WP2 OIH contractor overview

March 2021

Introduction

Goals:

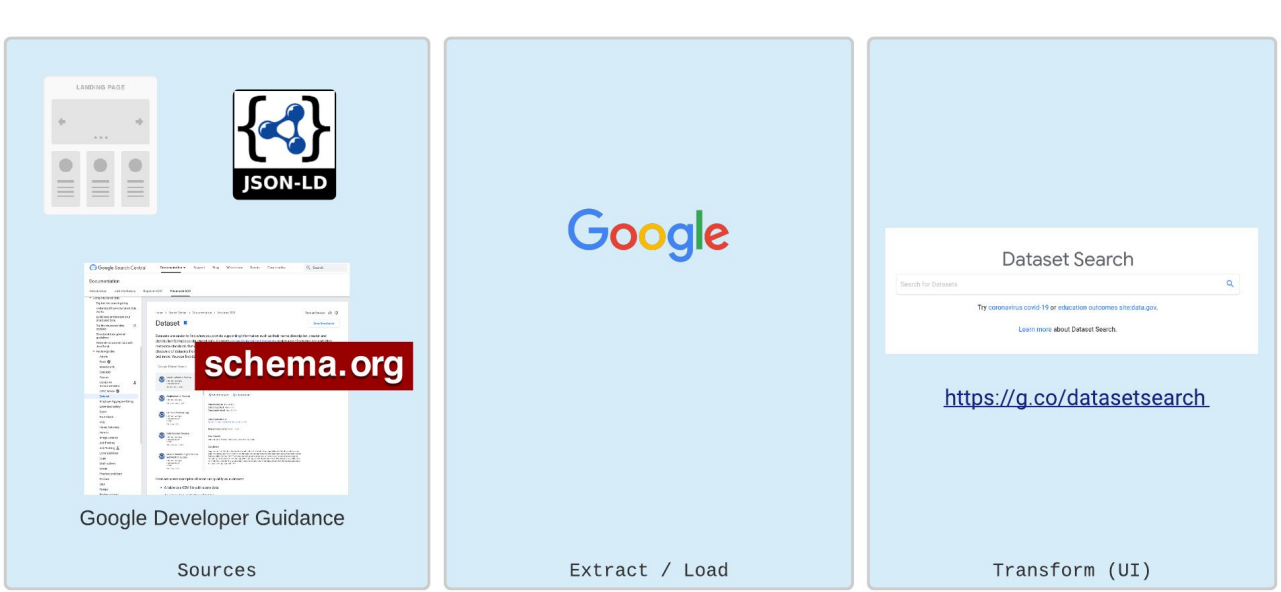
A hub providing a shared graph of metadata for community resources, open access to that graph and a reference search UI. (full UI in a later phase)

Design Principles:

- Leverage existing systems & skills
- Linked Open Data approach
- Decoupled/modular elements
- Encapsulation of metadata (carry context with metadata)

Architecture Decision:
These goals and design principles results in the selection of a ***structured data on the web*** approach leveraging ***web architecture, schema.org and JSON-LD***.

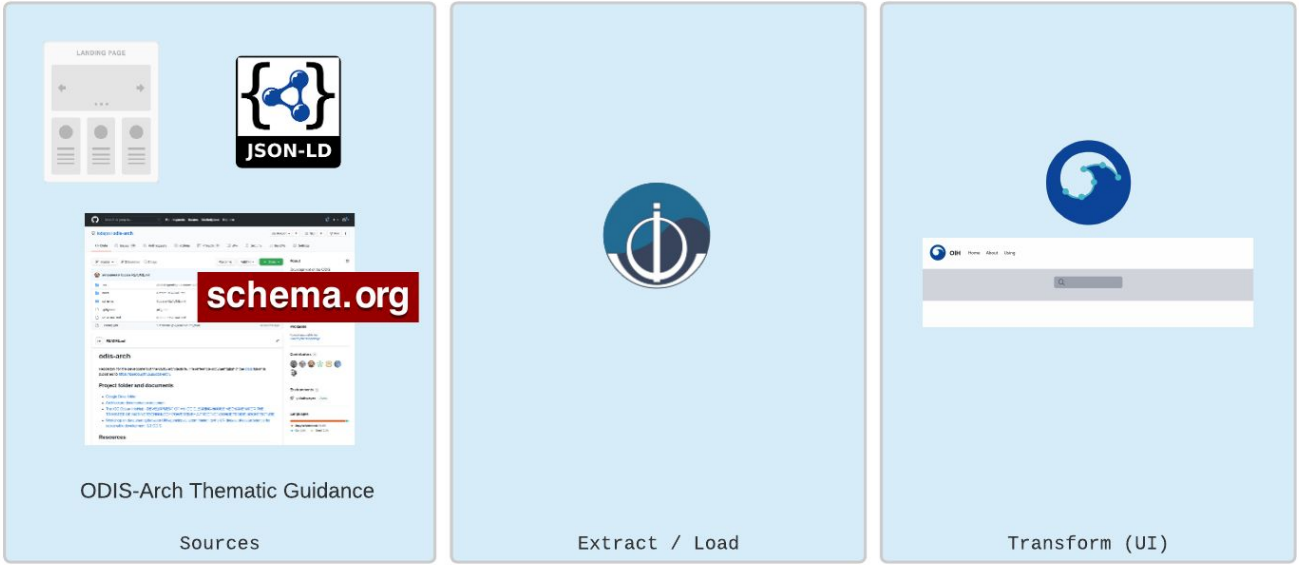
By example: Google Dataset Search



An example of this approach is seen today in the form of Google Dataset Search and the associated guidance.

This approach is also used by Google Job Search, Bing, Yandex as well as other groups such as NSF's EarthCube, BioSchemas, DataONE, etc..

ODIS Ocean Info Hub

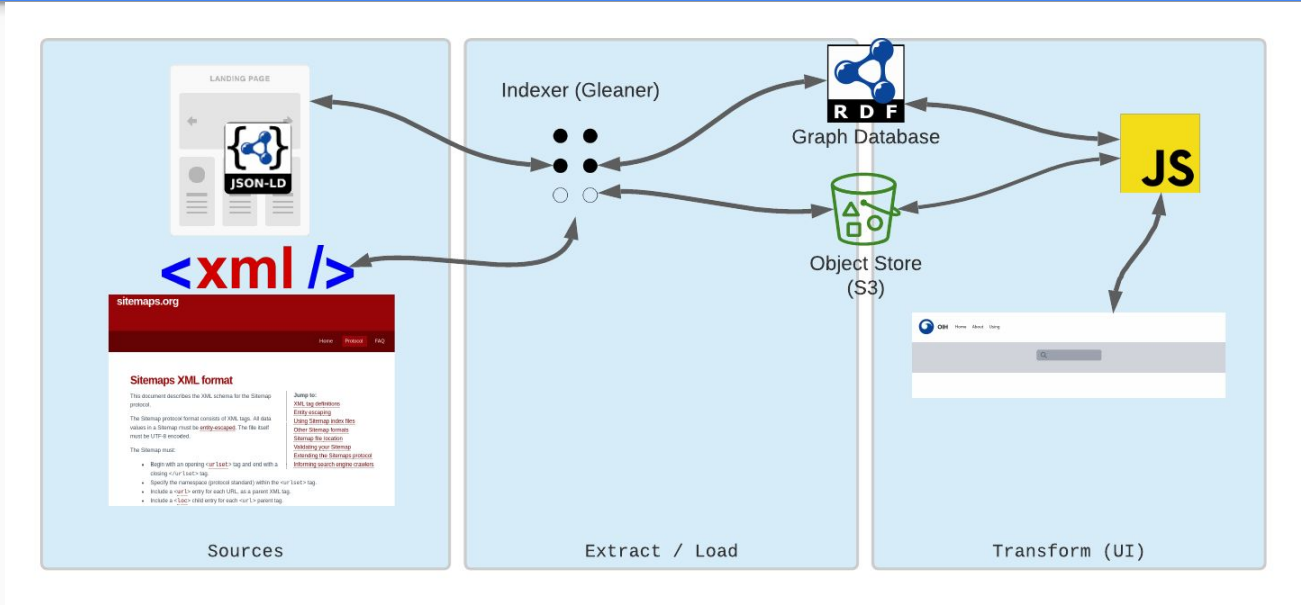


OIH follows a similar approach. Work with the community to provide shared guidance on thematic types. Index/generate an open graph, expose implementations of that graph.

Rather than type dataset, OIH is expressing schema.org based data for multiple thematic types, but across a defined community of practice.

Working as a community in a manner that is both open and scalable.

ODIS Ocean Info Hub (bit more detail)



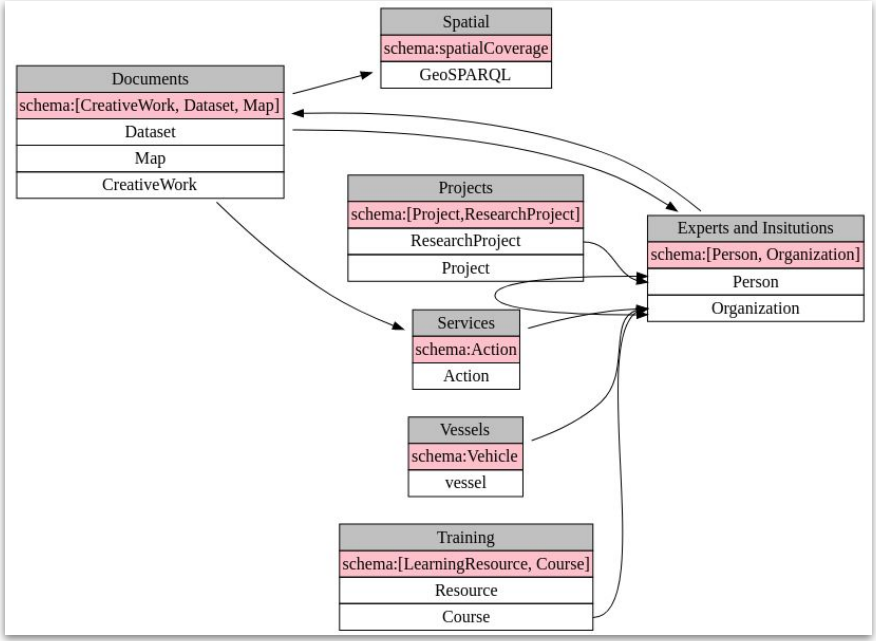
Sources expose resources (web pages) with JSON-LD data embedded

Indexers extract and load these resources (using sitemap.xml as a guide) into an object store and graph database

Transforms through such applications as interfaces and notebooks leverage the object stores and graph database to provide services and analysis.

Architecture Review

There are 5 keys types with two additional *support* types; spatial and services.



Example of possible relations between thematic types

Graph Relations:

As a graph, we can begin to expose relations in the graph and among the contributors.

Courses can be connected to people and organizations, documents to services and more.

Connects are NOT limited to this scope either. Connections to PID graph, WikiData and others mean community resources become both enriched and discoverable by more and more means.

References:

[ODIS Arch documentation]

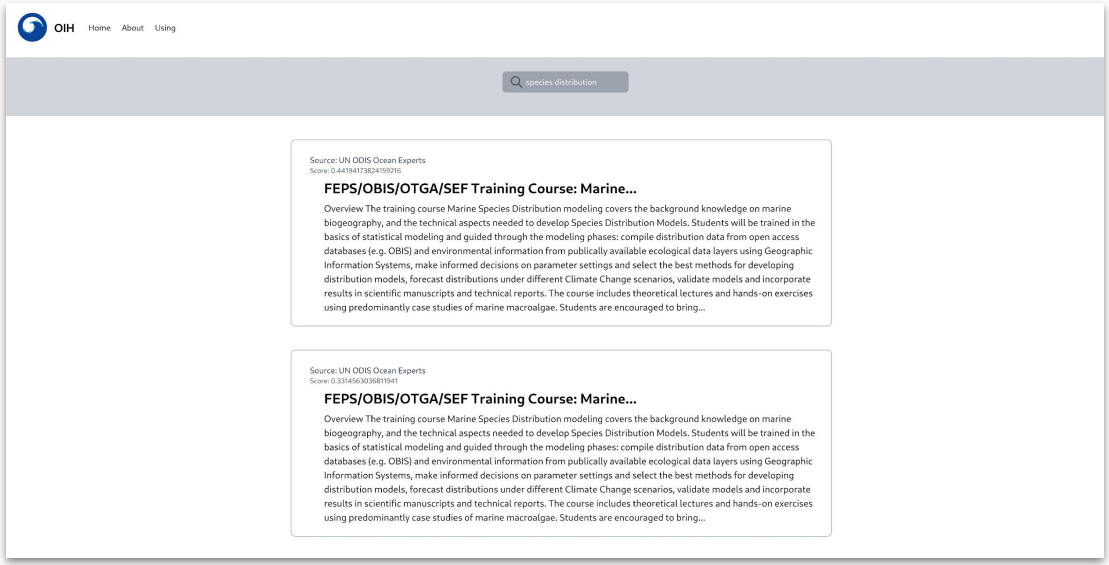
Conclusion Demo

Quick Demo:

Some examples include:

"species distribution"
"marine gis"
"Coral reef systems"

NOTE: The queries at this site are not optimized for speed yet. Also, they don't include a some group calls so it's possible a few duplicates will appear. These will be addressed as the query is evolved.



Demo at: https://oceans.collaborium.io/index.html

NOTE: This is ***NOT a sustained service***. It is ONLY a test site and it will be removed. Please do not publicly share or cite this URL. You are welcome to use it to test and evaluate the work.