

Increasing FAIRness of marine data within ENVRI-FAIR

Peter Thijsse (MARIS)



ENVRI-FAIR has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068

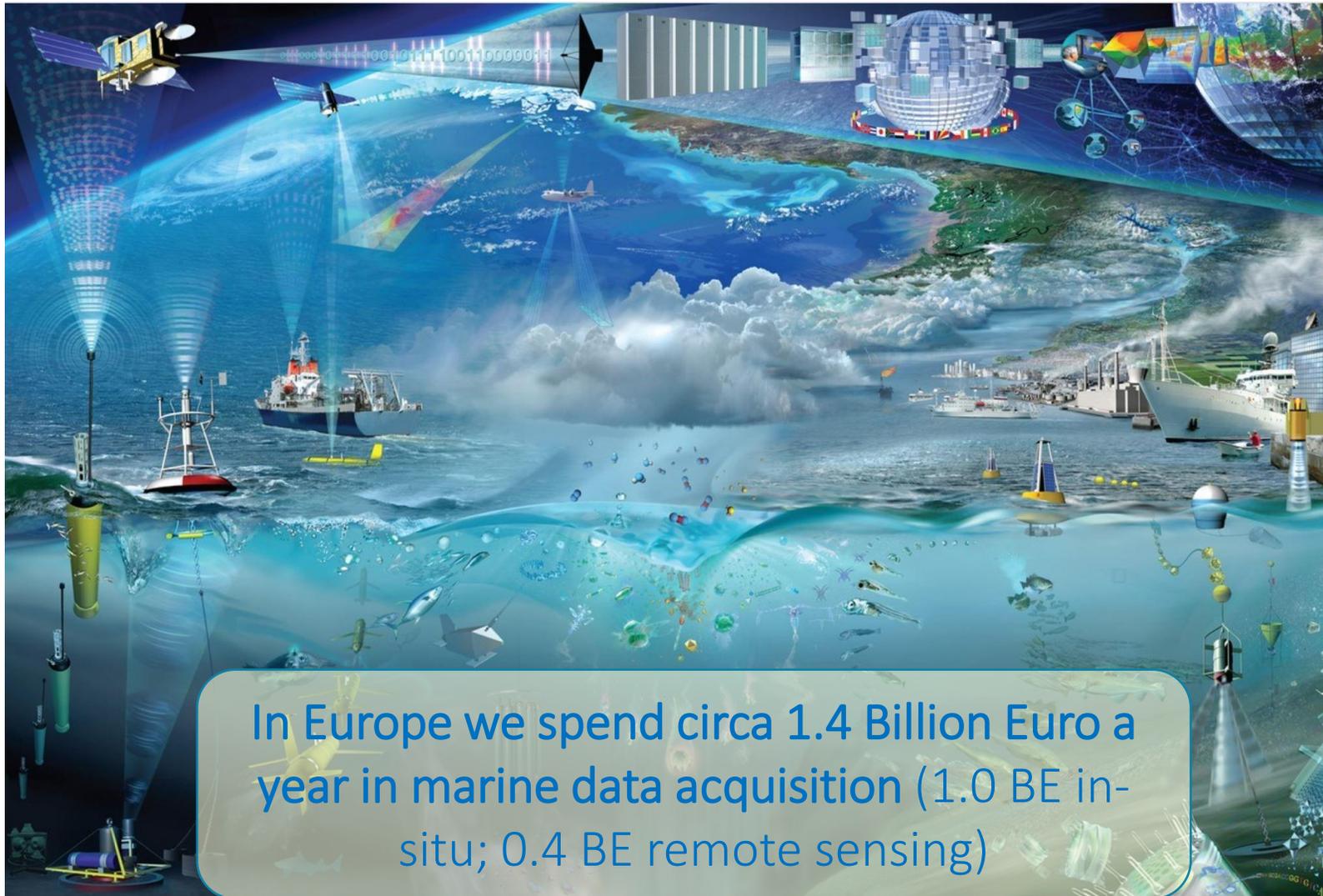


Content

- 🌊 Marine data landscape
- 🌊 What is FAIR?
- 🌊 ENVRIFAIR
- 🌊 Approach for FAIRification of the marine Research Infrastructures
- 🌊 Examples of developments

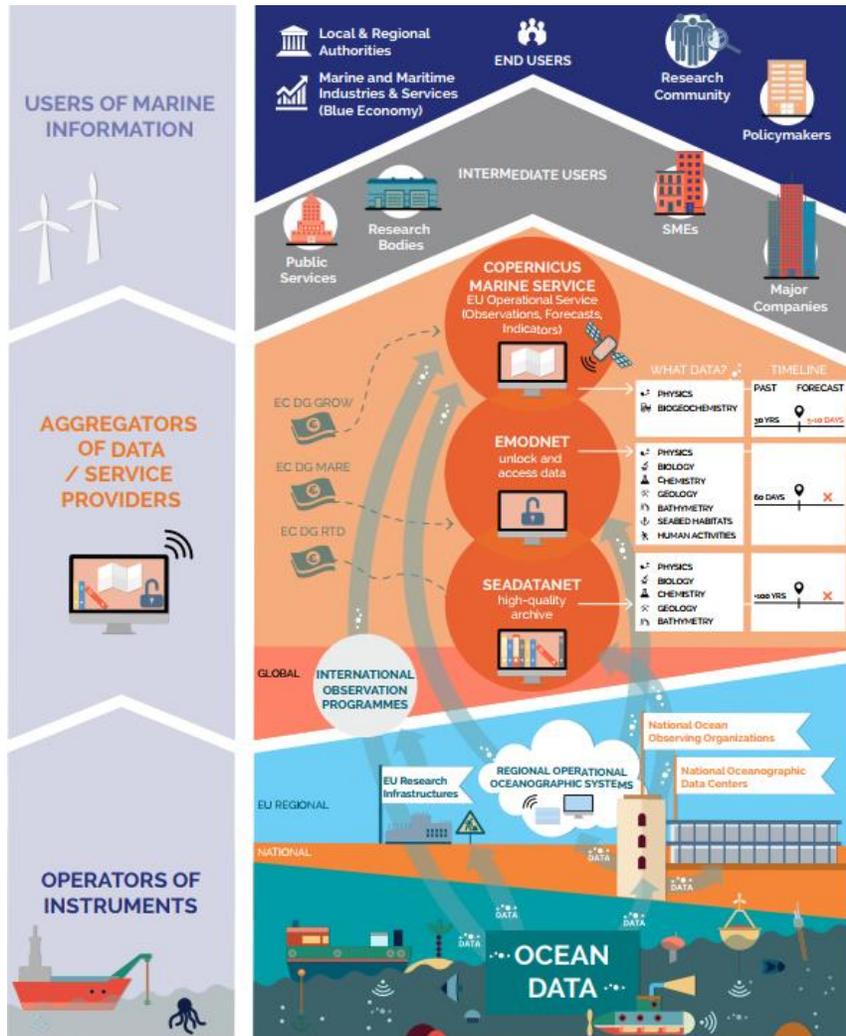


Marine landscape for data acquisition





European landscape marine data management



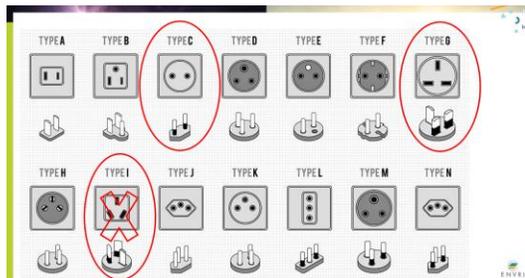
Data aggregators and providers of data products and services



What is FAIR?

The issue

The plug



The FAIR solution

The universal adapter



Findable 

Accessible 

Interoperable 

Reusable 

- FAIR is focused on machine2machine services
- Implementing FAIR principles don't require to change the data systems completely, but requires to **add information and services** at interface level. This leads to easier discovery and access to interpretable metadata and the data for machines (think of VRE's, Jupyter notebooks, etc.)





ENVRI-FAIR

- ENVRI-FAIR (2019 - 2022) is focused on the preparation of the connection of the Cluster of Environmental Research Infrastructures (ENVRI) to the European Open Science Cloud (EOSC).
- The overarching goal is that at the end of the project, **all participating Research Infrastructures have built a set of FAIR data services** which enhances the efficiency and productivity of researchers, supports innovation, enables data- and knowledge-based decisions and prepares the ENVRI Cluster for connection to EOSC.
- Marine RI's form one of the subdomains: **Euro-Argo, EMSO, ICOS-Marine, LifeWatch, SeaDataNet.**



www.envrifair.eu



Approach towards FAIRification

- Step 1: Do a **FAIRness analysis** for each RI's **data repositories** (documenting the status as a FAIR implementation profile)
- Step 2: Assessment of the FAIRness of each of the marine RI's, documenting current strengths, **identified gaps** and planned and/or suggested directions for solutions to overcome those gaps.
- Step 3: **Set priorities** to enhance RI's data FAIRness (bottom up)
- Step 4: Work out a conceptual implementation plan for each marine RI, resulting in an implementation plan per RI.
- Step 5: Creating a technical **specification** and planning of developments
- Step 6: **Implementation phase** (and repeating FIP to measure improvement)
- Step 7: Validation of services through expert tests, and the EOv product work

FAIR Implementation Profile



A machine-actionable list of all FAIR-enabling resources (like the FDP) and includes domain-relevant community standards.

https://fip-wizard.ds-wizard.org/

FIP Wizard

VODAN

Create Document More

Chapters

- I. General Information about Survey
- II. General Information about Participant
- III. Findability
- IV. Accessibility
- V. Interoperability
- VI. Reusability

More

- TODOs
- Summary Report

Help

III. Findability

Chapter text

F1 What globally unique, persistent, resolvable identifiers do you use for metadata records?

1.a.1 Choose your answer from FAIRsharing

Persistent Uniform Resource Locator

FAIRsharing <https://fairsharing.org/bsg-s001183>

1.a.2 Add your resource description here

Identifiers used in the FAIR Data Point



Examples of developments per RI

Some examples of work on data repositories:

- 🇪🇺 **F: Upgrading search/discovery options** for data and metadata:
 - 🇪🇺 Support for OAI-PMH/CSW
 - 🇪🇺 PID registration
- 🇪🇺 **A: Implement data and metadata API's** allowing interoperable access even if the data are distributed
 - 🇪🇺 Linked data, DCAT-AP, with Sparql endpoint
 - 🇪🇺 Restful API (e.g. ERDDAP service)
- 🇪🇺 **I: Upgrading metadata using community agreed vocabularies in metadata descriptions** that will allow mapping from one network to another, e.g.:
 - 🇪🇺 CF Convention (cfconventions.org) and/or NERC vocabulary system (vocab.nerc.ac.uk) for parameters, QC flags ,
 - 🇪🇺 EDMO codes for institutions: www.seadatanet.org/Metadata/EDMO-Organisations
 - 🇪🇺 ORCID for persons: orcid.org/
- 🇪🇺 **R: Upgrading provenance** information:
 - 🇪🇺 Use of unique ID's (PID and DOI) for datasets to be able to track the data all along its lifetime
 - 🇪🇺 **Adding provenance information to metadata (details of sensor, data origin, quality control, processing steps, software used, ..)**

→ Especially for Interoperability and Reusability data management at the source is crucial.

→ Experiences in the marine subdomain and other subdomains will lead to **training materials, Knowledge Base content and a set of recommendations** for others to use.



Take home messages

- Developments to increase FAIRness going on in many science domains. Important to learn from experiences.
- True FAIRness starts at data management at the source, during data acquisition.
- End-users will benefit of this in the long run.

Questions?

Please check <https://envri.eu/home-envri-fair/>
or contact me at: Peter@maris.nl



ENVRI
FAIR



envri.eu/envri-fair



[@envri_fair](https://twitter.com/envri_fair)



company/envri-fair



facebook.com/ENVRIcomm