctionnels

| 7

