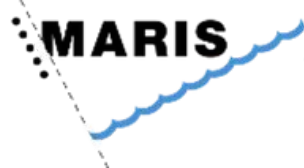


27-29 May 2024 



imdis

International conference on Marine Data and Information Systems





International conference on **Marine Data**
and **Information Systems**

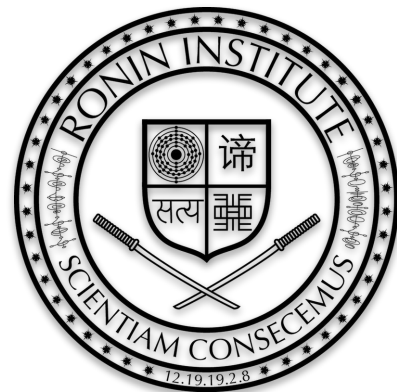
ir. Marc Portier, VLIZ

Functional «Semantic Web» Practice in the marine domain



credits

Partners, projects and co-authors



co-authors: **IODE (UN)** Lucy Scott, Arno Lambert, Pieter Provoost - **Ronin Institute (US)** Douglas Fils - **AWI (DE)** Pier Luigi Buttigieg - **Digitaal Vlaanderen (BE)** Simon Claus - **UGent-imec (BE)** Julian Rojas - **CNR (IT)** Giorgia Lodi, Alessandro Russo - **VLIZ (BE)** Cedric Decruw, Lennert Tyberghein, Bart Vanhoorne

... still waiting for the pages-web to data-web transition ...

Reading one book → querying one database

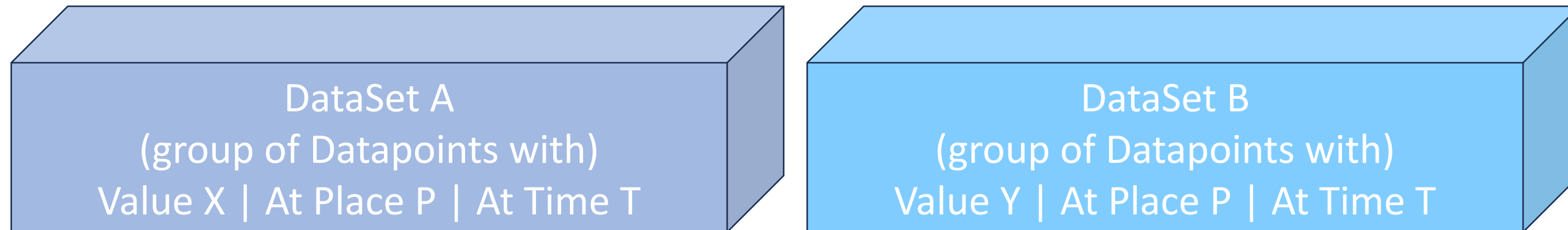
The experience of browsing the web
is like reading a single book...

The Semantic web promise is to similarly explore the web of data
like querying a single database...

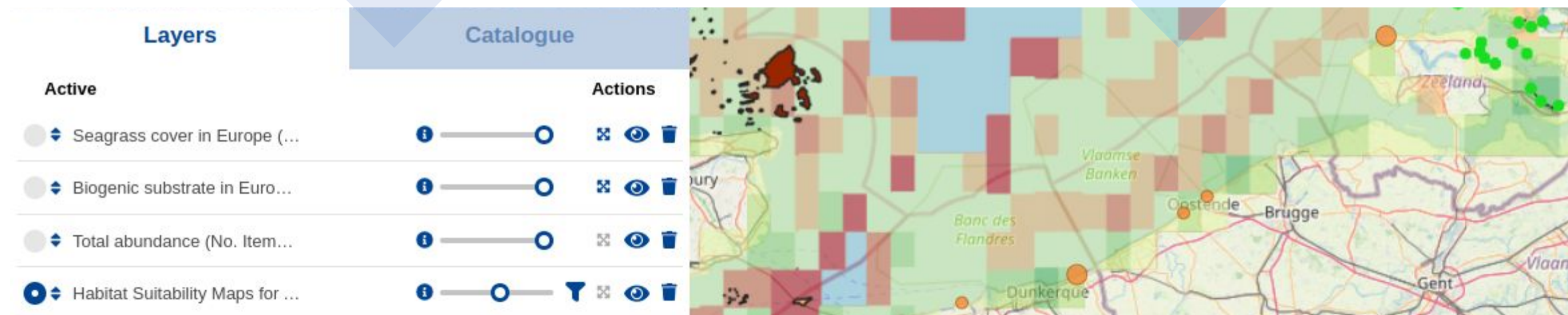
Why is that not happening?

*the technology stack is too complex !
(but is it?)*

Classic approach in Data Analysis

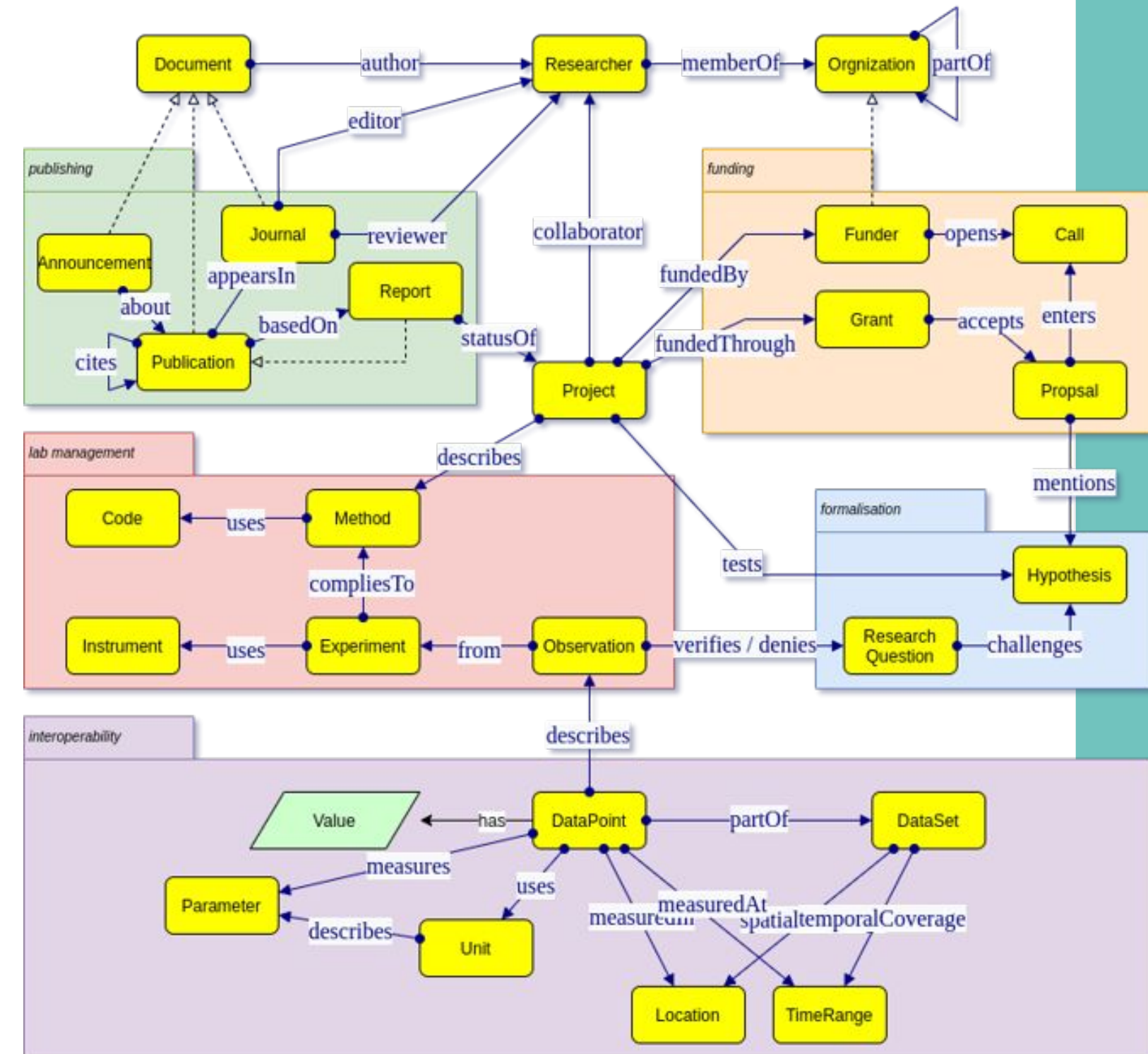


- Compare different values (like layers in a map)
--> datapoints with some shared geo-temporal properties
- To discover some hypothesis connection / causal relation



Research Knowledge Graph

- Adding all aspects of observations:
 - Where, When,
 - Value, Unit, Instrument, Platform, Procedure,
 - Taxname, Class, Function, Attribute, Trait,
 - Maintenance, Storage, Lab, Datasets,
 - Question, Hypothesis,
 - Project, Funding, Grant, Organisation, Person...
- ... and how they are connected



More opportunities for inference

- To extend the range of "coincidence" axes / dimensions
 - Not only geo-temporal (Where and When)
 - But along all conceivable links
- Enables: to detect more "connections"
- Feeds the Research Cycle:
 - *[Statistics]* → *[Coincidence]* → *[Hey, that's funny]* → *[Hypothesis]*
 - *[Causation]* → *[Knowledge]* → *[Policy]*
- ... and thus achieve the "Open Ended" ambition ?
 - Supporting unpredictable rehashing, mashups, ...

complexity through simplicity

RDF really **is simple**



"People think RDF is a pain because it is complicated. The truth is even worse. RDF is painfully simplistic, but it allows you to work with real-world data and problems that are horribly complicated. While you can avoid RDF, it is harder to avoid complicated data and complicated computer problems."

– Dan Brickley and Libby Miller
(in the foreword of the book "[Validating RDF](https://book.validatingrdf.com/)")

back to basics, make your life easy

- publishing Linked Open Data (LOD) is easier than publishing HTML
 - you don't need a SPARQL backend ! *(it is not the equivalent of a SQL store)*
 - you don't need a SPARQL backend ! *(forget endpoint - think startpoint)*
 - you don't need a SPARQL backend ! *(more like graphql a client support API)*
 - you don't need a SPARQL backend ! *(they still show bad uptime statistics)*
- publish first, craft models as you go
 - focus on the values ("nobody is interested in the types" – Michael Kay)
 - known property paths first ("as we may think" – Vannevar Bush)
 - don't overengineer ontologies ("semantics is use" – Ludwig Wittgenstein)
 - required / first : semantics, clarity, no-ambiguity, standards & formatting fit
 - optional / later: logic, type-constraints, shacl, validation

alternative road - keep it simple

- just produce RDF representations (<script> embed and/or linked),
 - static generators rule, converters exist, own code will do too
- pay attention to well designed durable URI
 - that double as working URL (dereferenceable) + optionally support conneg
- reuse (popular) terms and vocabs
 - General: schema.org, dublin-core, void, dcat,
 - Research: ror.org, orcid.org,
 - Marine: vocab.nerc.ac.uk/ (bodc) - ODIS-Arch, OIH (IODE) - maregraph.eu - edmos edios, edmerp (SeaDataNet)
- use smart web-linking tricks like fair signposting
- consider change-feed : LDES (~RSS for data)

basically: just update your website to use modern web standards ! (not the looks)

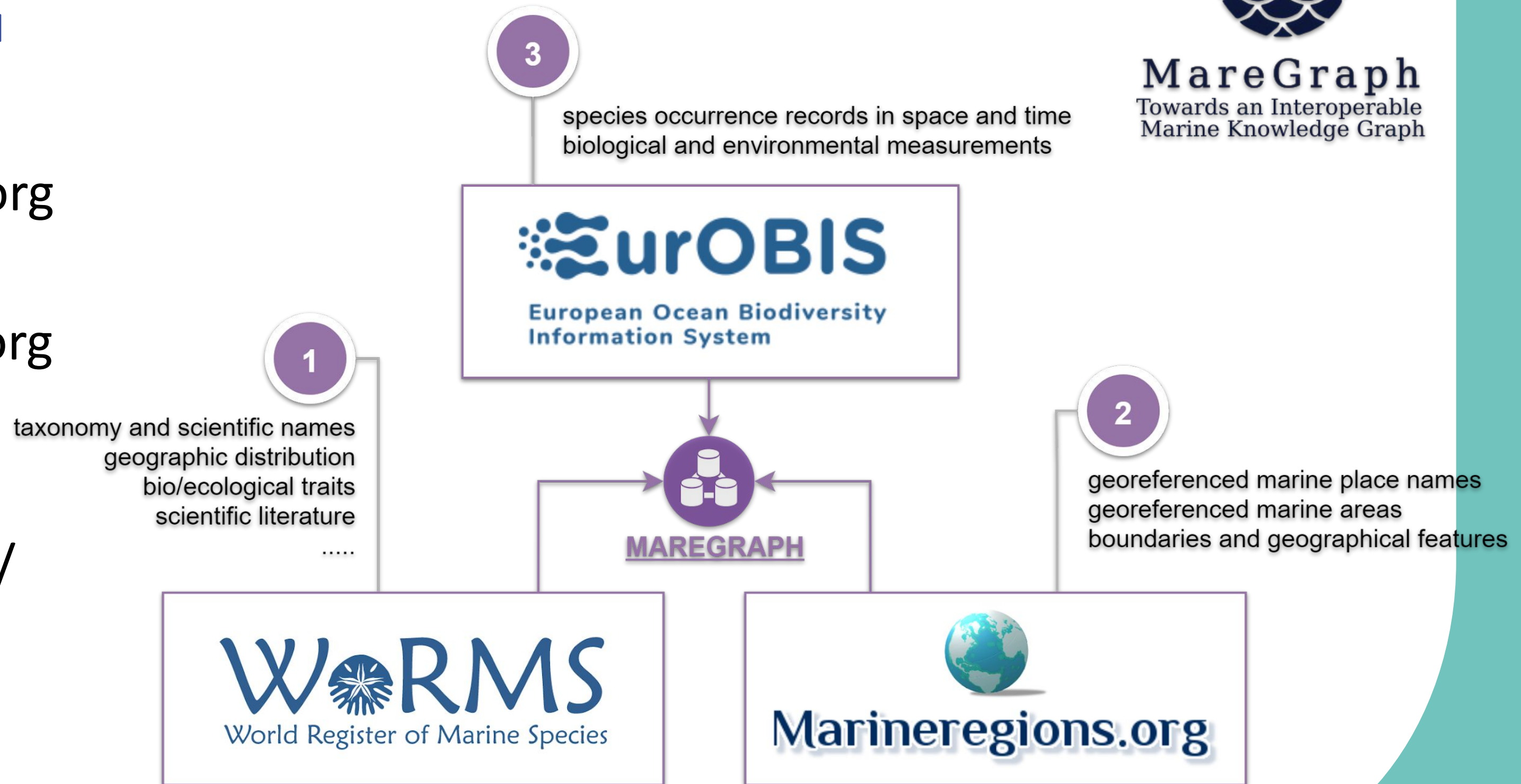
emerging / fresh work

maregraph.eu

- marinerregions.org
→ locations
- marinespecies.org
→ taxnames
- eurobis.org
→ occurrences /
biodiversity



MareGraph
Towards an Interoperable
Marine Knowledge Graph



achievable results

seeing the graph @work



- oceaninfohub.org | dashboard.oceaninfohub.org | SPARQL endpoint

The screenshot shows the ODIS Dashboard on a web browser. The page title is 'ODIS Dashboard' and the URL is 'dashboard.oceaninfohub.org'. The dashboard provides information about the ODIS graph, including a SPARQL endpoint, a node description, and a triple description. It also displays a summary of the graph's size and the number of nodes and entities described.

SPARQL Endpoint

SPARQL (SPARQL Protocol and RDF Query Language) is the query language that is used to query graphs. ODIS has a [SPARQL Endpoint](#) that allows you to directly query the ODIS graph. See the [ODIS book](#) for example queries.

Node

The ODIS graph consists of many nodes, which represent organizations, each with their own catalogue of data.

Triple

The ODIS graph connects information through a triple; *Subject, Predicate, Object*. In the information "Leonard Nimoy was an actor who played the character Spock", LeonardNimoy is the

This dashboard will help monitor the ODIS graph, as well as the nodes connected to it.

[Open Hub](#) \$2.57M Cost

ODIS Graph Endpoint Status

ODIS Graph Summary

Size of ODIS graph	Number of Nodes	Number of Entities Described
3014339 triples	47	665305

name	Sitemap Status	Node
acma	✓	African Coast
africaioc	✓	IOC Africa Dat

type	count
schmea:CreativeWork	288,427

Date Added	Node
2024-04-23	African Coastal and Marine Atlas

The screenshot shows the SPARQL endpoint interface on a web browser. The URL is 'http://graph.oceaninfohub.org/blazegraph/namespace/oih/sparql'. The query is:

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

The results are displayed in a table with columns 'sub', 'pred', and 'obj'. The results show information about Oceanneering International, Inc. and its reports.

sub	pred	obj
<https://edmo.seadatanet.org/report...>	<https://schema.org/name>	Oceanneering International, Inc.
<https://edmo.seadatanet.org/report...>	<https://schema.org/url>	https://edmo.seadatanet.org/report/3600
<https://edmo.seadatanet.org/report...>	<https://schema.org/descriptio...>	Oceanneering is a global oilfield provider of engineered services and products primarily to the offshore oil and gas industry, with a focus on d...
<https://edmo.seadatanet.org/report...>	<https://schema.org/identifier>	3600
<https://edmo.seadatanet.org/report...>	<https://schema.org/address>	<https://gleaner.io/xid/genid/ckatc63h2h49b33u6ijg>

(upcoming – target release 2024)

Showcasing LOD benefits

- Vocab-Term-Lookup-Service: consistent retrieval of URI in DM edits
 - widget for lookup of URI based on available human labels
 - e.g. BODC measurements, marine-regions, marine-info person
 - Auto-updates on available external changes
- Marine-Info-Affordances: add inline context and explanation online
 - widget for augmenting references to marine-entities
 - e.g. some online text that mentions "[the Belgian part of the North Sea](#)"
 - also for person, publication, ...
 - uses conneg and/or fair-signposting
- Knowledge Graph Analysis Platform (k-gap):
 - in practice:
 - linking usage statistics to "research knowledge graph"
 - reporting on the results



recap

connecting research data through "semantic web done right"

- RDF is easy, it needs to be to allow solutions to complex issues
- the things to do and to not do
 - keep things simple and easy
 - up your web-standards game
 - let the complex stuff to later / to others
 - no sparql - no (new) ontologies / type -structures / reasoning rules / ...
 - allow for innovation and prototypes into new directions
- here and now
 - simple showcases, not overselling
- more to come ... "federated querying"
 - see you in next talk :)

27-29 May 2024 



imdis

International conference on Marine Data and Information Systems

