

Blue Cloud Fisheries Monitor Demonstrator

Webinar on Fisheries Data Management IFREMER, IRD, FORTH, FAO





Agenda

- 14:00 Opening and introduction (Trust-IT Services F. Spagnoli)
- 14:10 Fisheries Atlas Objectives and tasks (IRD J.Barde)

- 14:40 Conclusion and Discussion (A. Ellenbroek & F. Spagnoli)



Blue-Cloud vision

Blue-Cloud aims to become the environment for the "Blue community", offering access to an unprecedented wealth of multi-disciplinary data resources and added-value services for the benefit of the future marine research and blue economy landscapes.

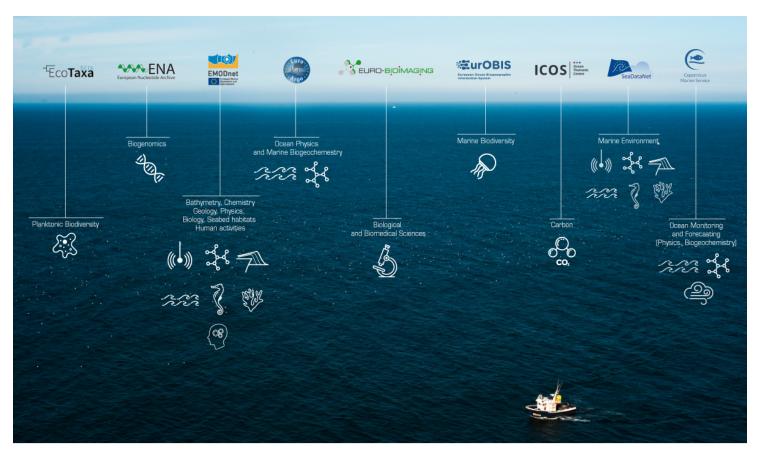


Blue-Cloud mission





A smart federation of existing data, computing, Blue-Cloud storage resources & analytical services









- a Data Discovery and **Access Service** to facilitate sharing with users of multidisciplinary datasets
- a Blue-Cloud Virtual **Research Environment** to facilitate the orchestration of computing and analytical services



Blue-Cloud in the policy context

- Blue-Cloud Roadmap to 2030
- European Open Science Cloud (thematic clouds)
- SDGs & UN Decade of Ocean Science for sustainable development
- EU Green Deal
- Digital Twin Ocean





Early results

- D2.1 Blue Data Infrastructures
- D3.1 Demonstrator general technical requirements
- Blue-Cloud Gateway (workspace)
- Blue-Cloud website <u>www.blue-cloud.org</u>
- ❷ Blue-Cloud workshop and presentation at the All-Atlantic Ocean Research Forum (Feb 2020) https://www.youtube.com/watch?v=iYQirONAyYU
- Blue-Cloud position paper on EOSC (March 2020)
- Blue-Cloud featured in <u>ECO Magazine</u> (May-June 2020)
- First Blue-Cloud webinar introducing the five demonstrators (June 2020)
- Activated collaborations with other projects & thematic clouds Synergies page
- Set up an <u>External Stakeholder Experts Board</u> (ESEB), featuring key representatives from the marine research community and the policy landscape
- Blue-Cloud featured in <u>ERCIM News special issue on Blue Growth</u> (October 2020)



Five real-life demonstrators

#3.FW

Zoo- and Phytoplankton EOV products

Biodiversity

Plankton Genomics

Environment

Marine Environmental Indicators

Fisheries

Fish, a matter of scales

Aquaculture



Aquaculture Monitor



Five demonstrators



Zoo- and Phytoplankton EOV products

Biodiversity



Plankton Genomics

Environment



Marine Environmental Indicators

Fisheries



Fish, a matter of scales

Aquaculture



Aquaculture Monitor



Agenda

- Opening and introduction (Trust-IT Services F. Spagnoli)
- Demonstrator Vision and Mission (FAO A.E. Nieblas)
- Fisheries Atlas Objectives and tasks (IRD J.Barde)
- Fisheries Atlas Activities and Tools (FAO E.Blondel)
- Global Record of Stocks and Fisheries (FORTH Y. Marketakis, FAO A.Gentile)
- Conclusion and Discussion (A. Ellenbroek & F. Spagnoli)

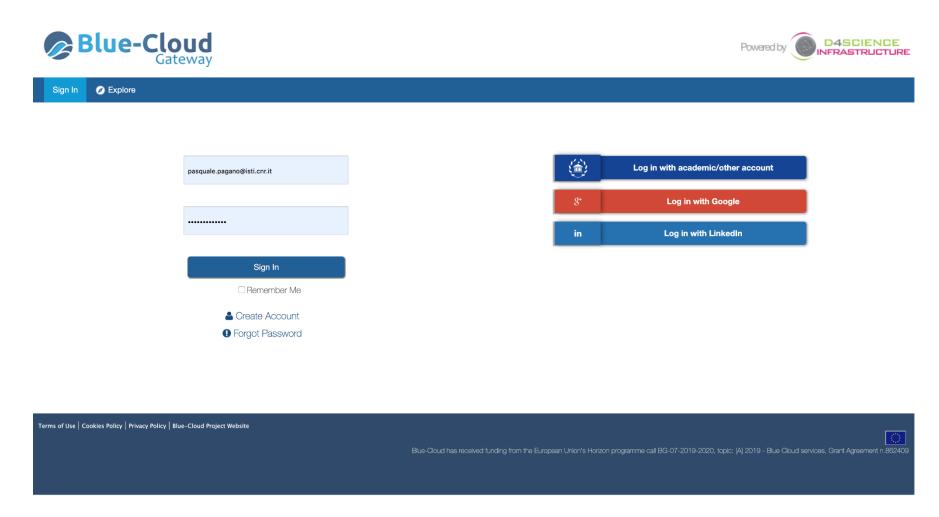


Fisheries Atlas

- Managed by FAO & IRD & FORTH
- Ambition: deliver tools to manage fisheries data and disseminate aggregate statistics and maps
- Expand analytic capabilities
 - Understand fisheries effort and contribution to food-systems;
 - Fisheries relation to the environment and biodiversity;
- Provide robust & replicable Open and FAIR environments
 - Databases, Open and Fair Mapviewers, Semantic KB, Registries



VREs to manage data...



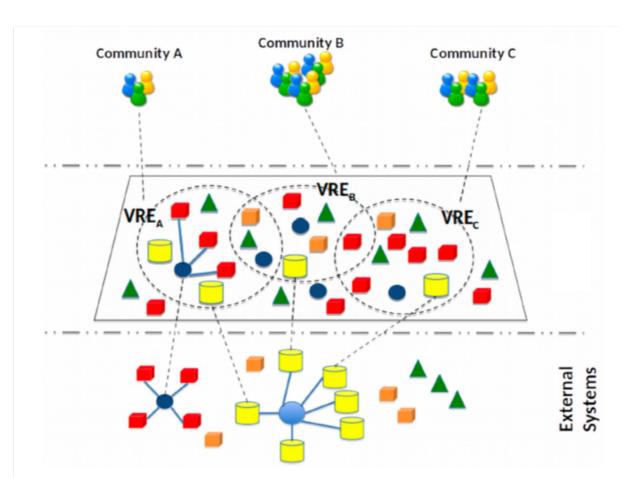


In a data Infrastructure...

Autonomous components

Distributed

Open



Assante, M., Candela, L., Castelli, D., Cirillo, R., Coro, G., Frosini, L., Lelii, L., Mangiacrapa, F., Marioli, V., Pagano, P., Panichi, G., Perciante, C., Sinibaldi, F. **The gCube System: Delivering Virtual Research Environments as-a- Service**. Future Generation Computer Systems (Vol. 95)



With loads of flexible integration A swiss army knife for nerds



Method available as-a-Service,



Invoked via communication standards (WPS)



Higher computational capabilities



Automatic creation of a Web interface



Provenance management



Storage of results on a high-availability system



Collaboration and sharing



Re-usability, Reproducibility, Repeatability, also from other software



Demonstrator Objectives Strengthen methodologies and tools in support of Fisheries Atlases

- An Atlas; Not for FAO, but for global community
 - A comprehensive overview of fisheries, supporting SDGs 2 and 14
 - Build and expand on the FAO/IRD H2020 Tuna Atlas wider range of stocks and products
- Methodology: Development through co-creation of a community
 - 1. Identify available data, analytics and indicators
 - 2. Standardise inputs to create standardised outputs
 - 3. Develop cross-platform visualisations and analytical services
 - 4. Disseminate the estimate statistics; OpenFairViewer and GRSF
- Tool: Virtual Lab as a Fisheries Atlas:
 - Start from the FAO/IRD H2020 Tuna Atlas
 - Continue integration in Blue Cloud Demonstrator: #4 –Fish, a matter of scales



Agenda

- Opening and introduction (Trust-IT Services F. Spagnoli)
- Demonstrator Vision and Mission (FAO A.E. Nieblas)
- Fisheries Atlas Objectives and tasks (IRD J.Barde)
- Fisheries Atlas Activities and Tools (FAO E.Blondel)
- Global Record of Stocks and Fisheries (FORTH Y. Marketakis, FAO A.Gentile)
- Conclusion and Discussion (A. Ellenbroek & F. Spagnoli)



Objectives & Tasks

WP3 Fisheries Atlas - Julien Barde (IRD)

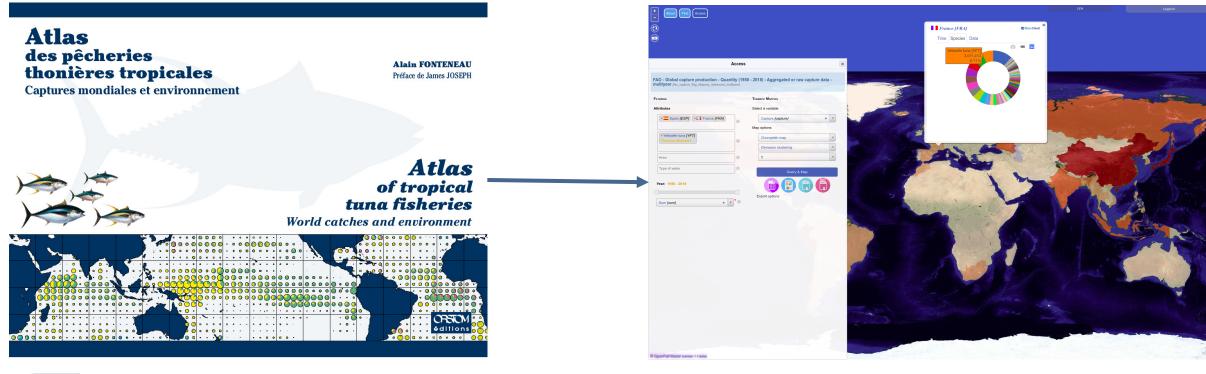


The picture can't be displayed.





Tuna Atlas: more than 20 years





No data, static outputs (GIS plots), made on a PC

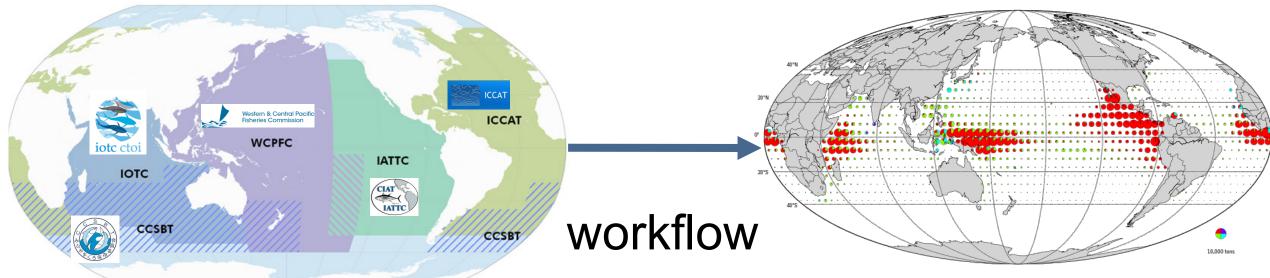


Open Data, interactive outputs (maps & dashboards) built on the fly Made on a VRE



Tuna Atlas: a typical use case for fisheries data flows





5 Tuna Regional Fisheries Management Organizations (tuna RFMOs) => 5 regional data providers / ocean areas

A single data provider (FIRMS) and global datasets





Objectives and Tasks (1)



- Build a generic tool to set up any fisheries atlas:
 - Proof of concept with data cubes and tuna fisheries use case
 - Main variables (dimensions) & indicators:
 - Catches (lat, lon, time, flag, gear, fads, species...),
 - Effort (lat, lon, time, flag, gear, fads...)
 - Upcoming variables and indicators (not only data cubes):
 - catch-at-size, tagging, biological data,
 - trajectories data: AIS / VMS, FADs...
 - environmental parameters (sst, chlo, wind...): e.g. climate change effect
 - Atlas => visualize & explore data and indicators:
 - maps,
 - interactive dashboards
 - maps and dashboards should be linked/interconnected



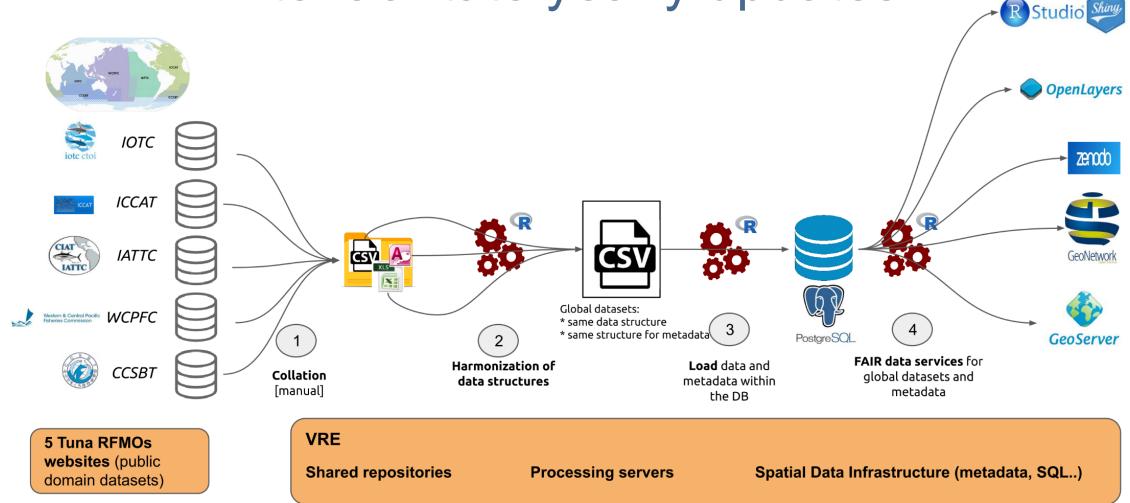
Objectives and Tasks (2)



- Comply with best practices for data management:
 - Make Fisheries data FAIR / a "ready to go" DMP for fisheries
 - Workflow to be managed by R programming language (the most relevant for this community)
 - Open & reproducible science approach: open data / source, standards...
- Use a VRE (EOSC / BlueCloud) to run the workflow and host applications:
 - Deploy a spatial data infrastructure with multiple software components
 - Process (R scripts): RStudio, Jupyter notebooks (tutorials), Dataminer
 - Host visualization tools: web mapping, interactive dashboards (shiny apps)



Tuna Atlas: a FAIR workflow to facilitate yearly updates











- Fisheries data FAIRification is achieved when making data:
- Findable => provide discovery metadata to foster discovery through various searching engines
- Accessible => through well known data formats & access protocols
- Interoperable => implement widely used standards in different domains (OGC, Datacite, CWP, TDWG)
- Reusable => usage metadata to understand how data can be used, open data formats...



Fisheries Atlas data are Findable





- Data discovery requires:
- Syntactic interoperability by implementing multiple metadata standards (OGC 19115, Datacite, Dublin Core):
 - Metadata mapping management
 - DOIs assignment (eg Zenodo, B2Share...)
- Semantic interoperability by using controlled vocabularies (URIs & keywords) rather than free text
- Metadata exposed or harvested by different catalogs:
 - Local or projects catalogs: geonetwork, CKAN...
 - Global search engines: DataCite Search, Google Dataset Search...



Fisheries Atlas data are Accessible





- Data access requires:
- Dataset found with a search engine can be browsed and filtered before downloading by exploring:
 - Spatial dimension with a Web mapping interface
 - Other dimensions with a set of indicators gathered on an interactive dashboard
- Main data formats & related protocols relevant for the community are used:
 - Simple download (by http) of widely used formats (CSV, Shapefile...)
 - Programmatic access with standardized protocols / Web services for (data) scientists, e.g.:
 - SQL access,
 - OGC: GML & WFS, WCS or NetCDF & OPeNDAP => libraries for all languages



Blue-Cloud Fisheries Atlas data are Interoperable

- Data interoperability is a key issue requiring the compliance with mu standards:
- **Current standards:**
 - Dublin Core & DataCite (DOI)
 - OGC for spatial (meta-)data formats & access protocols
 - CWP for fisheries data
- **Next steps:**
 - EML + Darwin Core for ecological datasets
 - SDMX for statistical datasets



Fisheries Atlas data are Reusable





- Data can be reused and cited because we provide :
- Rich usage metadata (beyond discovery metadata) to explain how uata can be used properly:
 - Description of the data structure with a data dictionary
 - Quality (protocols...) and Provenance
 - Ultimately writing a data paper to better describe each dataset
- DOIs to foster data citation:
 - DOI should be displayed in metadata sheets, in the viewer...
 - Proper citation and use with proper licensing
- Open Science: the R workflow (on github) and data can be
 - checked, replicated => Tuna Atlas data provenance is crystal clear.
 - customized by tuning the parametrization of Tuna Atlas R scripts.



Agenda

- Opening and introduction (Trust-IT Services F. Spagnoli)
- Demonstrator Vision and Mission (FAO A.E. Nieblas)
- Fisheries Atlas Objectives and tasks (IRD J.Barde)
- Fisheries Atlas Activities and Tools (FAO E.Blondel)
- Global Record of Stocks and Fisheries (FORTH Y. Marketakis, FAO
 - A.Gentile)
- Conclusion and Discussion (A. Ellenbroek & F. Spagnoli)



- Enable an operational web Fisheries Atlas
 - Gathering multiple datasources...
 - Fisheries datasets (Catch & Effort, surveys, etc)
 - Other relevant data (if FAIRly available): vessel trajectories, tagging/biological/ecological data, environmental data, etc.
 - Handled and maintained:
 - By different institutions
 - In heterogeneous way

data structures, formats, stores, locations



- Set-up a common methodology to enable FAIR data flows
 - Use of international standards for (meta)data interoperability
 - Coherence with EC INSPIRE directive
 - Cross-domain: broader scope than thematic e-infrastructures



- Our vision
 - No 'FAIRification' possible without standards
 - No real sustainability without standards: methodology first, technology after
 By nature, the technology becomes obsolete; the methodology continues
 - Data access should be metadata-driven: Are we FAIR or just.. F(lying) in the AIR?
 - Standards are a cornerstone for FAIR DMPs and a sustainable EOSC vision
- Backed up by Blue-Cloud strategical objectives
 - ISO 19115/19139 common metadata model / OGC catalogue service for the web (CSW) for discovery
 - OGC data services (WMS, WFS and WCS) for viewing and distribution



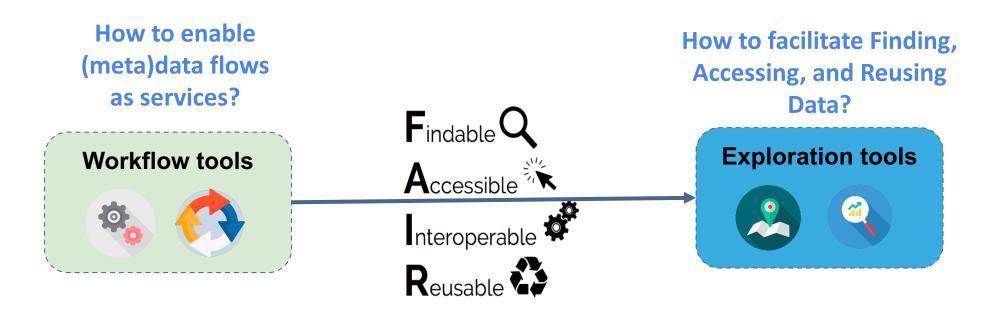
To exploit existing data flows (from Blue-Cloud data providers and beyond) and if not possible, FAIRify some of them, as proof-of-concept.

Multiple constraints and challenges

- IT constraints...
 - Protocol security issues: http vs. https
 - Origin security issues: CORS, Tainted canvas images
- Backed up by institutional constraints / policies (or not)
 - Lack of openess, accounting requirements
 - Lack of willingness to share data using open protocols
 - (Un)FAIR metadata driving: metadata with no data access endpoints, or redirects to user web-forms
- Lack of consideration of the importance of standards
 - Dismissal of standard protocols in favour of trendy technologies
 - Prevalence of system accounting / business considerations



- Build on and Enhance existing tools
 - Open-source software projects
 - Metadata-driven Workflow and Exploration tools



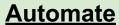


Tools

geoflow R package



- Data processings
- Turnkey actions (meta/data production / publication, data upload, OGC services enabling)



- Avoid web-forms
- Increased reproducibility
- Simplified use of standards
- Multiply (meta)data dissemination endpoints from a single source

Extend

- Plug in-house metadata sources (eg database)
- Plug in-house actions









OpenFairViewer

Metadata driven

 Rely on OGC Catalogue to explore datasets:



- Dataset metadata → Find
- Structural metadata → Query
- Access data through OGC services exposed in metadata:







Visualization

Maps













Share/Export







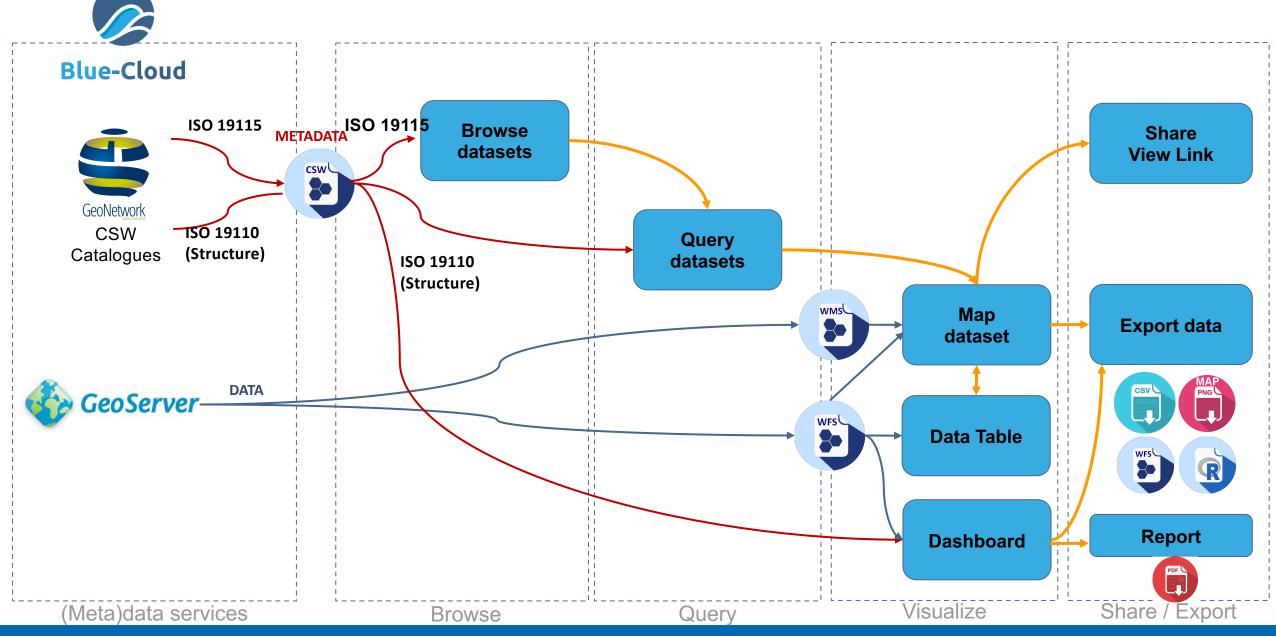


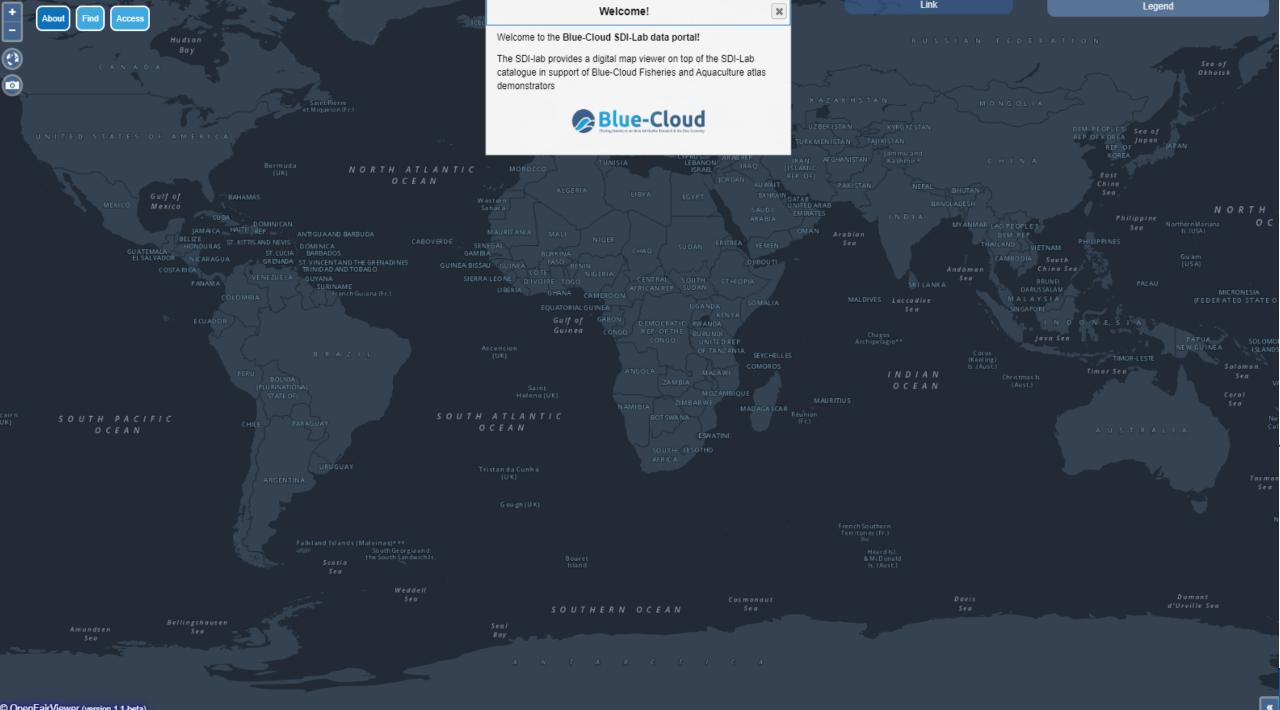


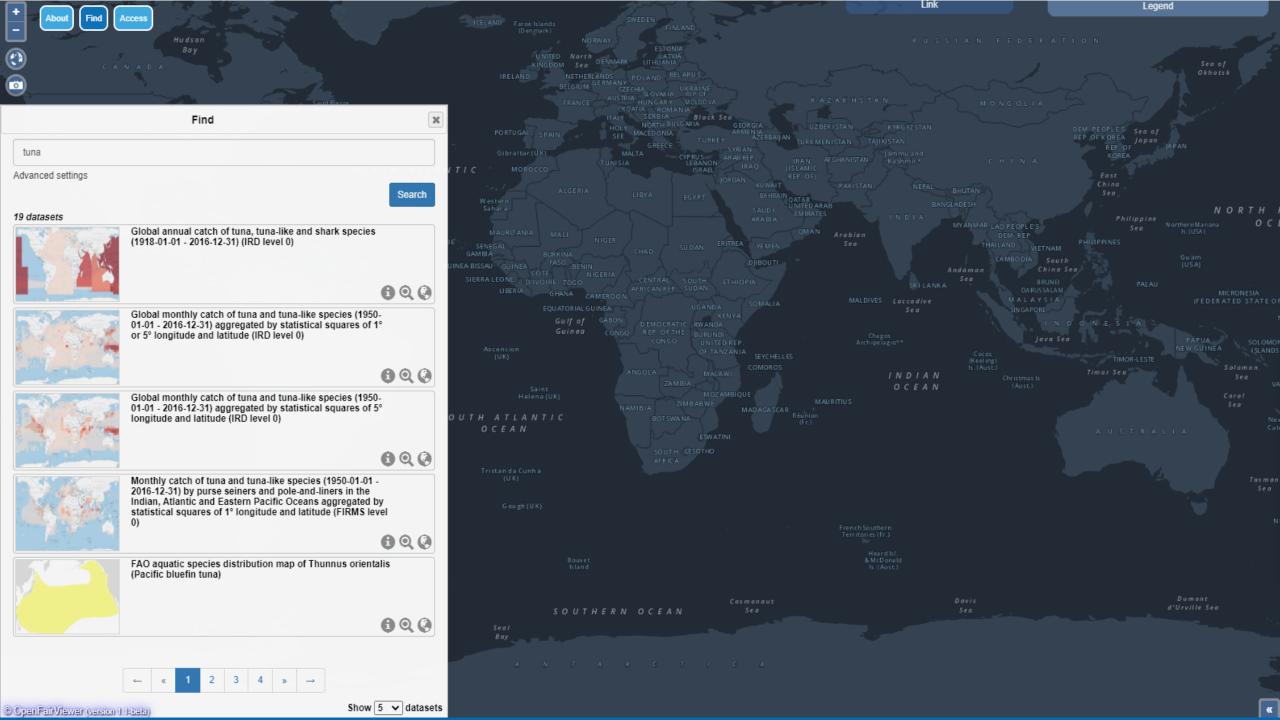


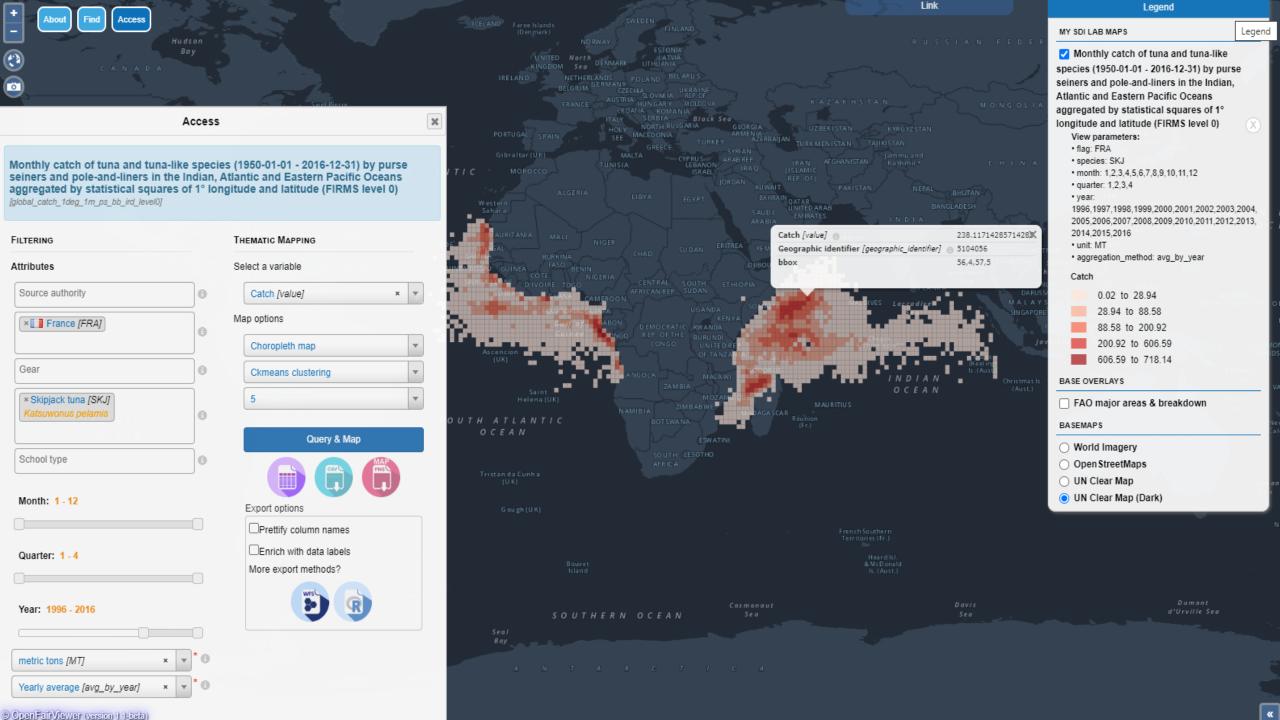
Geoflow – Schematic view of a complete workflow Blue-Cloud CREATE ISO 19115 **PUBLISH** geometa **METADATA** geonapi / ows4R **CSW Catalogues** Contacts CREATE ISO 19110 **OpenFairViewer** geometa R Shiny apps DB UPLOAD **PUBLISH** 🌇 GeoServer Metadata DATA **DBI / RPostgres** geosapi SPATIAL FILE UPLOAD geoflow geosapi Data Entities zenodo **DEPOSIT** RESERVE DOI Registers (Reference data + Actions Software Controlled vocabularies)

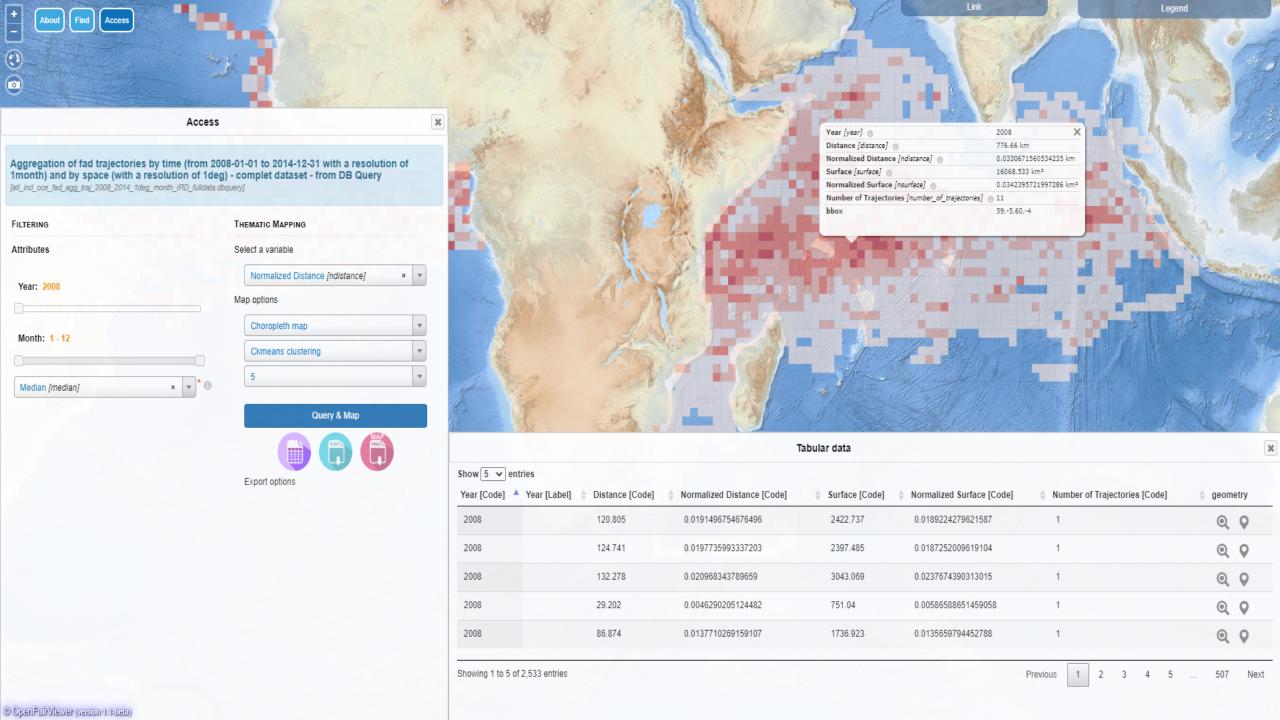
OpenFairViewer – Schematic view



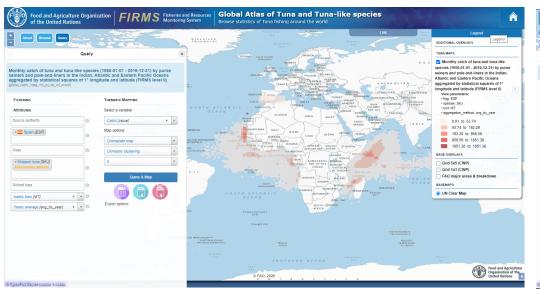


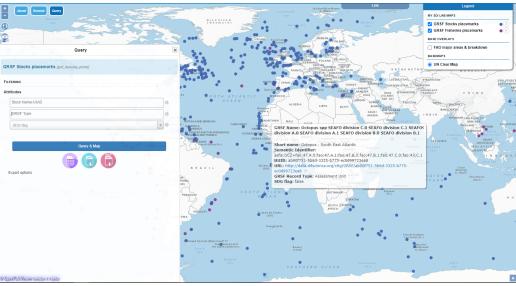


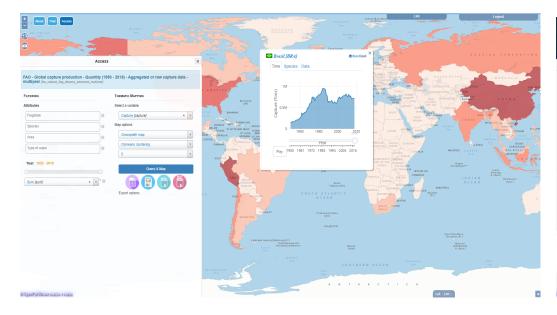


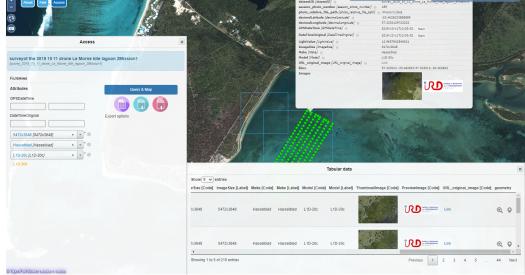














Agenda

- Opening and introduction (Trust-IT Services F. Spagnoli)
- Demonstrator Vision and Mission (FAO A.E. Nieblas)
- Fisheries Atlas Objectives and tasks (IRD J.Barde)
- Fisheries Atlas Activities and Tools (FAO E.Blondel)
- Global Record of Stocks and Fisheries (FORTH Y. Marketakis, FAO A.Gentile)
- Conclusion and Discussion (A. Ellenbroek & F. Spagnoli)

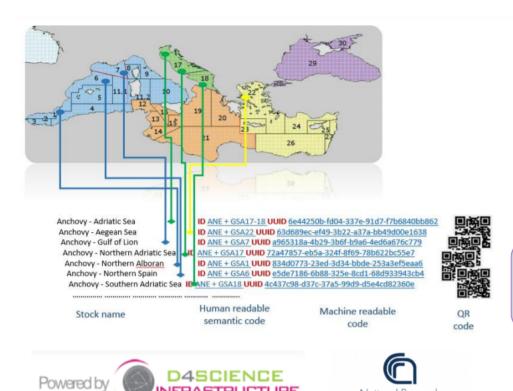


- A database of stocks and fisheries inventoried at global level with standard unique identifiers
- Sources of information: FIRMS Fisheries and Resources Monitoring System, FishSource, and RAM Legacy Stock Assessment Database
- A tool to support:
 - SDG 14.4.1 "Proportion of fish stocks within biologically sustainable levels"
 - Traceability and ecolabelling schemes connecting seafood industries and consumers to the status of stocks and fisheries



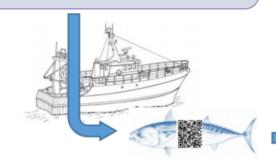
- An aggregator of global knowledge from 3 top institutes
- Data Integration through the construction of a Semantic Knowledge Base
- Publicly accessible; Global Unique Identifiers
- Ready to be **extended** with external resources (e.g. Food data, oceanographic data)
- Re-uses many components from parallel task
 - Blue Cloud enables Reusable, Interoperable, Scalable data sharing







- Stock status (reported at national, regional level) - SDG 14.4.1 indicator
- Traceability schemes











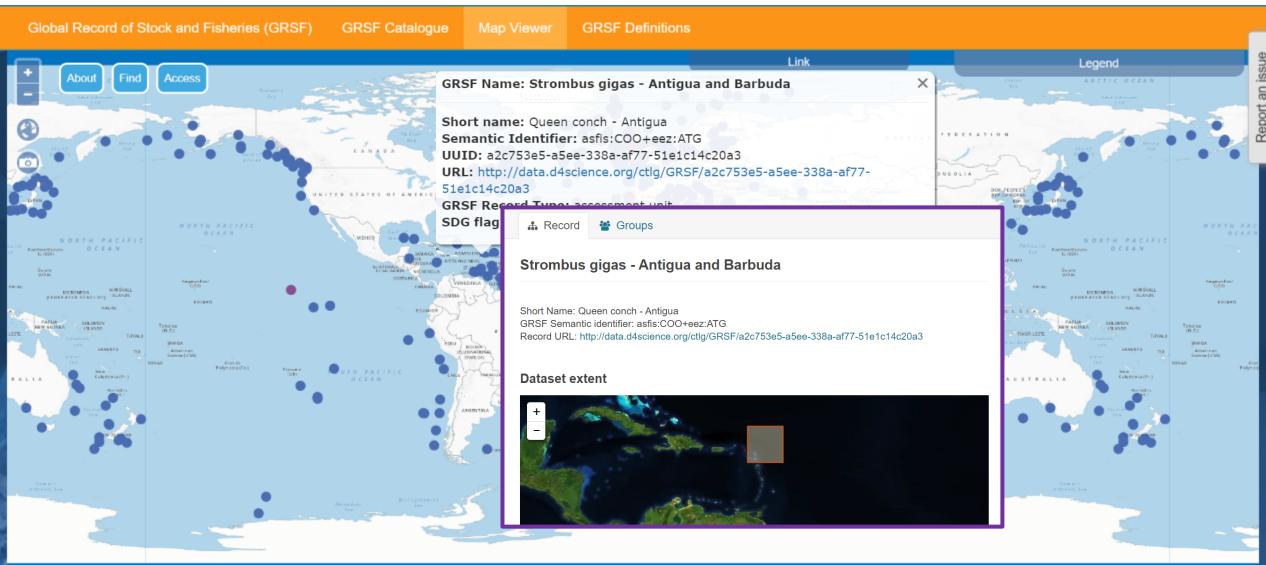


National Research Council of Italy

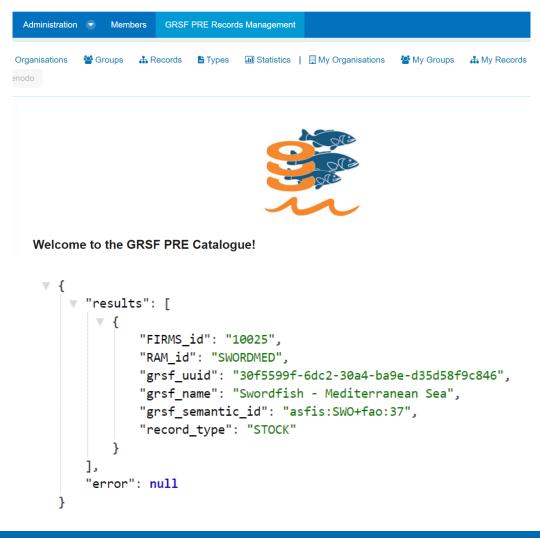












- Testing/Staging new data and software
- Enrichment of the records with additional data from the BlueCloud, e.g. Food data or other demonstrators
- Competency Queries and APIs available and more are in progress...



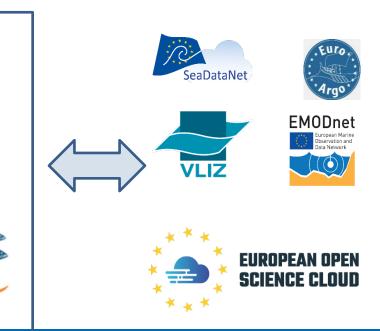
The backbone of GRSF is **GRSF Knowledge Base**, a semantic warehouse, built on top of a top level ontology (MarineTLO), that semantically integrates data from the three main sources of GRSF

BlueCloud Services will "interconnect" GRSF with information from external

Blue-Cloud

RAM legacy

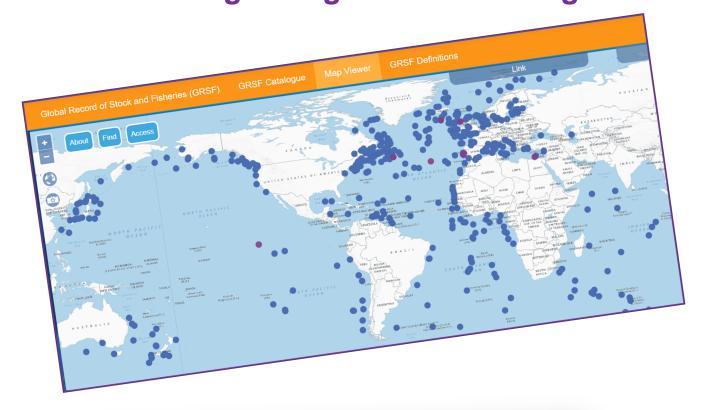
Status and environmental performance of fisheries worldwide





Visit the GRSF at https://i-marine.d4science.org/web/grsf/data-catalogue







Blue Cloud Data Access and Processing Services

- Virtual Research Environment
 - To help researchers to collaborate by providing
 - Harmonize access to various data sets / data sources
 - Discipline-specific tools, such as data analysis, visualization, ...
- The Blue Cloud implementation of VRE's: D4Science
 - Developed by CNR, application to «Open Science» paradigm
 - To allow
 - Data Hosting and curation (metadata)
 - Data Analytics and visualization (e.g. geospatial data)
 - Software Dynamic Deployment
 - Resources (data, software, ...) controlled sharing and accounting



Agenda

- Opening and introduction (Trust-IT Services F. Spagnoli)
- Demonstrator Vision and Mission (FAO A.E. Nieblas)
- Fisheries Atlas Objectives and tasks (IRD J.Barde)
- Fisheries Atlas Activities and Tools (FAO E.Blondel)
- Global Record of Stocks and Fisheries (FORTH Y. Marketakis, FAO A.Gentile)
- Conclusion and Discussion (A. Ellenbroek & F. Spagnoli)

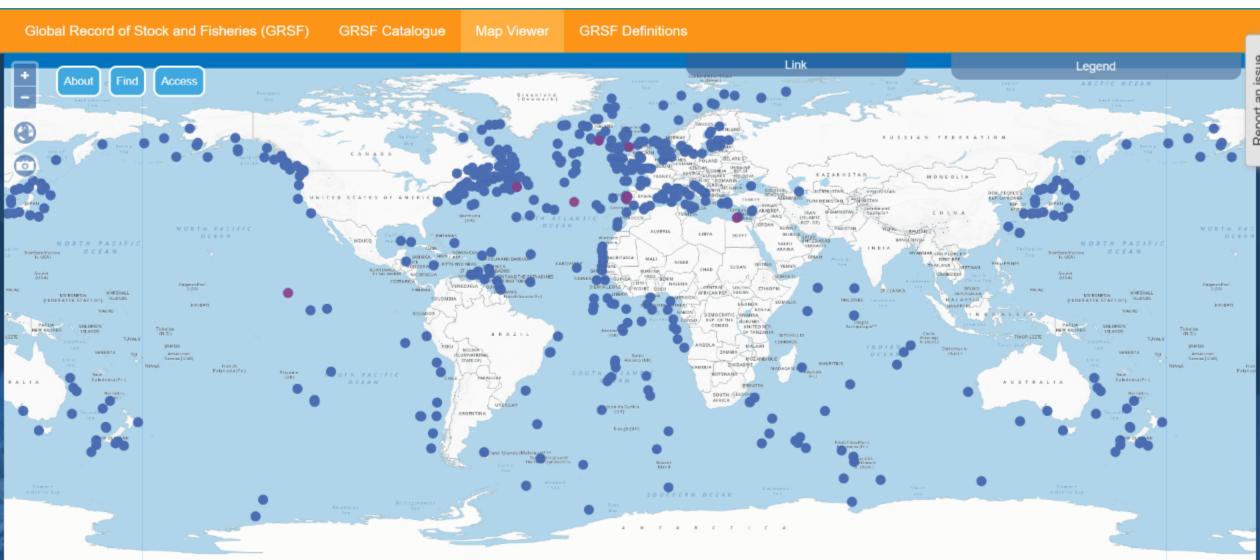


D4 Science main features

- Metadata Catalogue (gCat)
 - to publish and to search collections of metadata for items including data, services, and related information objects
- Workspace (Storage Hub)
 - to browse, to upload and to download user's files and folders
- Data processing platform (named DataMiner)
 - DataMiner online interfaces (harmonized for users)
 - Integration of new methods/algorithms
- Spatial Data Infrastructure
 - Spatial Data Storage and Publishing
 - Spatial Data Discovery and Access
 - SDI Services (OGC Compliant)
- Integrated Users' Authentication and Authorisation framework
- Profile & Social Networking API
 - to get user information and/or to boost content's reach by making easier to share it



End of the FAIRy tale...





D4Science: some useful links

- DataMiner overview: <u>Data Mining Facilities</u>
 - (https://wiki.gcube-system.org/gcube/Data-Mining-Facilities)
- More about Data Miner (developers oriented):
 - Developers website (commonly used APIs, ...): Dev Web Site (https://dev.d4science.org/)
 - Implementation of custom new algorithms for DataMiner: <u>Software Algorithm Importer</u> (https://wiki.gcube-system.org/gcube/Category:Statistical Algorithms Importer)
 - Supported languages for new Methods/Algorithms for DataMiner: Create a new project with SAI (https://wiki.gcube-system.org/gcube/Statistical Algorithms Importer: Create Project)
 - DataMiner online interfaces <u>DataMiner Manager</u> (https://wiki.gcube-system.org/gcube/DataMiner Manager),
 - Web Processing service: Web Processing Service | OGC (https://www.opengeospatial.org/standards/wps).
- Spatial Data Infrastructure capabilities
 - SDI-Service
 - (https://gcube.wiki.gcube-system.org/gcube/SDI-Service),
 - Spatial <u>Data Storage and Publishing</u> (https://gcube.wiki.gcube-system.org/gcube/Spatial Data Storage and Publishing),
 - Spatial Data Discovery and Access
 - (https://gcube.wiki.gcube-system.org/gcube/Spatial_Data_Discovery_and_Access)



Thank you!

Follow Blue-Cloud

Website: <u>www.blue-cloud.org</u> E-mail: <u>info@blue-cloud.org</u>

Twitter: @BlueCloudEU

LinkedIn: Blue-Cloud Org

Join our newsletter: https://bit.ly/2VdZD95