Sharing Ocean Data Through Standards

An Introduction to the Ocean Data & Information System (ODIS)

Jeff McKenna
OIH technical team

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

Blue Rocks, Nova Scotia, Canada

OSM node (5750230905)

FOSS4G
THAILAND 2022
KhonKaen ◇ 25th - 27th November

FOSS4G:UK
LOCAL 2022

FOSS4G AKL

October 27, 2022 – Auckland, New Zealand

Reunion of the 3 original founders of the global FOSS4G event (from left to right: Markus Neteler, Jeff McKenna, Venka Raghavan); photo taken at FOSS4G Europe, July 2017, in Paris

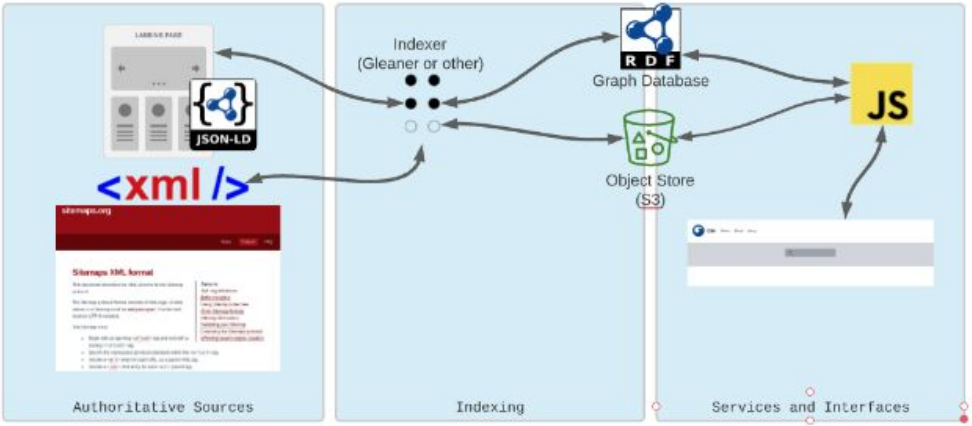
Core of ODIS: patterns and schema.org

Ocean InfoHub

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

Leverage existing standards

- Structured data for the Web: schema.org
- JSON-LD exposed through the sitemap



This is nothing new

- Same standards leveraged by Google Dataset search
- https://datasetsearch.research.google.com/

Google Dataset Search Beta

Initial profiles (“patterns”)

The initial priorities for ODIS (as identified by partners) were to develop specifications to facilitate discovery for six priority themes:

- (i) Experts and institutions/organizations,
- (ii) Documents,
- (iii) Spatial data and maps,
- (iv) Research vessels,
- (v) Education and training opportunities,
- (vi) Projects.

ODIS Book

Ocean InfoHub

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

ODIS Book

- https://book.oceaninfohub.org

Search this book...

The Ocean InfoHub Project

INTRODUCTION

Structured Data on the Web

Personas

Publisher

JSON-LD Foundation

PROFILES

Thematic Patterns

Core 1: Experts and Institutions

Core 2a: Documents

Core 2b: Datasets

Core 3: Spatial Maps

Core 4: Projects

The Ocean InfoHub Project

Introduction

Organizations are increasingly exposing data and resources on the Web. A popular approach to this is using web architecture to expose structured data on the web using the schema.org vocabulary. Doing this makes resources discoverable by a range of organizations leveraging this architecture to build indexes. These include major commercial indexes, large domain focused groups and community focused services.

The Ocean Data and Information System (ODIS) will provide a schema.org based interoperability layer and supporting technology to allow existing and emerging ocean data and information systems, from any stakeholder, to interoperate with one another. This will enable and accelerate more effective development and

Contents

Introduction

Guidance for the implementation of the ODIS-architecture

Key links to the OIH GitHub repository

Info-graphic

ODIS-arch

- https://github.com/iodepo/odis-arch

ODIS Architecture

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

Architecture

The "provider" interface is via web architecture, so the ODIS/OIH Arch can be leveraged; using simple web servers with statically generated pages, groups using CKAN/DKAN (with plugins), ERDAP (DAP systems), etc

Properties of OIH Architecture

- Docker container deployment for local and cloud based implementation
- Compatable with multiple scheduling approaches
- No software lock-in, multiple solutions/options at all stages
- Web architecture based
- Standards body based (both standards and recommendations)

Expected Outcomes

- Multiple indexes can be leveraged (text, spatial, etc)
- Easy to implement or re-implement based on established patterns
- Validation approaches to the data resources

Provenance

Provenance provides temporal record of the activities. OIH uses complementary and unique provance provided by the provider & indexing activity.

Provider Prov:

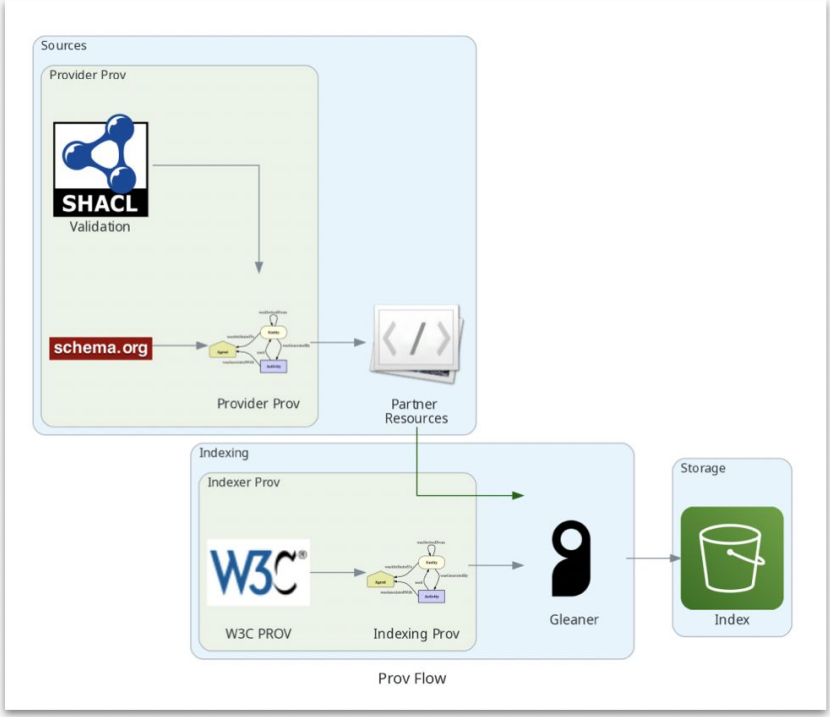
- Authored by provider
- Based on Schema.org
- Provides information from the source on history of and relations between resources.

Indexer Prov:

- Authored by indexer
- Based on W3C PROV
- Potentially different or absent in other implementations of the OIH architecture
- Provides information on resource used in indexing

Expected Outcomes

- Richer query results (freshness, connections, etc)
- Exposing both capacity and gaps in graph relations
- Better credit/citation reporting



Patterns and Validation

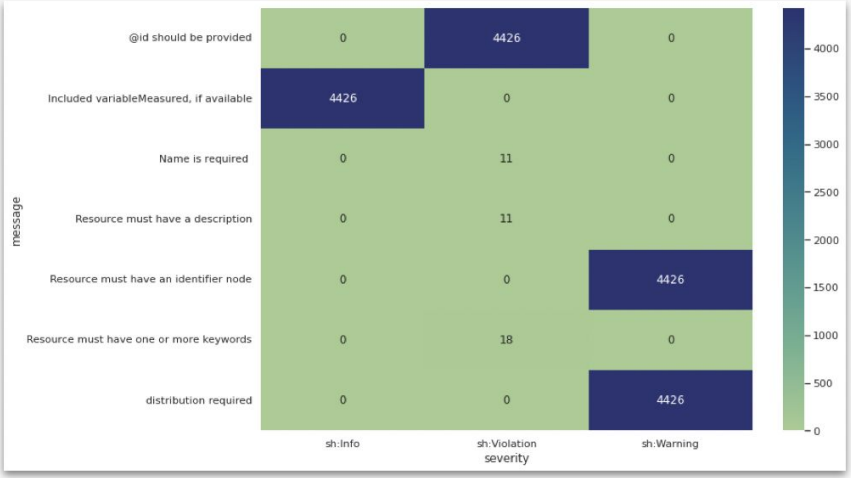
Ocean InfoHub has developed and continues to refine as set of validation shapes.

These shapes allow us to provide partners feedback on how well their resources align to OIH guidance and how well they will respond to portal queries.

Through this approach, OIH can provide required, recommended and informational guidance on partner resources.

Expected Outcomes:

- Fit for use: Are partner resource aligned with query (Findable) or expected or potential use patterns (Accessible, Interoperable).
- Integration: Validation shapes can also help mediate connections between OIH and other sources.



Validation Reports

SHACL shapes defined with the community provide guidance on using the profiles.

Severity levels (violation, warning, info) help convey importance. Notebooks and workflows have been created to provide overview and resource level feedback to partners on alignment

Alignment and Interoperability

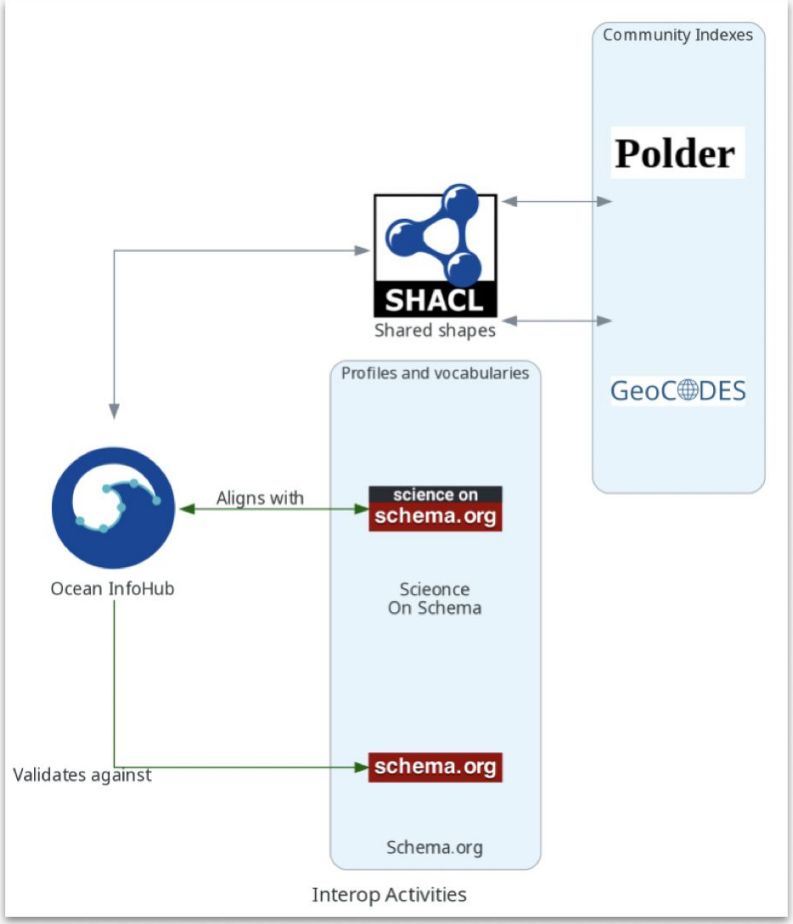
OIH maintains alignment with upstream works like Schema.org and ESIP Science on Schema.org.

A pattern for alignment with others like Bioschemas, EOV community, PID community etc.

Discussions with other groups (POLDER and NSF GeoCODES, etc) on how simple SHACL shapes could help facilitate federated queries between our respective graphs

Expected Outcomes:

- Further FAIR goals
- Leverage other community graphs
- Leverage PID graphs (DataCite, Orcid, ROR, others?)



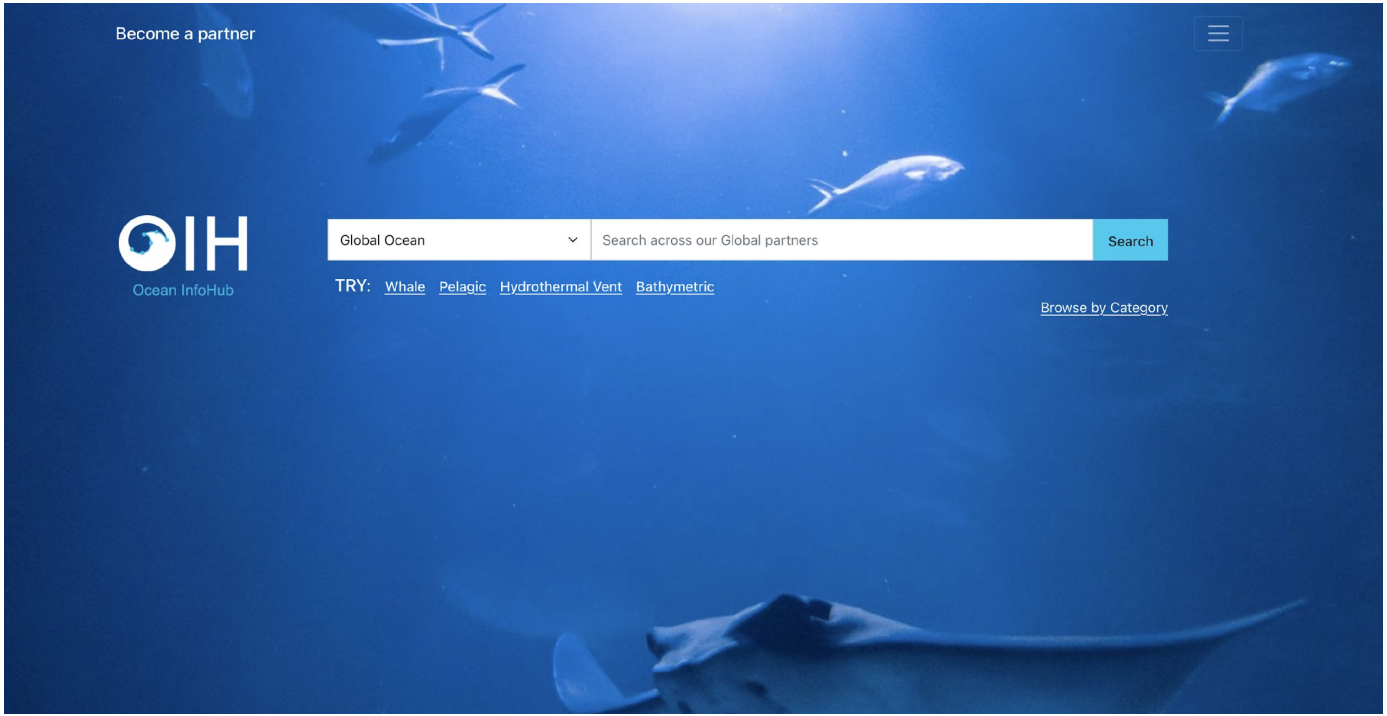
ODIS endpoints

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

Config (YAML)

- Point to sitemap, ODISCatalogue entry
- https://github.com/iodepo/odis-arch/blob/schema-dev/config/sources.yaml

OIH Search


Become a partner
Ocean InfoHub
Global Ocean
Search across our Global partners
Search
TRY: Whale Pelagic Hydrothermal Vent Bathymetric
Browse by Category

SPARQL endpoint

- https://oceans.collaborium.io/sparql.html

Query

https://ts.collaborium.io/blazegraph/namespace/development/sparql

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT * WHERE {
  ?sub ?pred ?obj .
} LIMIT 10
```

Dashboard

Ocean InfoHub

Developing a Dashboard for ODIS partners

ODIS Dashboard

ODIS Node Summary

Select an ODIS node

Ocean Biodiversity Information System

Types indexed

	count
https://schema.org/Person	345127
https://schema.org/Organization	87007
https://schema.org/CreativeWork	41058
https://schema.org/Event	40393
https://schema.org/Dataset	8509
https://schema.org/ResearchProject	3354
https://schema.org/CourseInstance	1651
https://schema.org/Course	1282
https://schema.org/Vehicle	85

Keywords

	count
Fisheries	14655
Biology	6930
Occurrence	6819
Observation	5311
Aquaculture	3932
Ecology	3516
GCFI	2952
Iran	2794
Management	2693
Limnology	1675

Number of Catalogues

4382

Date indexed to Graph

2021-02-26

Size of ODIS graph

6712460 triples

Types indexed

	count
https://schema.org/Person	345127
https://schema.org/Organization	87007
https://schema.org/CreativeWork	41058
https://schema.org/Event	40393
https://schema.org/Dataset	8509
https://schema.org/ResearchProject	3354
https://schema.org/CourseInstance	1651
https://schema.org/Course	1282
https://schema.org/Vehicle	85

Number of Nodes

12

Timeline: when added to Graph

	propername	dates
0	Ocean Best Practices	2021-01-26
1	Ocean Biodiversity Information System	2021-02-26
2	AquaDocs	2021-02-26
3	Marine Training EU	2021-05-26
4	EDMERP SeaDataNet	2021-07-26
5	EDMO SeaDataNet	2021-07-26
6	OceanExpert UNESCO/IOC Project Office for IODE	2021-08-26
7	INVEMAR	2022-01-26
8	EurOcean	2022-03-26
9	MASPAWIO: Marine Spatial Atlas for the Western In	2022-03-30

Number of Catalogues

41667

Keywords

	count
Fisheries	43873
Biology	20834
Aquaculture	11790
Ecology	10524
GCFI	8856
Iran	8382
Management	7809
Occurrence	6825
Observation	5317
Limnology	5025

About the Dashboard

ODIS nodes

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

What software are ODIS nodes leveraging to expose their Ocean metadata catalogues?

ckan

GeoNode

pygeoapi

ArcGIS

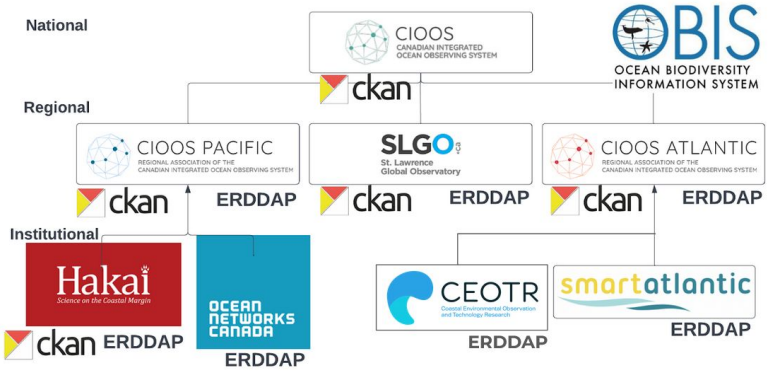
pycsw

ODIS Partner examples

Ocean InfoHub

CIOOS

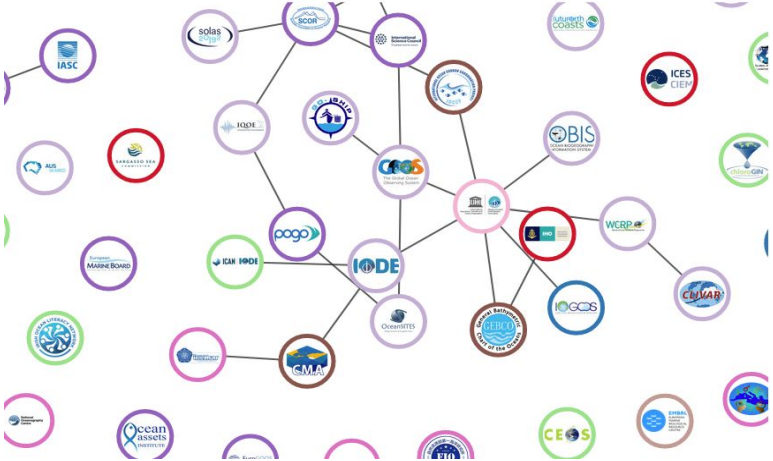
- Meta-node containing regional partners (Pacific, St.Lawrence, Atlantic)
- Strong focus on EOVs
- Brings challenges of how to handle overlapping records with OBIS/other partner catalogues



ODIS Partner examples

OceanScape

- Pragmatic solution
- Existing WordPress site, with an *Organizations* database stored in MySQL
- Worked together in weekly tech meetings to implement an existing schema.org plugin to expose metadata through schema.org, and connect into ODIS



Patterns in development

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

Protected Areas

- https://github.com/iodepo/odis-arch/issues/101

Times-series observations

- https://github.com/iodepo/odis-arch/pull/128

World Ocean Database

Other needed patterns?

- fishing impacts
- coastal pollution
- illegal fishing
- ?

Homework

GEO Blue Planet Symposium 2022
Accra, Ghana
2022-10-24

OceanTeacher : ODIS module

- https://classroom.oceanteacher.org/enrol/index.php?id=722
  - Online modules to introduce ODIS concepts (patterns, JSON-LD, etc.)

Ocean InfoHub

https://book.oceaninfohub.org
https://oceaninfohub.org
https://www.odis.org

Join the Ocean InfoHub and build a truly shared Ocean

Contact us through e-mail at info@oceaninfohub.org

We can help you share your organisation's Ocean data

The Ocean InfoHub is leading OceanData-2030, a registered Programme of the UN Decade for Ocean Science for Sustainable Development

unesco Intergovernmental Oceanographic Commission

2021 2030 United Nations Decade of Ocean Science for Sustainable Development

IODE

Flanders State of the Art