

FAIR data management

What is it, and how does it fit to LandSeaLot work?

Peter Thijsse, MARIS



LandSeaLot has received funding from the European Union's Horizon Europe Framework Programme for Research and Innovation under grant agreement No 101134575. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Content



- The FAIR metrics – GO-FAIR (short overview/recap)
 - Interpretation of data reliability
- Use and reuse of reliable marine observation data
 - European perspective
- Some practical examples how SeaDataNet has worked towards “as FAIR as possible data
- How to make use of this in LandSeaLot?



1. About FAIR

Growing need for data FAIRness

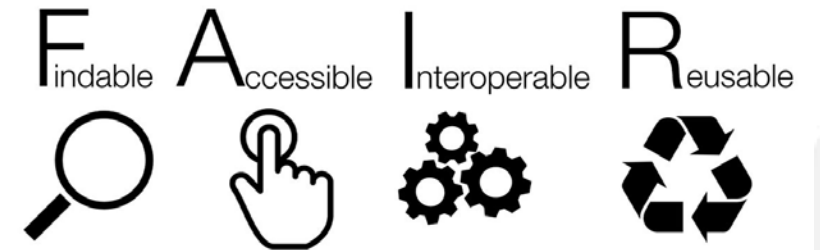


Research and observation data should be more and more **FAIR**

“Findable Accessible Interoperable Reusable” for machines

This is driven by:

- Internet and cloud processing possibilities
 - EOSC
 - Blue-Cloud
 - VRE's/Jupyter notebooks
- Research requirements for reproducibility and traceability
- Growing data availability and need for better findability of data
- And on the other side the cost of observation data and benefit of multiple use.



=> *In the marine domain this has always been important.*

About FAIR





- Movement
- GO-FAIR an important driver
- Important: FAIR for machines


FAIR Principles

GO FAIR is committed to making data and services **findable, accessible, interoperable** and **reusable** (FAIR).

 **Findable:** Metadata and data should be easy to find for both humans and computers.

 **Accessible:** The exact conditions under which the data is accessible should be provided in such a way that humans and machines can understand them.

 **Interoperable:** The (meta)data should be based on standardized vocabularies, ontologies, thesauri etc. so that it integrates with existing applications or workflows.

 **Reusable:** Metadata and data should be well-described so that they can be replicated and/or combined in different research settings.

About GO FAIR

A growing number of scientists, science funders and policy makers advocate the **transition to Open Science**. Various communities have already launched projects designed to make unlinked research data findable, accessible, interoperable and reusable. Too often, however, these important efforts lack synergy.

A **common environment** enabling data-driven research and innovation around the world needs to be **built collectively**. That's where GO FAIR comes in.

GO FAIR aims to avoid silo formation, undue competition and fragmentation by fostering **concerted "FAIRification" efforts** of its participants. This happens on the basis of three pillars: **GO CHANGE, GO TRAIN** and **GO BUILD**.



Benefits

GO FAIR offers an **open and inclusive ecosystem** for anyone committed to defining and creating materials and tools as elements of the Internet of FAIR Data and Services. It is a **community-led** and **self-governed** initiative working across disciplines and countries.

Individuals, institutions and organisations from all over the world engage as **GO FAIR implementation networks**. At the country level, **EU Member States and Associated Countries** join GO FAIR to work towards a smooth and swift involvement of their country's interested communities.

GO FAIR participants benefit from:

-  **knowledge exchange and knowledge transfer** at GO FAIR workshops and meetings
-  a **network** that can be seamlessly integrated into the **European Open Science Cloud (EOSC)**
-  broad dissemination and visibility of their **best practices and project outcomes**
-  effective **information flows** facilitated by the initiative's office

1. The FAIR metrics



Findable

- F1. (Meta)data are assigned a globally unique and persistent identifier
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

Accessible

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
- A1.1 The protocol is open, free, and universally implementable
- A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

Interoperable => think domain specific metadata and vocabulary solutions

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data

Reusable => the biggest challenge for achieving reliability assessment

- R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
 - Context, purpose of data collection, lab conditions, parameter settings, etc.
- R1.1. (Meta)data are released with a clear and accessible data usage license
- R1.2. (Meta)data are associated with detailed provenance
 - Where does the data come from, who collected it, how was it processed, which original data was included.
- R1.3. (Meta)data meet domain-relevant community standards
 - Formal data standard, combine various datasets

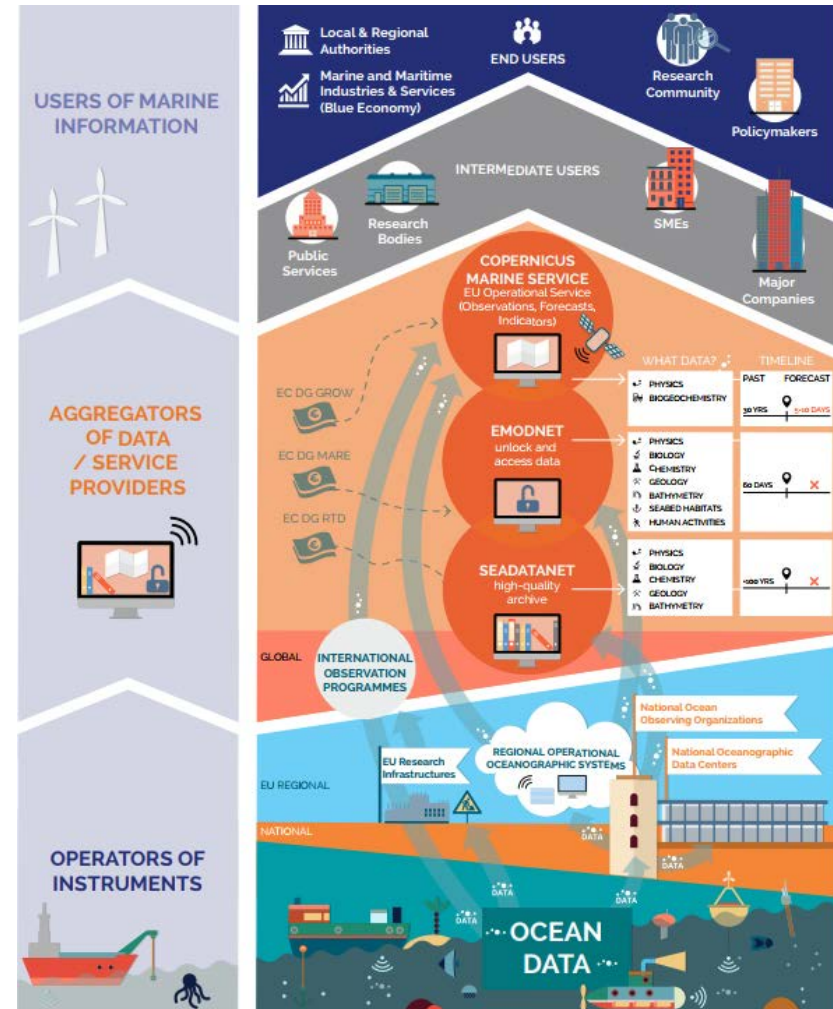


2. European marine perspective

Use and reuse of reliable marine data



- EU perspective
- Intensive observations of seas and oceans
- First use: Local, regional, national
- Reuse: Via national datacentres, EU aggregators
- Eventual end-users: machines (models, data products) and humans (industry, research etc)
- FAIR data, and the means of assessment, is an important key to make data reusable. This starts at the source (metadata, unique ID's etc.)
=> see next example





3. Seadatanet actions to be FAIR

How does SeaDataNet work towards FAIR data?



- F: Registered data discovery services for machines and humans
- A: PID's, Open metadata, authentication layer for download
- I: Standardised data exchange formats, community vocabularies,
links
between metadata
- R: Extensive metadata covering context (originator, instrument,
parameters,
link to sensor), usage conditions/policy, processing software, links to
source data (for data products)

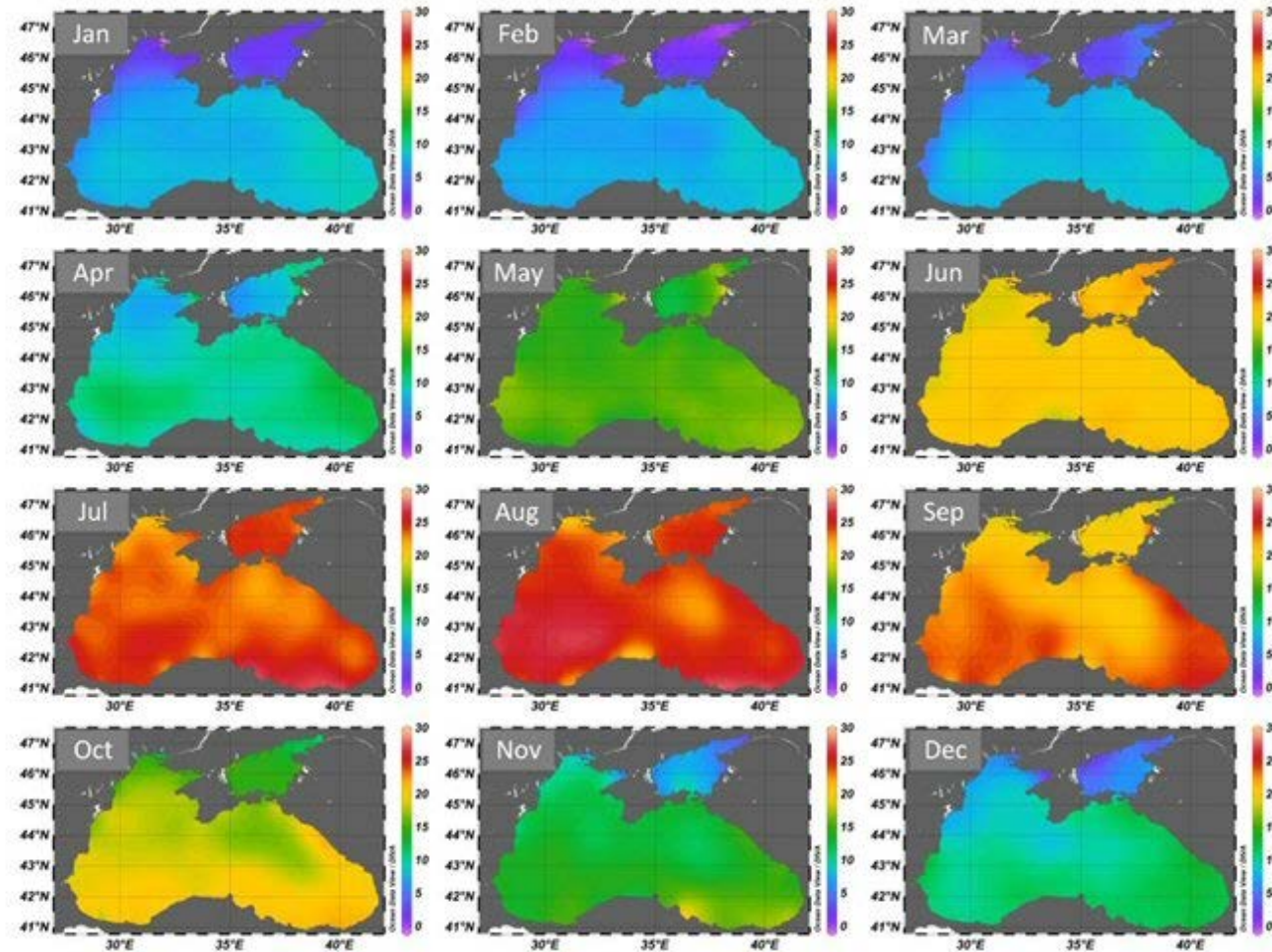
F&A: Catalogue of SeaDataNet data products



Example of
SeaDataNet
climatologies:

Black Sea
monthly
variation of
temperature at
the surface for
the time period
1955 – 2019

All products in
catalogue with
DOI landing
pages



I: SeaDataNet standards



“Making Data and

Services:

- **Findable**
- **Accessible**
- **Interoperable**
- **Re-usable**

*for machines and
people.”*

- **Common standards for the marine domain**, adapting ISO and OGC standards and achieving INSPIRE compliance:
 - **Metadata formats for data sets, research cruises, monitoring networks, organisations, and research projects**
 - **Standard data exchange formats : ODV ASCII and NetCDF (CF)**, fully supported by controlled vocabularies
 - **Controlled Vocabularies** for the marine domain (>89.000 terms in 110+ lists), with international governance and web services
- Maintenance and dissemination of standard QA-QC procedures, together with IOC/IODE and ICES





4. How to make use of this in LSL?

For LandSeaLot data

REDUCING GAPS IN ENVIRONMENTAL OBSERVATIONS AT THE LAND-SEA INTERFACE TO PREDICT SYSTEM BEHAVIOR AND INFORM SCIENCE-BASED MANAGEMENT AND POLICY RESPONSE

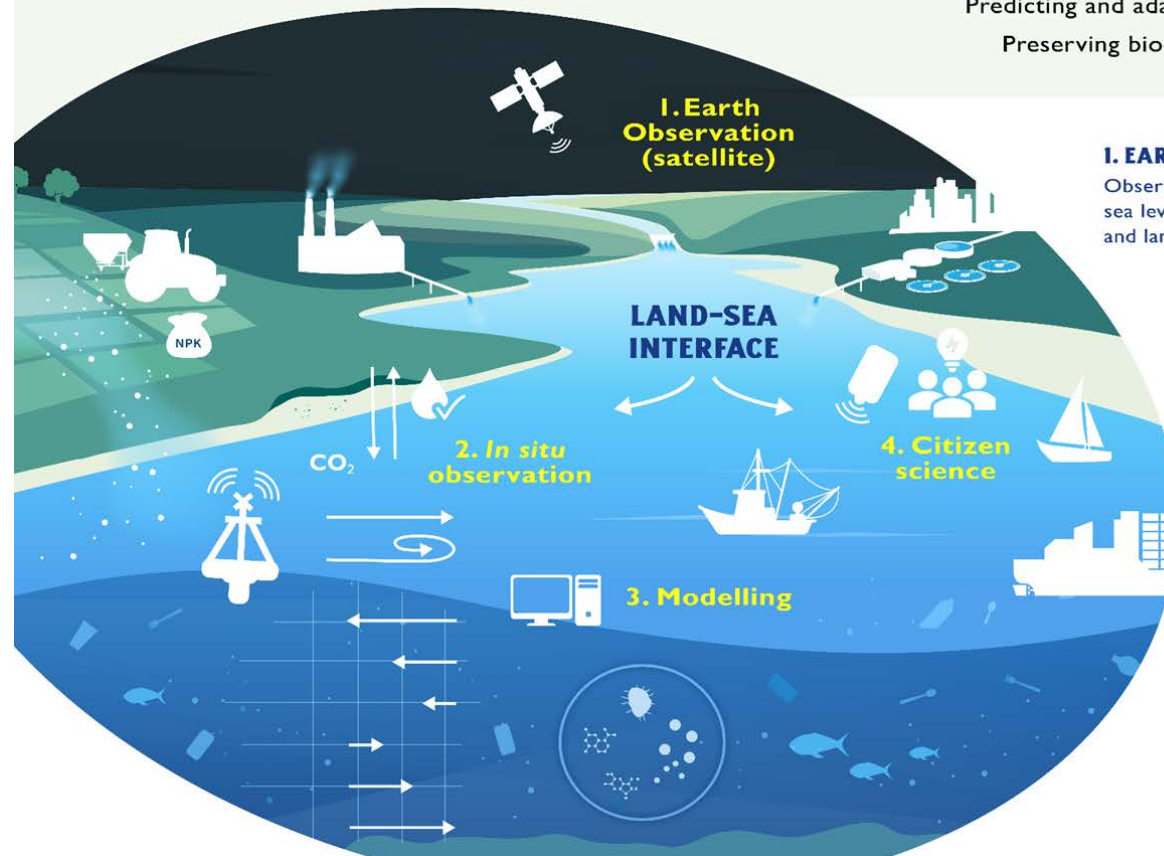


- Dataflows are very diverse
- Data products will be created in the LILs
- Requires reliable data
- Some data we have to use “as-is” and build methods to aggregate
- Other data types LSL can work towards next level of FAIRness
 - River (discharge) data
 - Low-cost sensor data

Testing the benefits of integrated land-sea observations: Understanding lateral carbon fluxes and marine carbon stocks

Predicting and adapting to climate change threats

Preserving biodiversity and reducing pollution



1. EARTH OBSERVATION (SATELLITE)

Observing change in e.g., coastal erosion, sea level rise, turbidity and chlorophyll and land use patterns

2. IN SITU OBSERVATION

Observing change through fixed and mobile platforms in e.g., water quality, coastal erosion and morphological changes, carbon fluxes and nutrients

3. MODELLING

Predicting change in e.g., water temperature, sea level, wave patterns, carbon fluxes, plastic pathways, morphological changes, salt intrusion, water quality, habitats for biodiversity

4. CITIZEN SCIENCE

Gathering data from a wider range of locations and perspectives on e.g., temperature, water level, and plastics

Our ideas from the proposal



- ❖ Some data flows we can not directly improve = external: Remote sensing (Copernicus/ESA), model outputs, in situ data (Euro-Argo, EuroBIS, ..), other citizen science:
 - create syntax and semantic bridges and mappings
 - smart mappings
 - Vocabulary mapping

- ❖ Other data flows we can influence:
 - River data:
 - existing national solutions, integration in EMODnet (as-is)
 - work with the Danubius-RI community
 - work on developing new metadata and data exchange standards
 - expand vocabularies (only develop new ones if needed)

 - Low-cost sensor data (empowering Citizen Science):
 - fully new domain, all open => **we can guide this from scratch!**
 - work on developing new metadata and data exchange standards
 - expand vocabularies (or develop new ones if needed)
 - develop connections and connector to mature sensor developers

- ❖ More in WP4 and WP6

1. EARTH OBSERVATION (SATELLITE)

Observing change in e.g., coastal erosion, sea level rise, turbidity and chlorophyll and land use patterns

2. IN SITU OBSERVATION

Observing change through fixed and mobile platforms in e.g., water quality, coastal erosion and morphological changes, carbon fluxes and nutrients

3. MODELLING

Predicting change in e.g., water temperature, sea level, wave patterns, carbon fluxes, plastic pathways, morphological changes, salt intrusion, water quality, habitats for biodiversity

4. CITIZEN SCIENCE

Gathering data from a wider range of locations and perspectives on e.g., temperature, water level, and plastics



Let's observe together!

Visit landsealot.eu

Be part of the conversation  



LandSeaLot has received funding from the European Union's Horizon Europe Framework Programme for Research and Innovation under grant agreement No 101134575. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.