Towards an e-infrastructure in agriculture?

Donatella Castelli, CNR-ISTI

CNR-ISTI

donatella.castelli@isti.cnr.it

Euragri Workshop

Big data in agriculture: consequences for research and research organizations

9 March 2016, Paris

BlueBRIDGE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 675680

Outline

- iMarine & its Virtual Research Environments
- BlueBRIDGE Virtual (Research) Environments
- Final remarks

Terminology (1)

Data infrastructure

"e-Infrastructure" offering services for collection, deposition, storage, preservation, access, retrieval, analysis/mining/processing, publication, etc.

e-Infrastructure

Elecronic platform operated by a responsible entity offering an open set of basic enabling services (including access to resources) to a Community of Practice (CoP). Through the e-Infrastructure the CoP realises economy of scale.

Terminology (2)

The Community of Practice
(the community)

The e-infrastructure
(the operational platform)

The system
(the enabling sw system)

iMarine

Launch an initiative aimed at **establishing and operating an e-infrastructure** contributing to the implementation of the principles of the **Ecosystem Approach to Fisheries Management and Conservation of Marine Living Resources**

With the support of European Commission

SEVENTH FRAMEWORK PROGRAMME

iMarine CoP needs

How to facitate the implementation of the an **Ecosystem Approach to fisheries management and conservation of marine living resources**?

Knowledge generation

- Access to large amount of heterogeneous, distribuited, across-domain data
- Rich data mining, analysis and processing capabilities

Interdisciplinary & multifacets collaboration

at the local, national, regional and international levels

Data Infrastructure for the iMarine CoP

Services
“in the style of cloud”

Leveraging others' resources

D4SCIENCE INFRASTRUCTURE

Leveraging others' resources

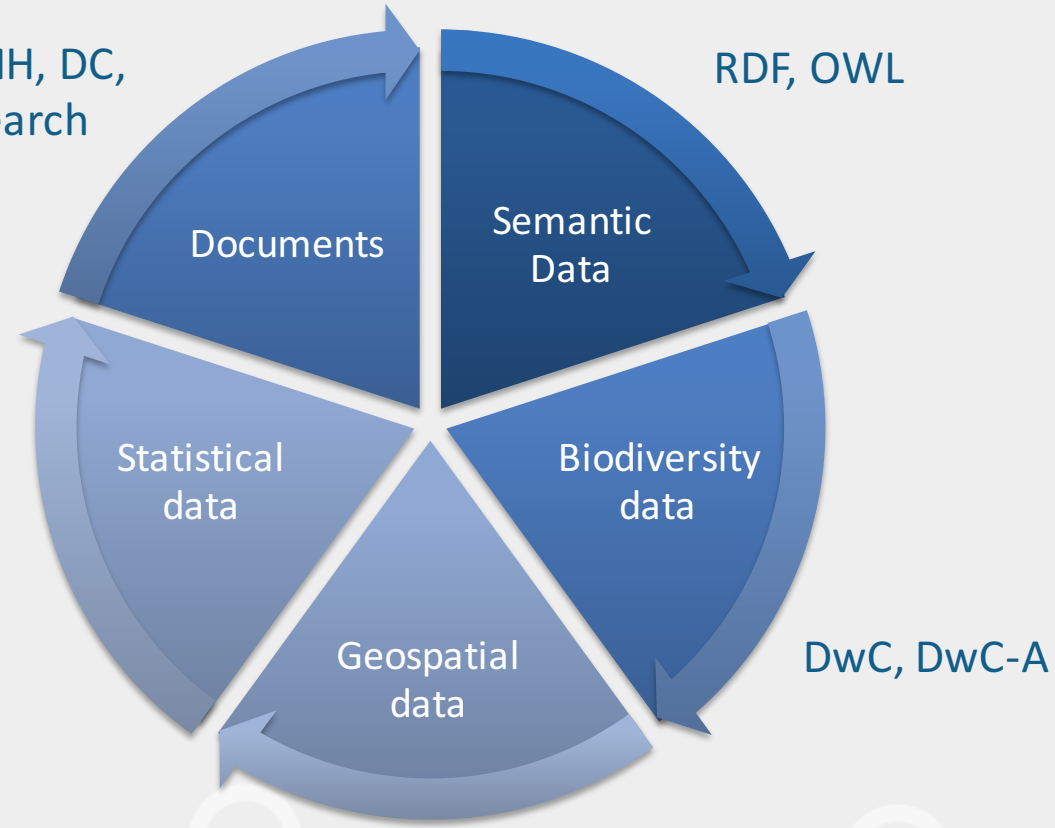
Own resources

Third-party providers

Hetogeneous datasets

Services

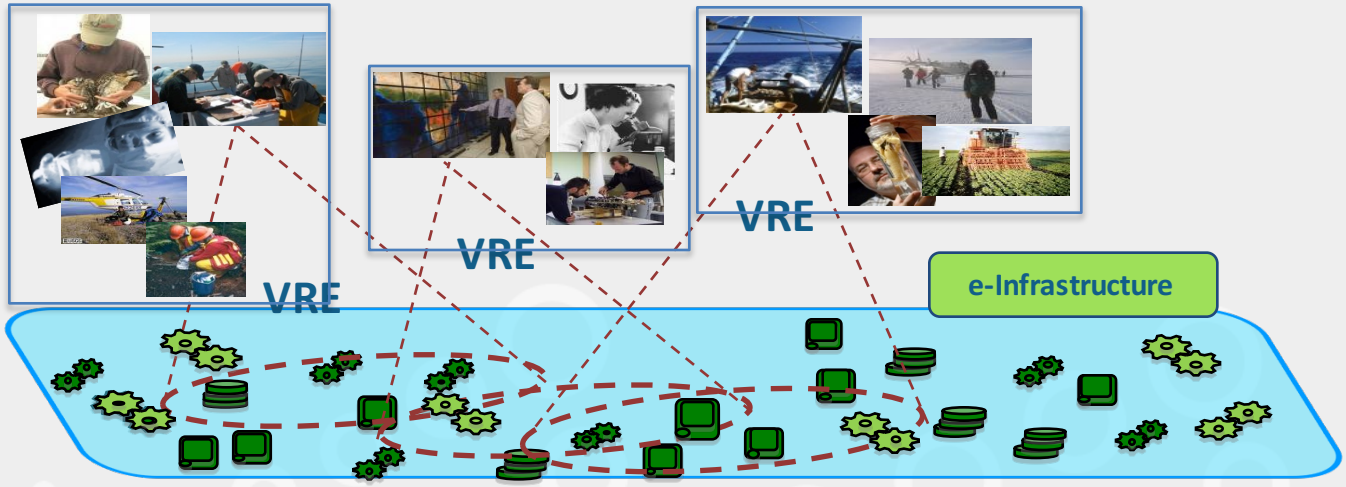
- Collaboration
- Sharing
- Publication
- Storage
- Processing
- Analysis
- Modeling
- Integration
- Enrichment
- Transformation
- Access
- Discovery
- Collection



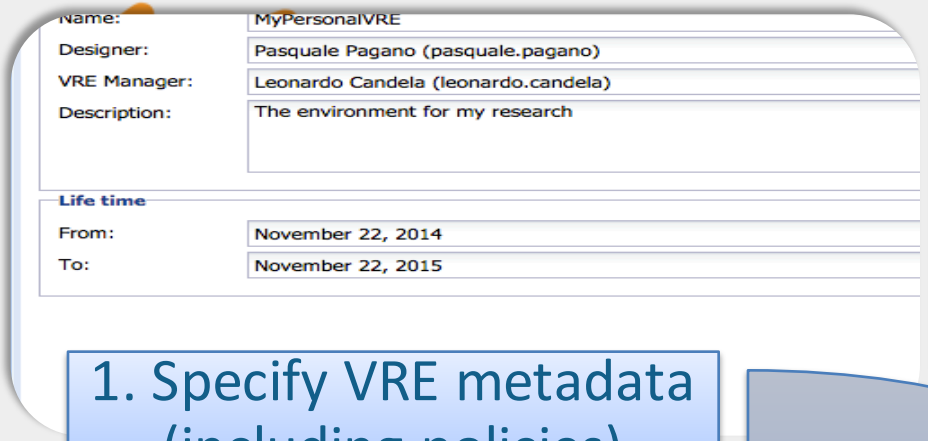
VRE as-a-Service

Virtual Research Environment

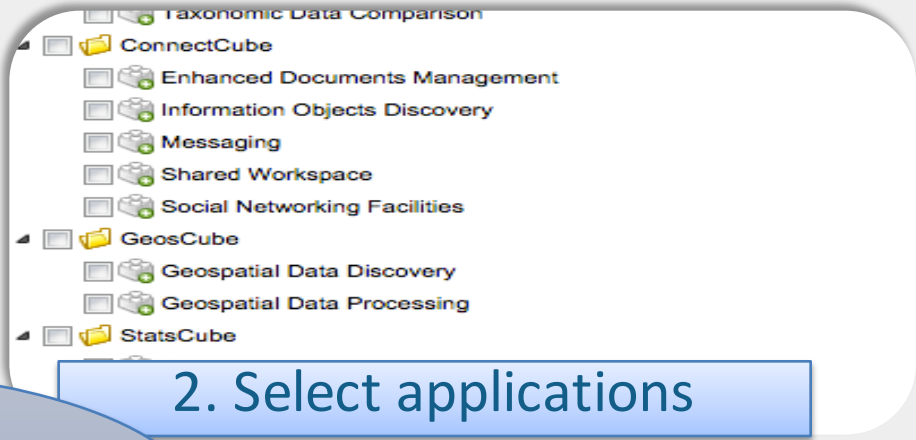
- **web-based working environment**
- providing access to services and resources tailored to serve the needs of a community of practice **in accomplishing a specific goal**
- **open and flexible** with respect to service offering and lifetime
- providing fine-grained **controlled sharing** of both intermediate and final research results
- **Very low cost of creation and operation**



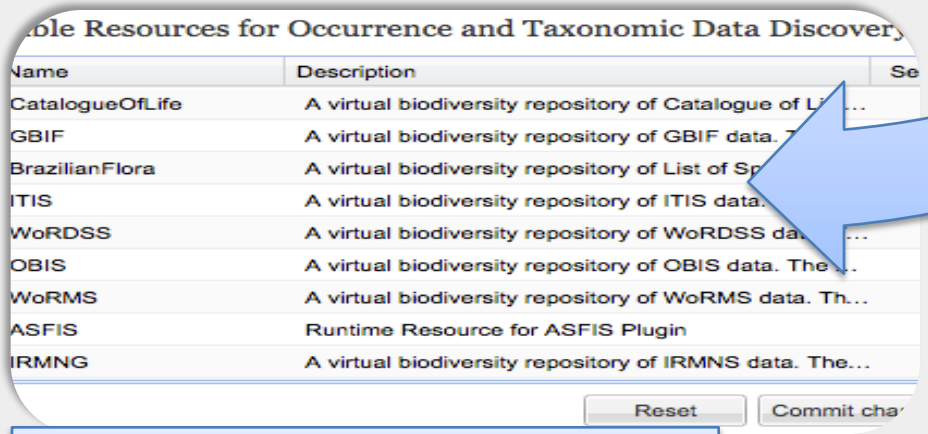
BlueBRIDGE



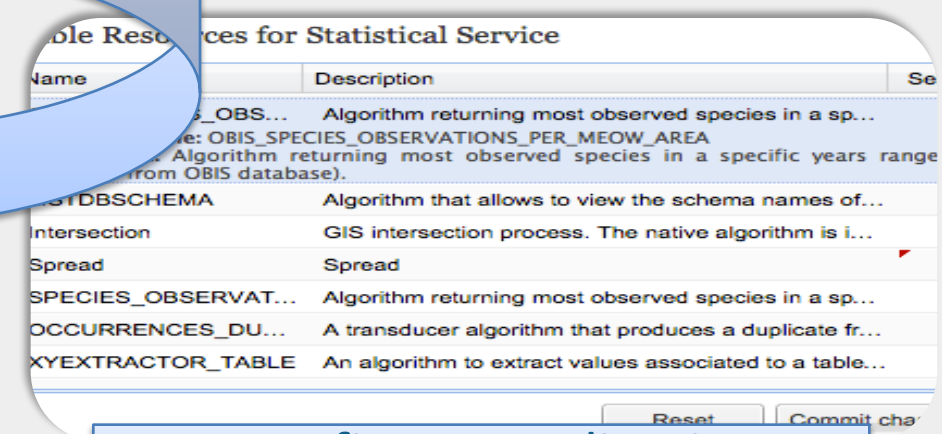
1. Specify VRE metadata (including policies)



2. Select applications



4. Select data collections



3. Configure applications

Hardware setup and software deployment completely hidden

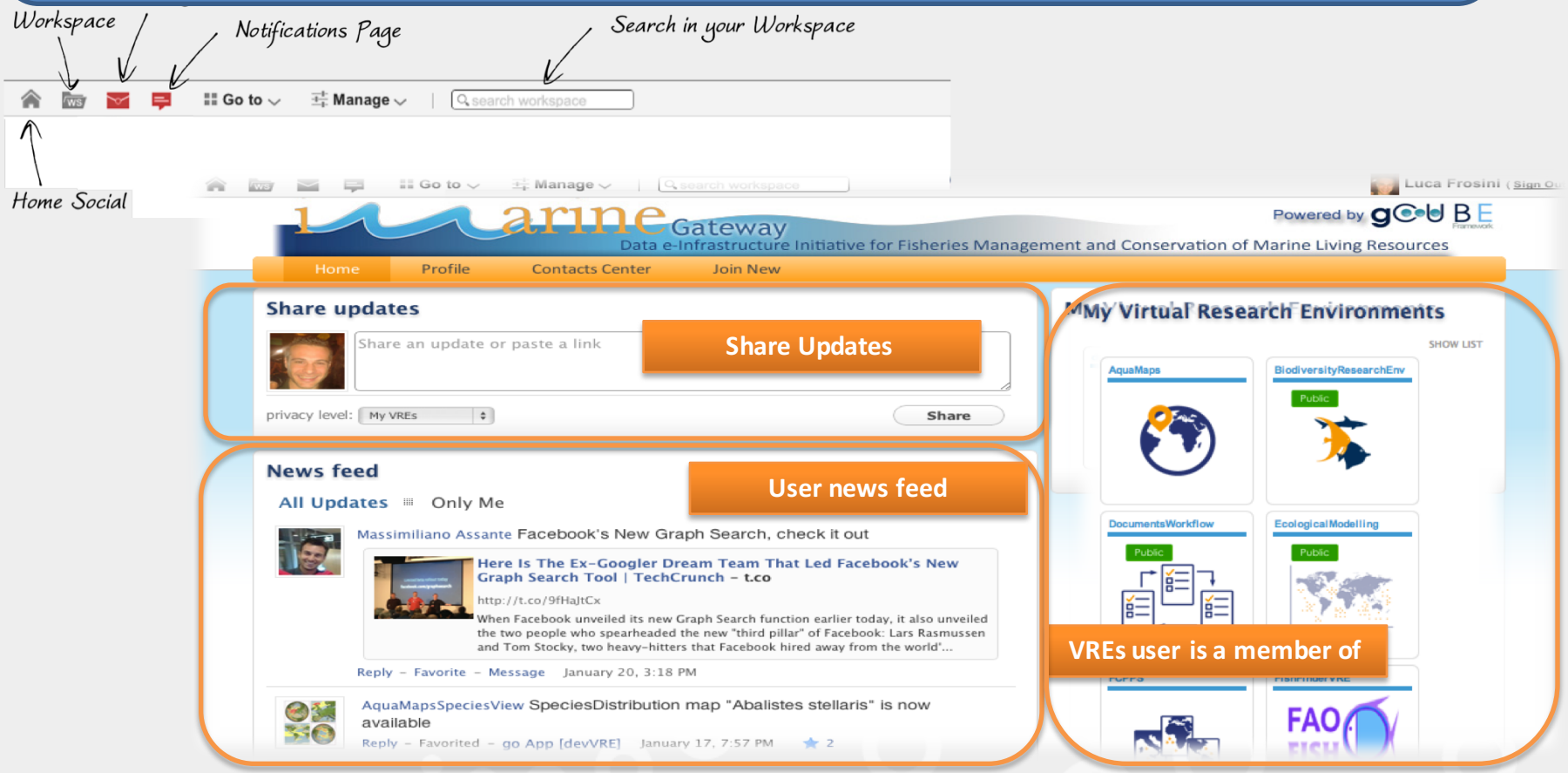
Evolving needs of its users completely supported

Collaborative Environment

A single point to

Get status and updates from applications and other users they are interested in;

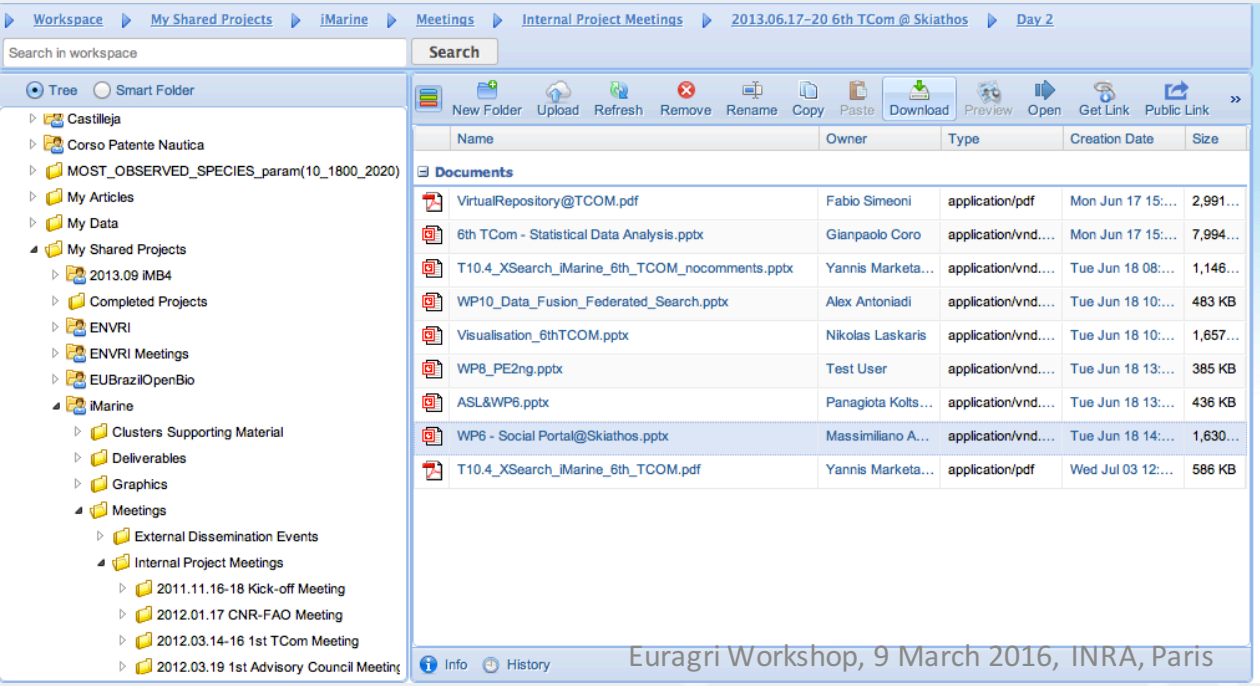
Get notifications about messages, jobs completion, new generated products, etc.



Workspace

A single place to

- Manage data, store and preserve them
- Share data



Workspace Public Link

Create links to files in your Workspace to download them. You can send the links to anyone by pasting them into Workspace Message, your emails, instant messages, etc.

Workspace Share Folders and Files

The quickest way to share something is using the Share Folder. Locate the folder with the files you want to share and then click 'Share'

Workspace Upload Files and Archives

You can upload files in the Workspace in several ways:

1 – Drop your files from Desktop;

2 – Click 'Upload' and Browse Files;

3 – Upload a zip file to unzip directly its content in the Workspace.

Working within a VRE (1)

The user

- Stores data in a personal workspace
- Visualizes and harmonizes data
- Saves the results for further exploitations
- Shares with his/her colleagues

Working within a VRE (2)

The user

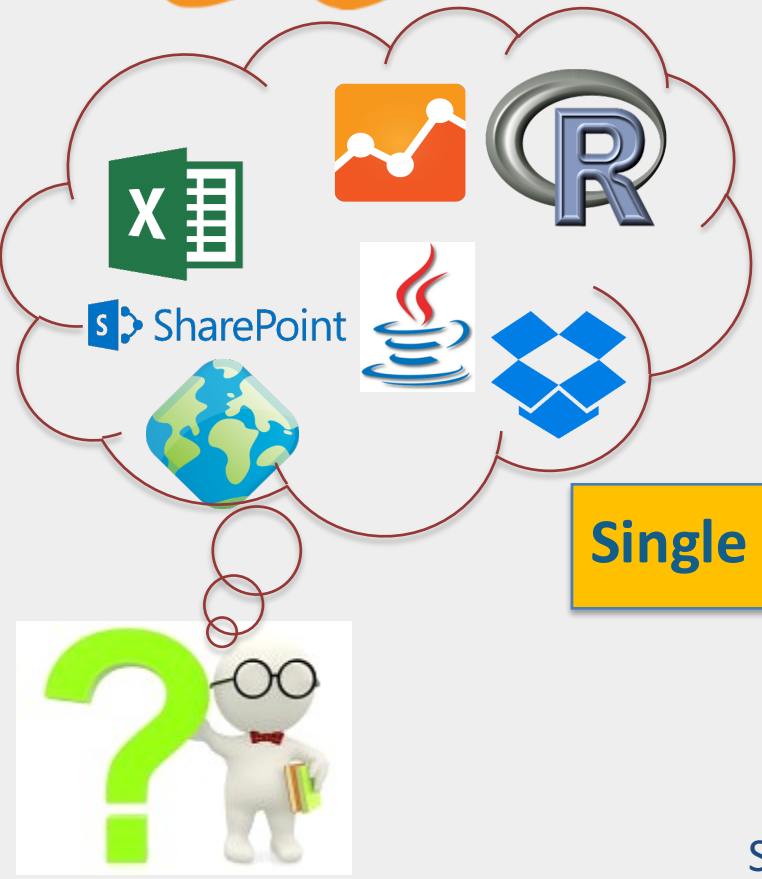
- Stores software in a personal workspace
- Prepares it for execution with a simple interface (one shot process, never to repeat)
- Executes it and analyses the results
- Modifies the code in the workspace
- Shares software and/or results with his/her colleagues

iMarine VREs

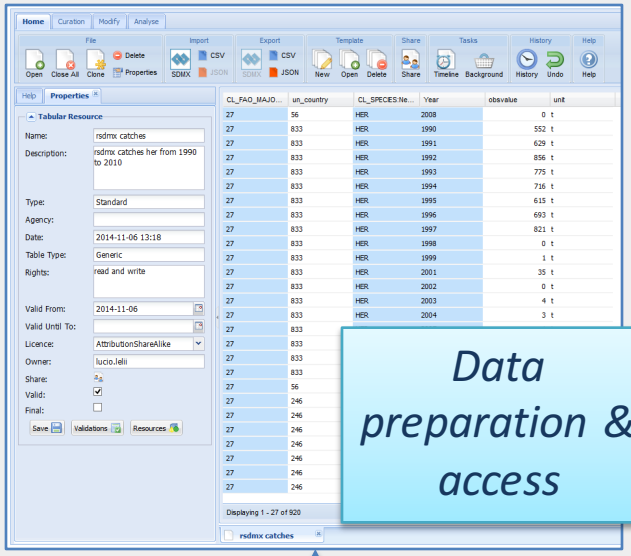
iMarine Gateway
https://i-marine.d4science.org/

- *Public VREs* (used to offer an application environment to a subset of users of a community)
- *Private VREs* (used for experiments, data access and preparation, and data analytics)

Scalable Data Mining VRE



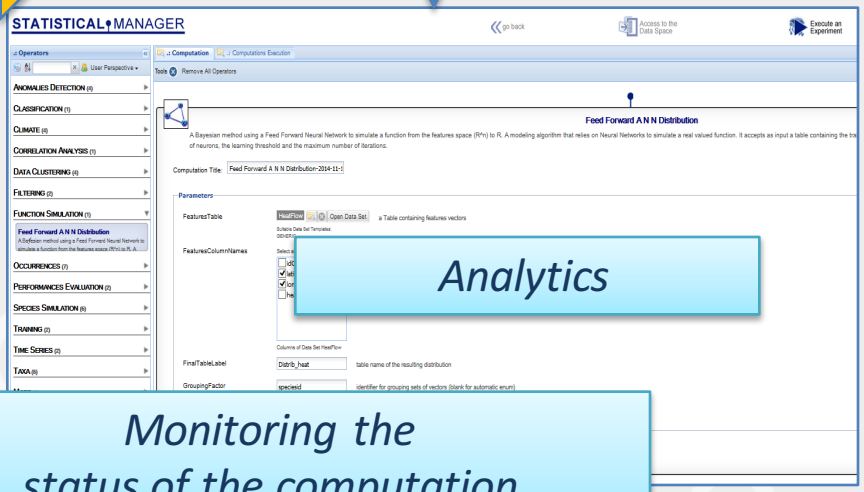
Tabular Data Manager (**TabMan**)



Data preparation & access

Single environment

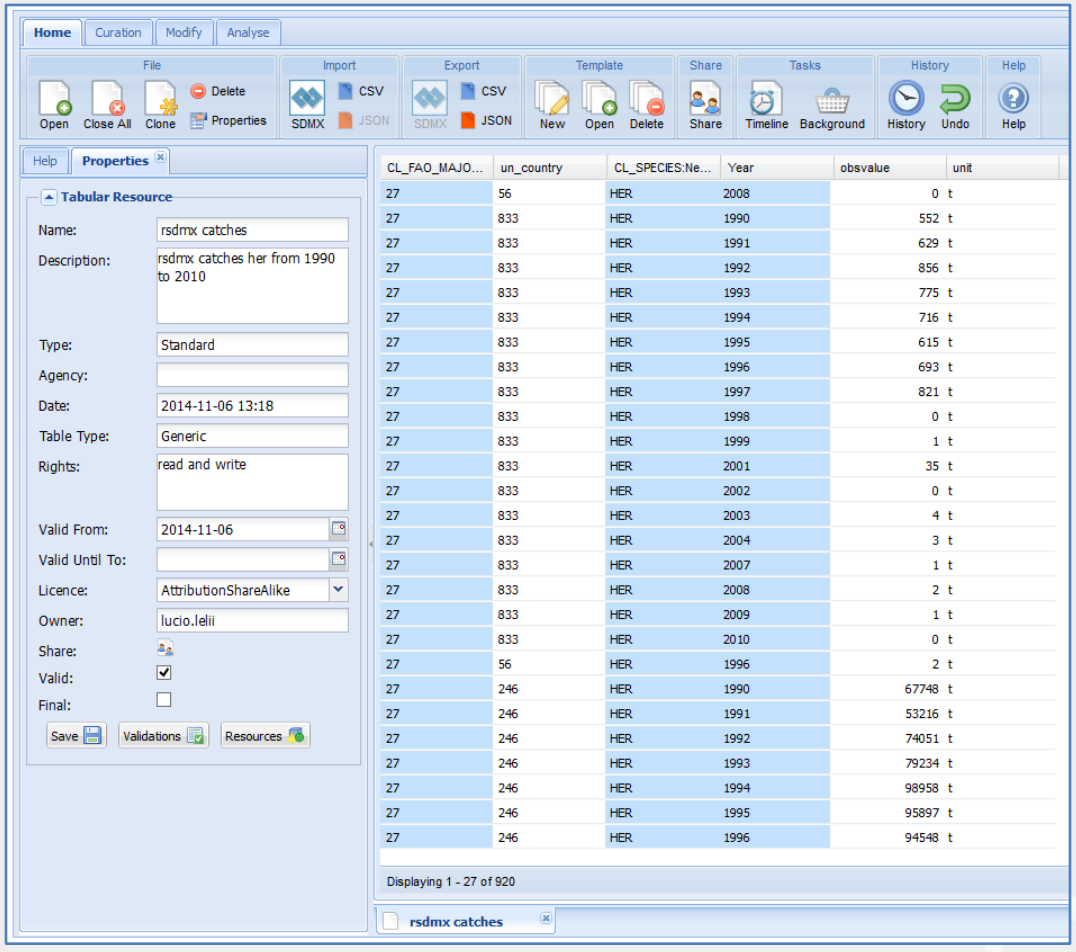
Statistical Manager (**StatMan**)



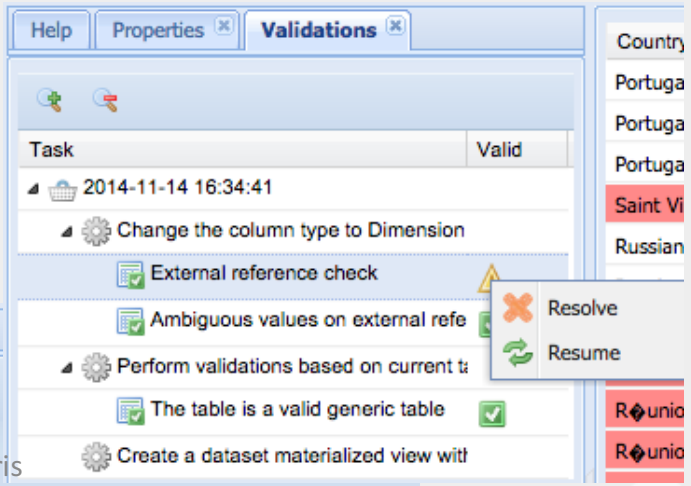
Analytics

Monitoring the status of the computation

Tabular Data Manager (TabMan)



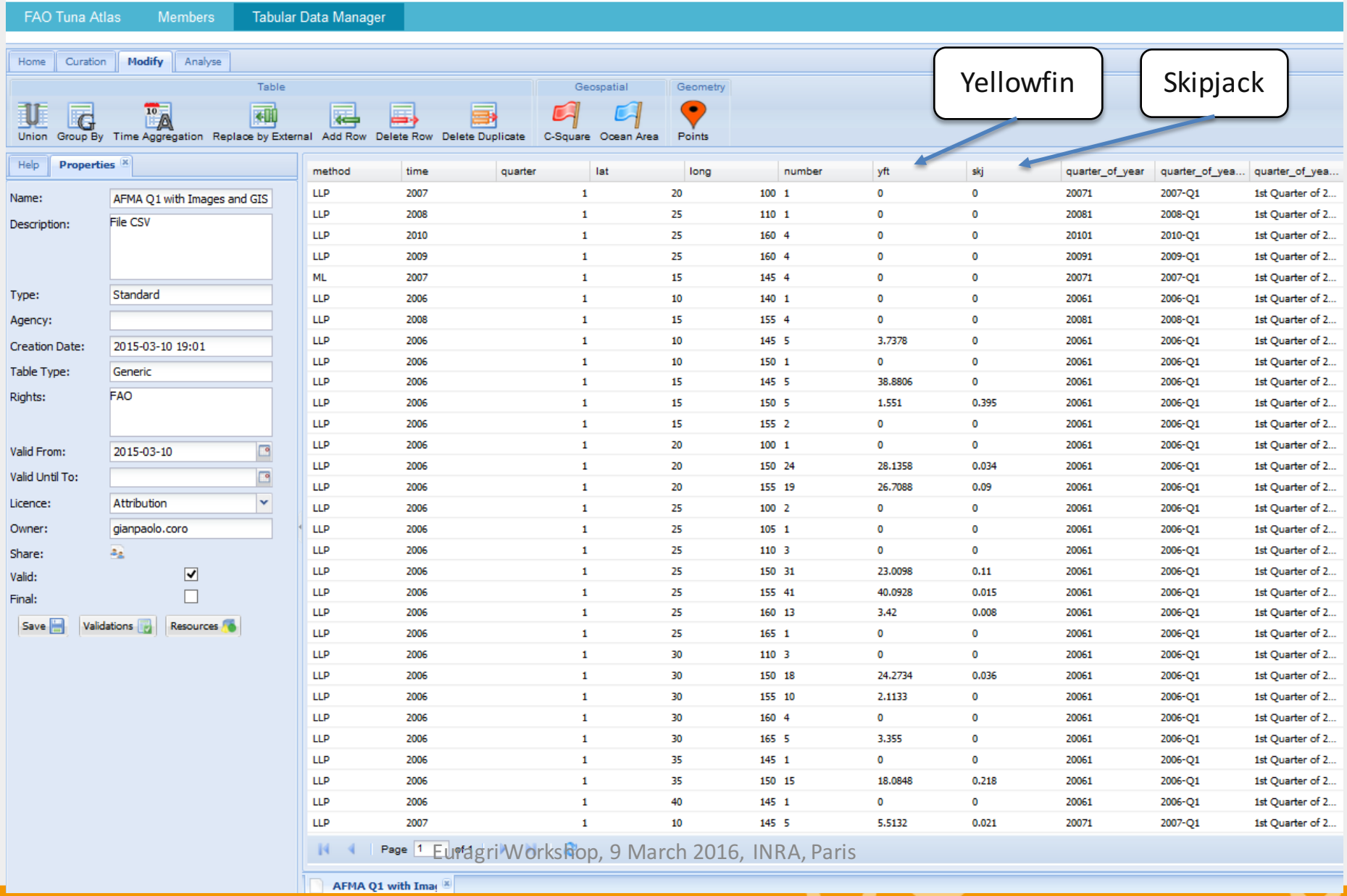
- Manipulates Big Tabular Datasets
- Prepares data for analyses
- Makes data compliant with external code lists
- Visualizes, represents and inspects data



BlueBRIDGE

TabMan

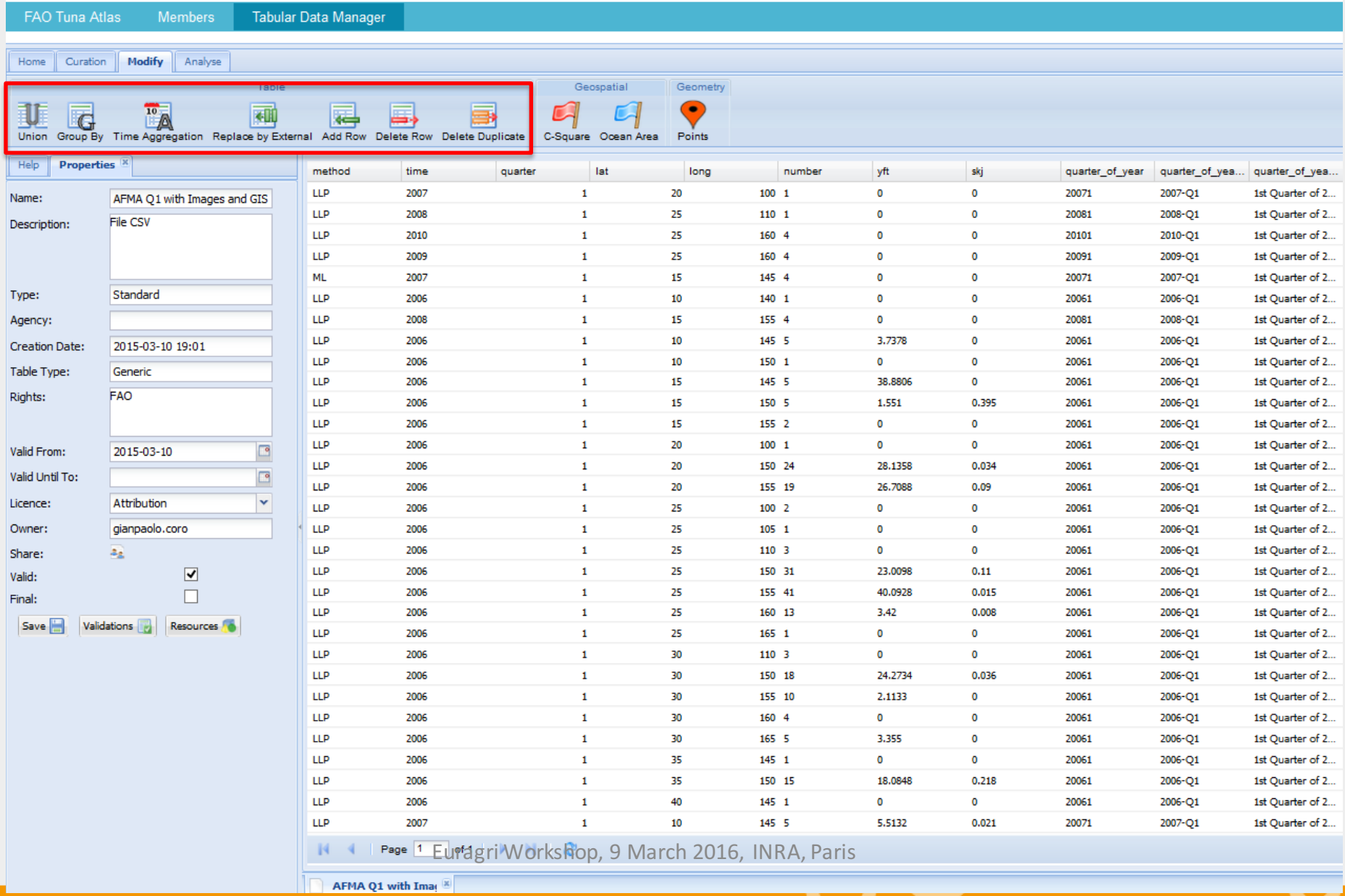
TabMan imports datasets with tuna catch statistics



BlueBRIDGE

TabMan

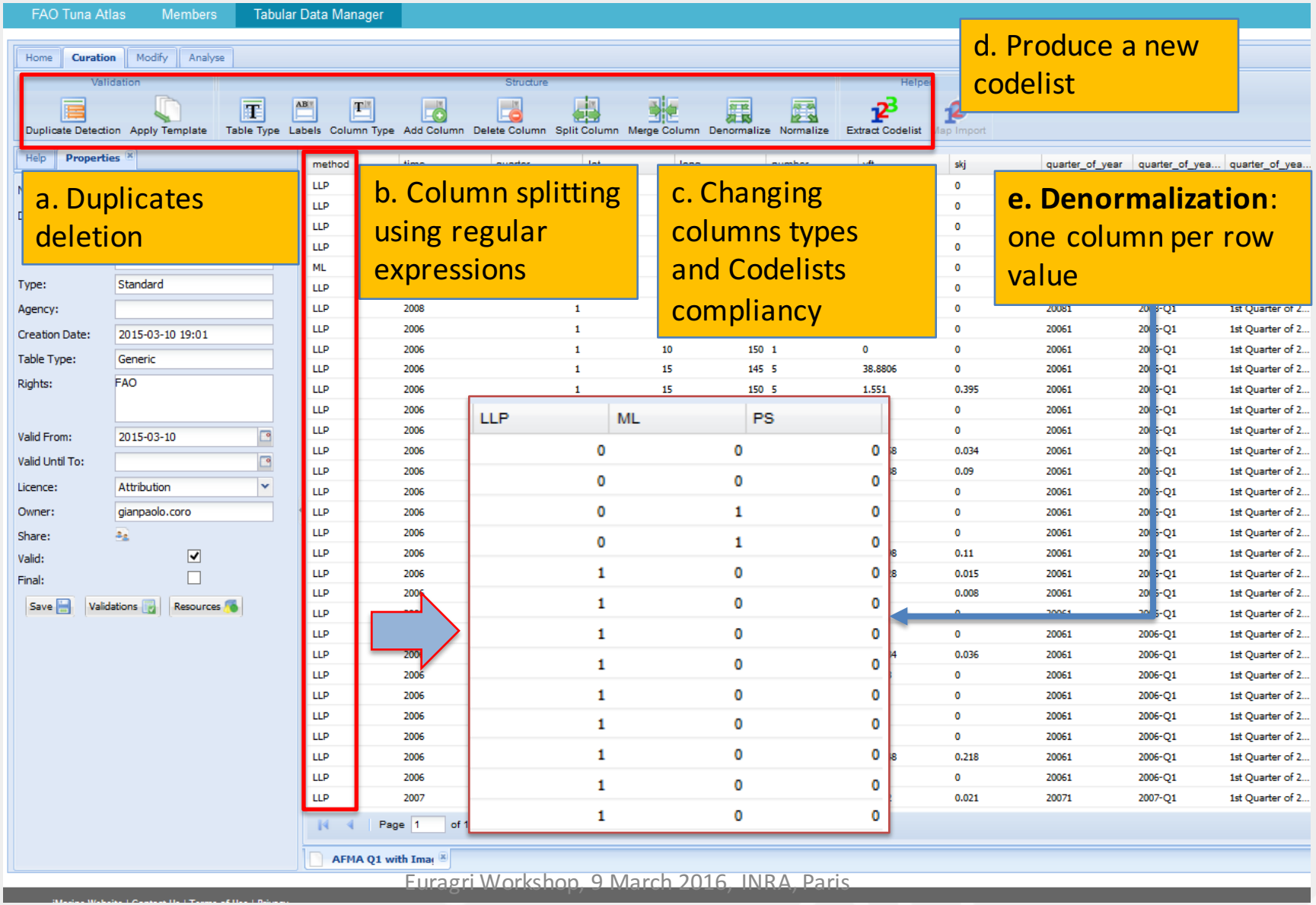
Modifying columns



BlueBRIDGE

TabMan

Data Curation



BlueBRIDGE

TabMan

Adding Geometries

FAO Tuna Atlas | Members | Tabular Data Manager

Home | Curation | **Modify** | Analyse

Table: Union, Group By, Time Aggregation, Replace by External, Add Row, Delete Row, Delete Duplicate

Geospatial: C-Square, Ocean Area

Geometry: Points

Help | **Properties**

Name:	AFMA Q1 with Images and GIS
Description:	File CSV
Type:	Standard
Agency:	
Creation Date:	2015-03-10 19:01
Table Type:	Generic
Rights:	FAO
Valid From:	2015-03-10
Valid Until To:	
Licence:	Attribution
Owner:	gianpaolo.coro
Share:	
Valid:	☑
Final:	☐

Save | Validations | Resources

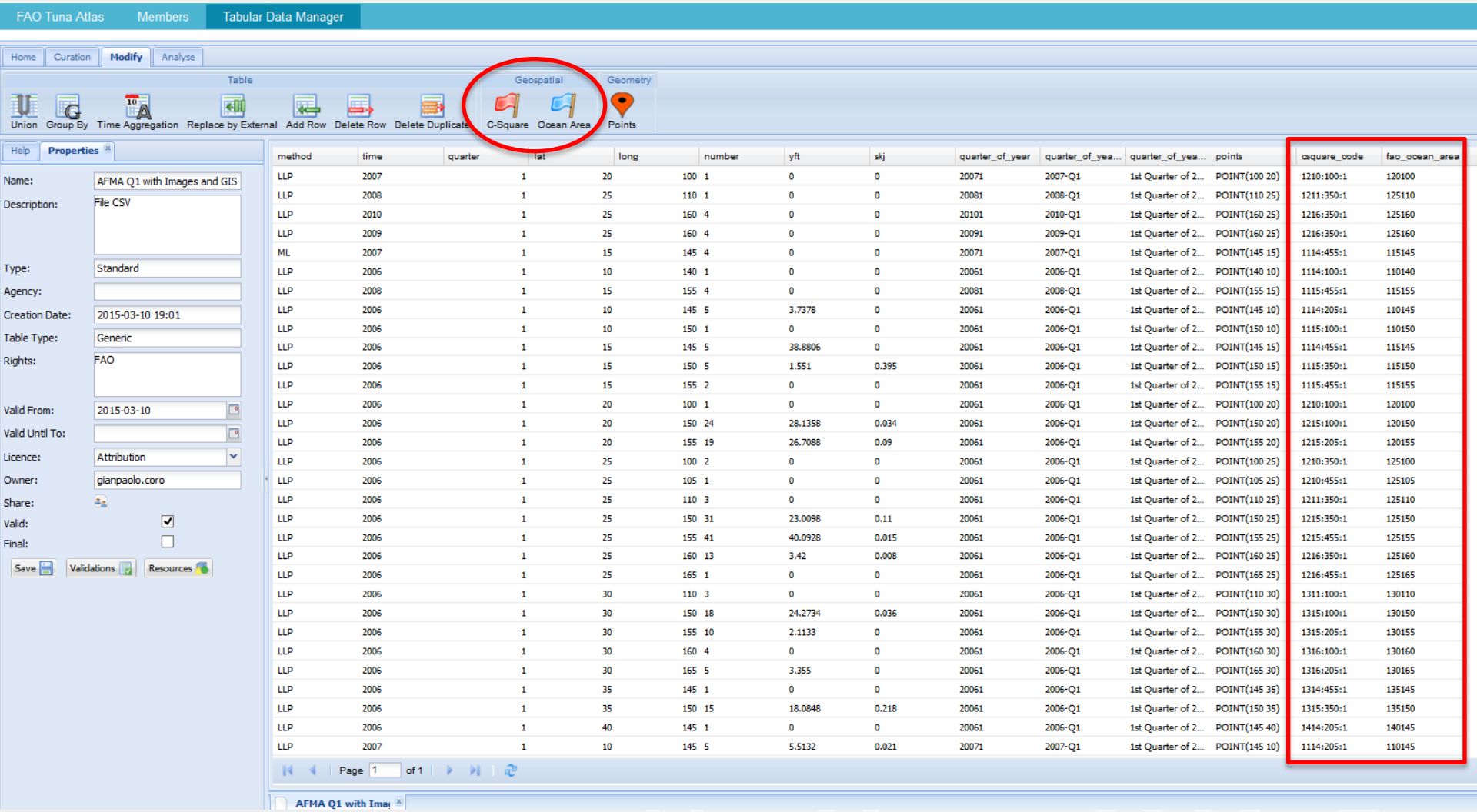
method	time	quarter	lat	long	number	yft	skj	quarter_of_year	quarter_of_yea...	quarter_of_yea...	points
LLP	2007	1	20	100	1	0	0	20071	2007-Q1	1st Quarter of 2...	POINT(100 20)
LLP	2008	1	25	110	1	0	0	20081	2008-Q1	1st Quarter of 2...	POINT(110 25)
LLP	2010	1	25	160	4	0	0	20101	2010-Q1	1st Quarter of 2...	POINT(160 25)
LLP	2009	1	25	160	4	0	0	20091	2009-Q1	1st Quarter of 2...	POINT(160 25)
ML	2007	1	15	145	4	0	0	20071	2007-Q1	1st Quarter of 2...	POINT(145 15)
LLP	2006	1	10	140	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(140 10)
LLP	2008	1	15	155	4	0	0	20081	2008-Q1	1st Quarter of 2...	POINT(155 15)
LLP	2006	1	10	145	5	3.7378	0	20061	2006-Q1	1st Quarter of 2...	POINT(145 10)
LLP	2006	1	10	150	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(150 10)
LLP	2006	1	15	145	5	38.8806	0	20061	2006-Q1	1st Quarter of 2...	POINT(145 15)
LLP	2006	1	15	150	5	1.551	0.395	20061	2006-Q1	1st Quarter of 2...	POINT(150 15)
LLP	2006	1	15	155	2	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(155 15)
LLP	2006	1	20	100	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(100 20)
LLP	2006	1	20	150	24	28.1358	0.034	20061	2006-Q1	1st Quarter of 2...	POINT(150 20)
LLP	2006	1	20	155	19	26.7088	0.09	20061	2006-Q1	1st Quarter of 2...	POINT(155 20)
LLP	2006	1	25	100	2	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(100 25)
LLP	2006	1	25	105	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(105 25)
LLP	2006	1	25	110	3	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(110 25)
LLP	2006	1	25	150	31	23.0098	0.11	20061	2006-Q1	1st Quarter of 2...	POINT(150 25)
LLP	2006	1	25	155	41	40.0928	0.015	20061	2006-Q1	1st Quarter of 2...	POINT(155 25)
LLP	2006	1	25	160	13	3.42	0.008	20061	2006-Q1	1st Quarter of 2...	POINT(160 25)
LLP	2006	1	25	165	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(165 25)
LLP	2006	1	30	110	3	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(110 30)
LLP	2006	1	30	150	18	24.2734	0.036	20061	2006-Q1	1st Quarter of 2...	POINT(150 30)
LLP	2006	1	30	155	10	2.1133	0	20061	2006-Q1	1st Quarter of 2...	POINT(155 30)
LLP	2006	1	30	160	4	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(160 30)
LLP	2006	1	30	165	5	3.355	0	20061	2006-Q1	1st Quarter of 2...	POINT(165 30)
LLP	2006	1	35	145	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(145 35)
LLP	2006	1	35	150	15	18.0848	0.218	20061	2006-Q1	1st Quarter of 2...	POINT(150 35)
LLP	2006	1	40	145	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(145 40)
LLP	2007	1	10	145	5	5.5132	0.021	20071	2007-Q1	1st Quarter of 2...	POINT(145 10)

Page 1 of 1

AFMA Q1 with Ima

TabMan

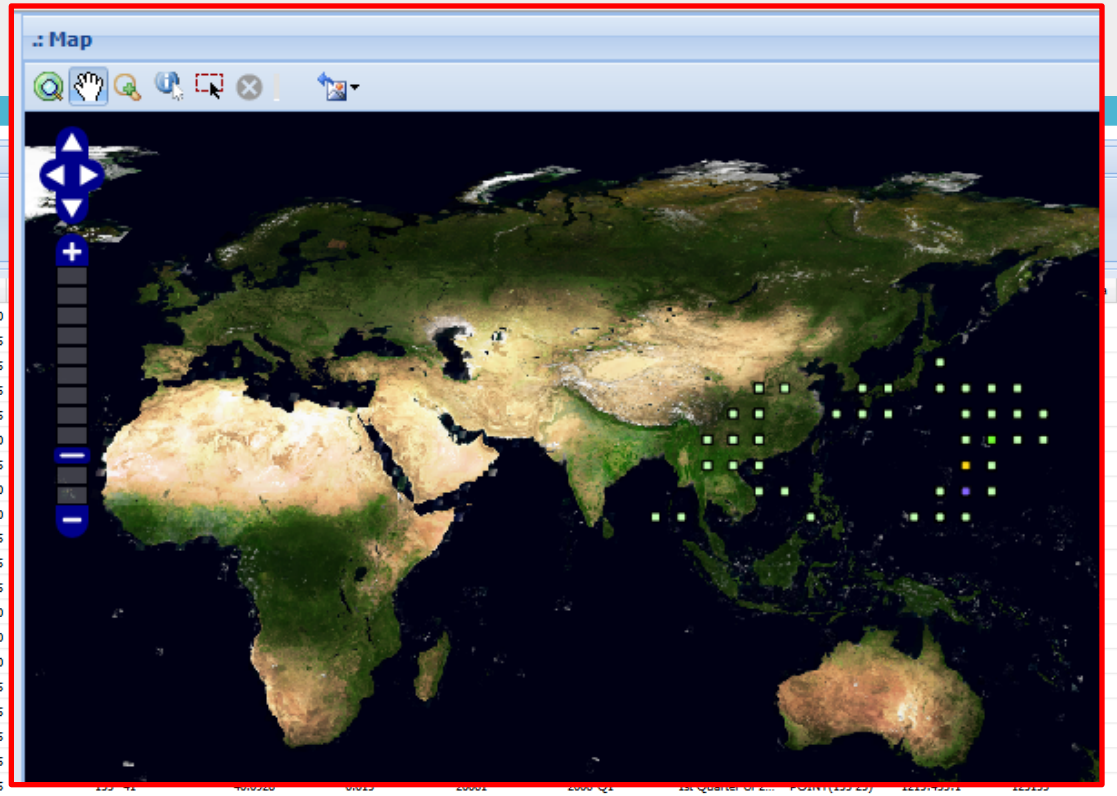
Csquare codes and FAO Ocean Areas codes



BlueBRIDGE

TabMan

GIS maps published under standard formats (WMS, WFS, WCS)



FAO Tuna Atlas | Members | Tabular Data Manager

Home | Curation | Modify | Analyse

Charts | Process | GIS

Create Chart | R Studio | Statistical | Create Map

Help | Properties

Name: AFMA Q1 with Images and GIS

Description: File CSV

Type: Standard

Agency:

Creation Date: 2015-03-10 19:01

Table Type: Generic

Rights: FAO

Valid From: 2015-03-10

Valid Until To:

Licence: Attribution

Owner: gianpaolo.coro

Share:

Valid: ☑

Final: ☐

Save | Validations | Resources

method	time	quarter	lat								
LLP	2007	1	20								
LLP	2008	1	25								
LLP	2010	1	25								
LLP	2009	1	25								
ML	2007	1	15								
LLP	2006	1	10								
LLP	2008	1	15								
LLP	2006	1	10								
LLP	2006	1	10								
LLP	2006	1	15								
LLP	2006	1	15								
LLP	2006	1	15								
LLP	2006	1	20								
LLP	2006	1	20								
LLP	2006	1	20								
LLP	2006	1	25								
LLP	2006	1	25								
LLP	2006	1	25								
LLP	2006	1	25								
LLP	2006	1	25	160	13	3.42	0.008	20061	2006-Q1	1st Quarter of 2...	POINT(160 25)	1216:350:1	125160
LLP	2006	1	25	165	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(165 25)	1216:455:1	125165
LLP	2006	1	30	110	3	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(110 30)	1311:100:1	130110
LLP	2006	1	30	150	18	24.2734	0.036	20061	2006-Q1	1st Quarter of 2...	POINT(150 30)	1315:100:1	130150
LLP	2006	1	30	155	10	2.1133	0	20061	2006-Q1	1st Quarter of 2...	POINT(155 30)	1315:205:1	130155
LLP	2006	1	30	160	4	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(160 30)	1316:100:1	130160
LLP	2006	1	30	165	5	3.355	0	20061	2006-Q1	1st Quarter of 2...	POINT(165 30)	1316:205:1	130165
LLP	2006	1	35	145	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(145 35)	1314:455:1	135145
LLP	2006	1	35	150	15	18.0848	0.218	20061	2006-Q1	1st Quarter of 2...	POINT(150 35)	1315:350:1	135150
LLP	2006	1	40	145	1	0	0	20061	2006-Q1	1st Quarter of 2...	POINT(145 40)	1414:205:1	140145
LLP	2007	1	10	145	5	5.5132	0.021	20071	2007-Q1	1st Quarter of 2...	POINT(145 10)	1114:205:1	110145

Page 1 of 1

AFMA Q1 with Ima

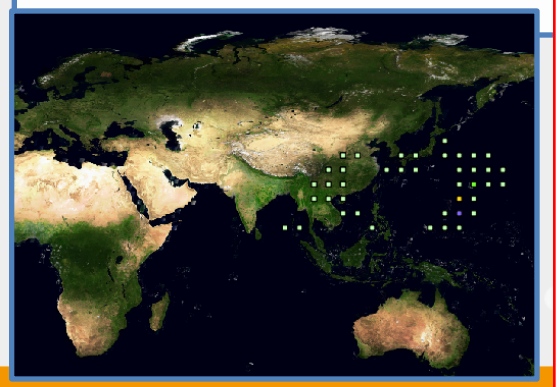
Tuna Atlas VRE

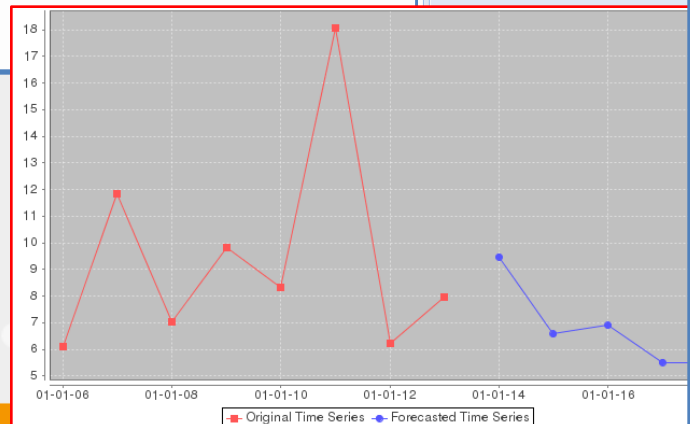
FAO Tuna Atlas

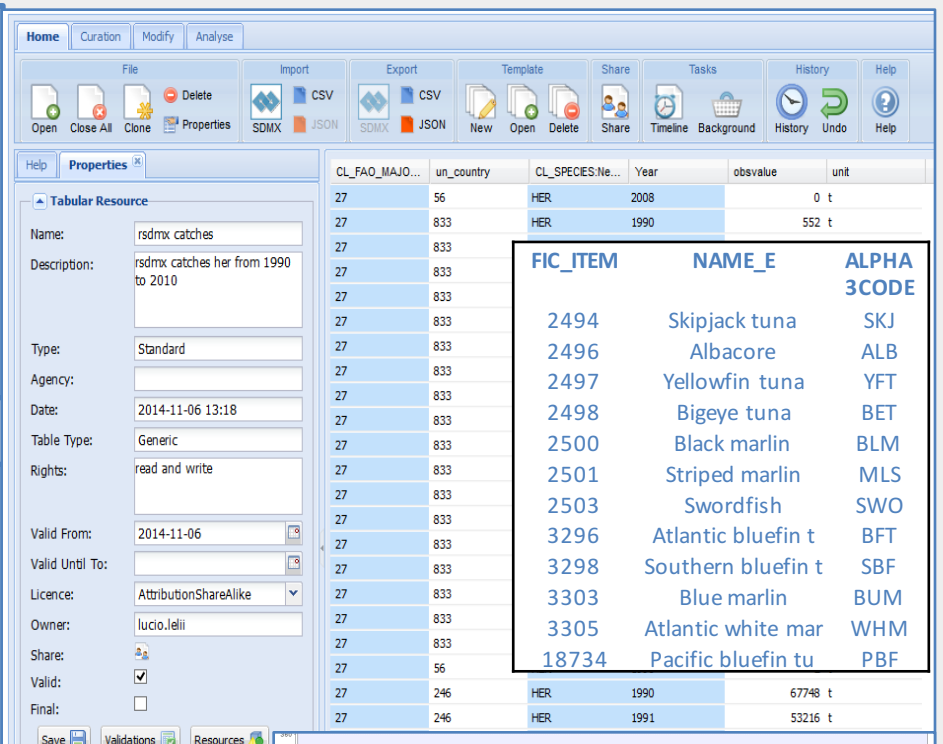
- Represents world wide Tuna catches in an interactive map
- Centralize and harmonize Tuna RFMOs Statistics with a common format
- Share the approach with other institutions with different exploitations

D4Science support to Tuna Atlas

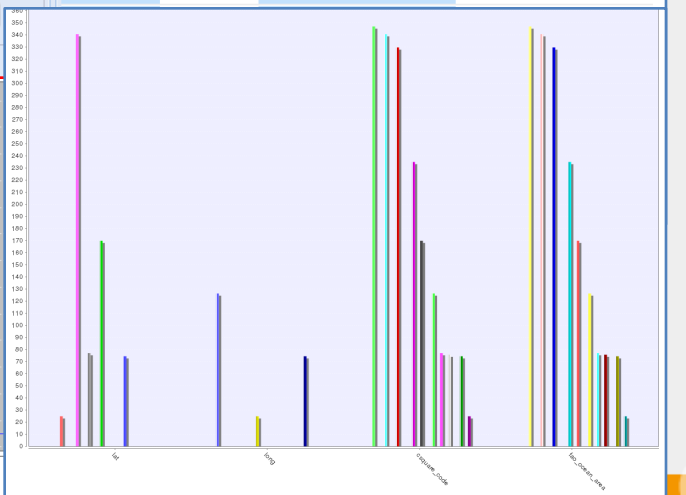
- Create a Virtual Research Environment
- Centralise and harmonise catch statistics coming from heterogeneous sources
- Modify and analyse tabular datasets
- Visualise and summarize information and trends



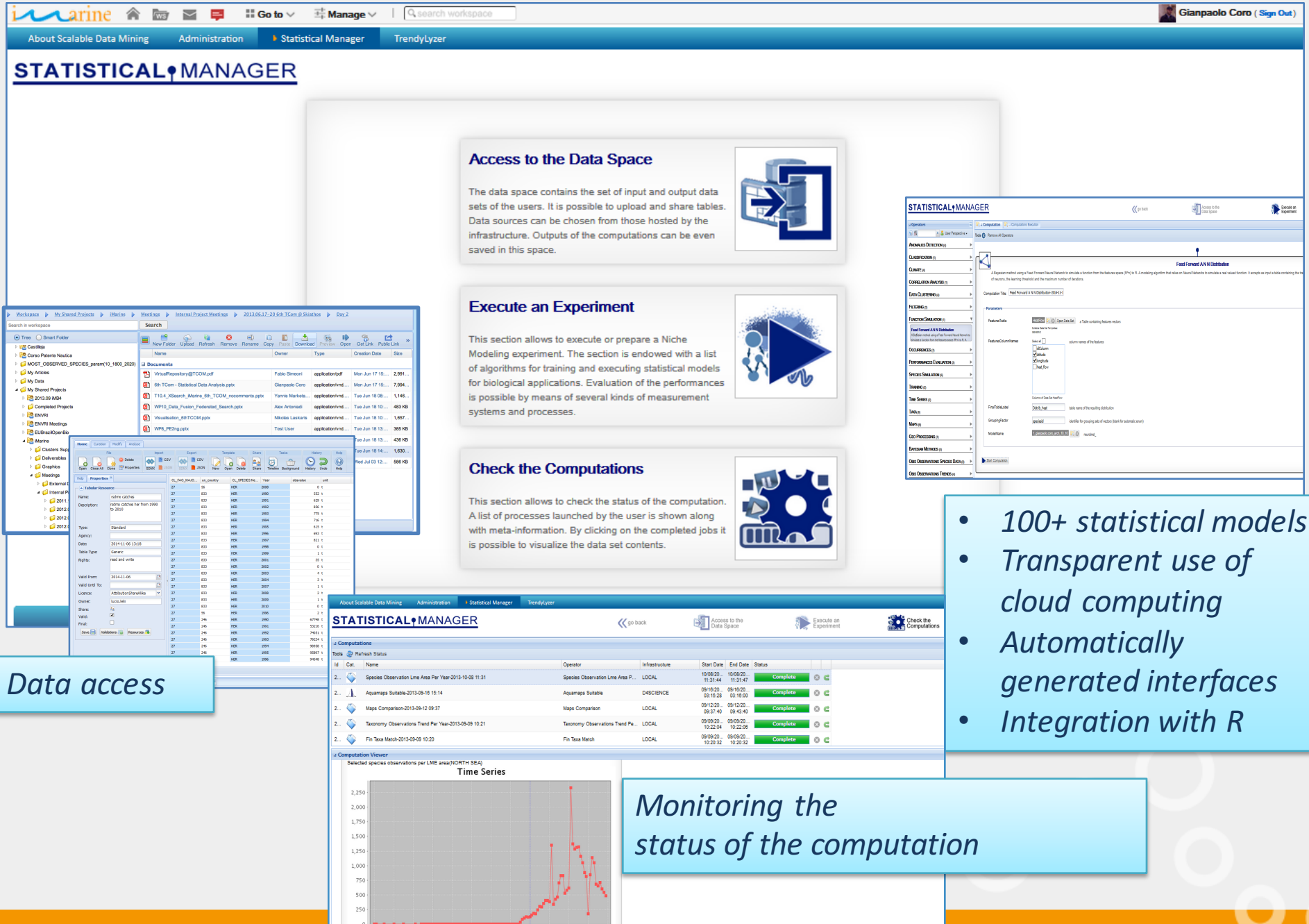




FIC_ITEM	NAME_E	ALPHA 3CODE
2494	Skipjack tuna	SKJ
2496	Albacore	ALB
2497	Yellowfin tuna	YFT
2498	Bigeye tuna	BET
2500	Black marlin	BLM
2501	Striped marlin	MLS
2503	Swordfish	SWO
3296	Atlantic bluefin t	BFT
3298	Southern bluefin t	SBF
3303	Blue marlin	BUM
3305	Atlantic white mar	WHM
18734	Pacific bluefin tu	PBF



Statistical Data Manager (StatMan)



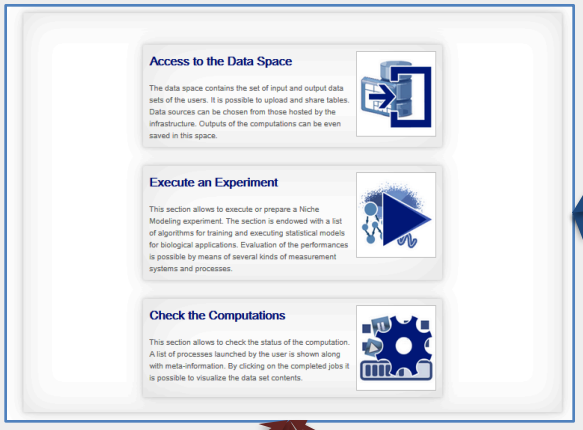
Data access

- *100+ statistical models*
- *Transparent use of cloud computing*
- *Automatically generated interfaces*
- *Integration with R*

Monitoring the status of the computation

Statistical Manager (StatMan)

Statistical services



Java

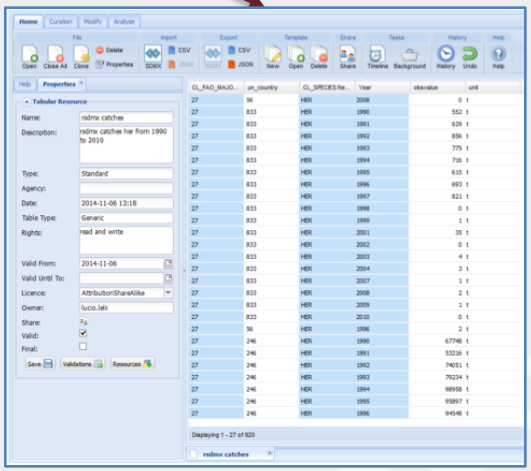
OGC

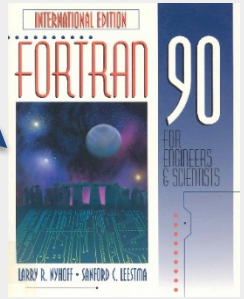
OGC WPS Interface

External Computing Facility

R

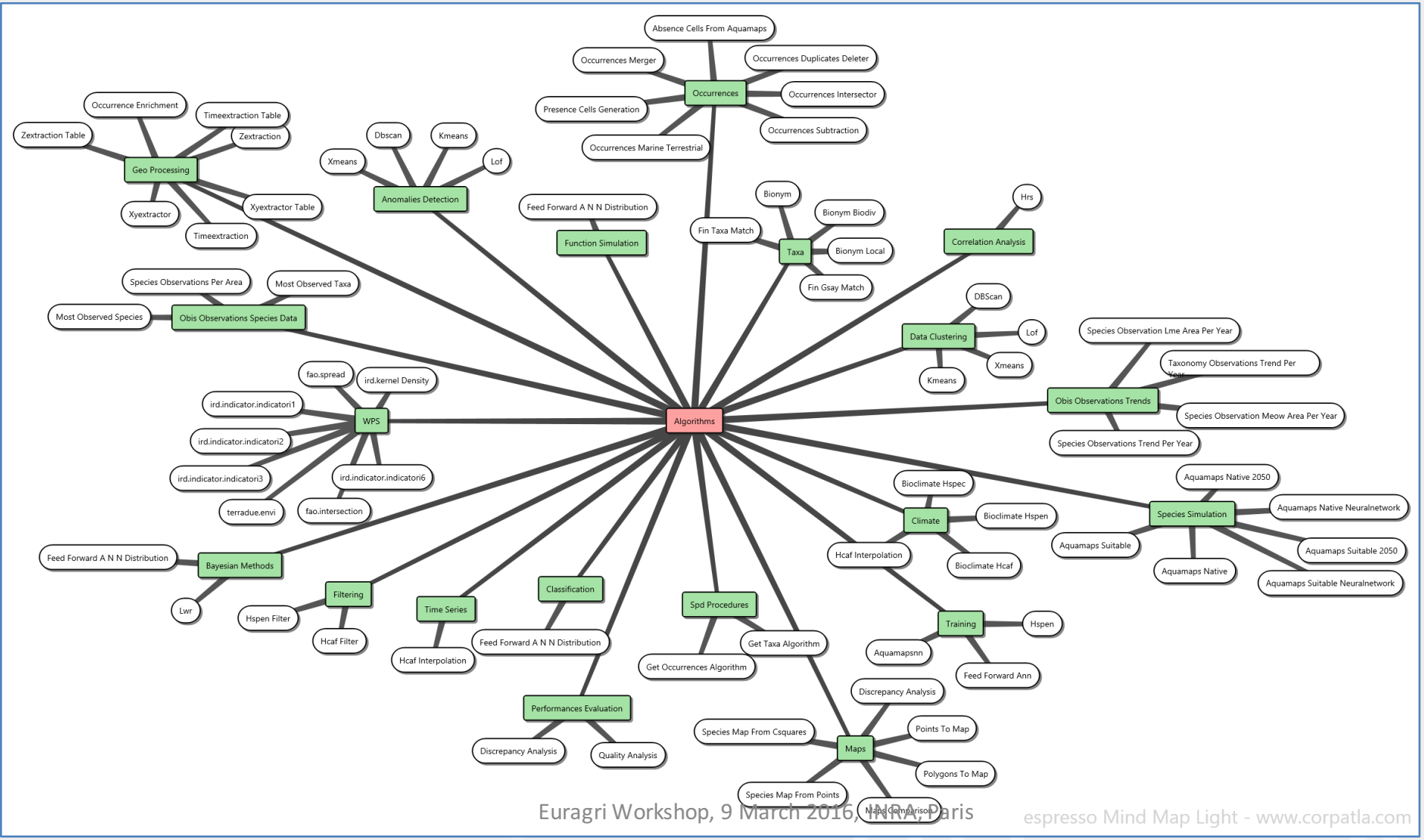
WPS





Data standardisation services

100+ Hosted algorithms

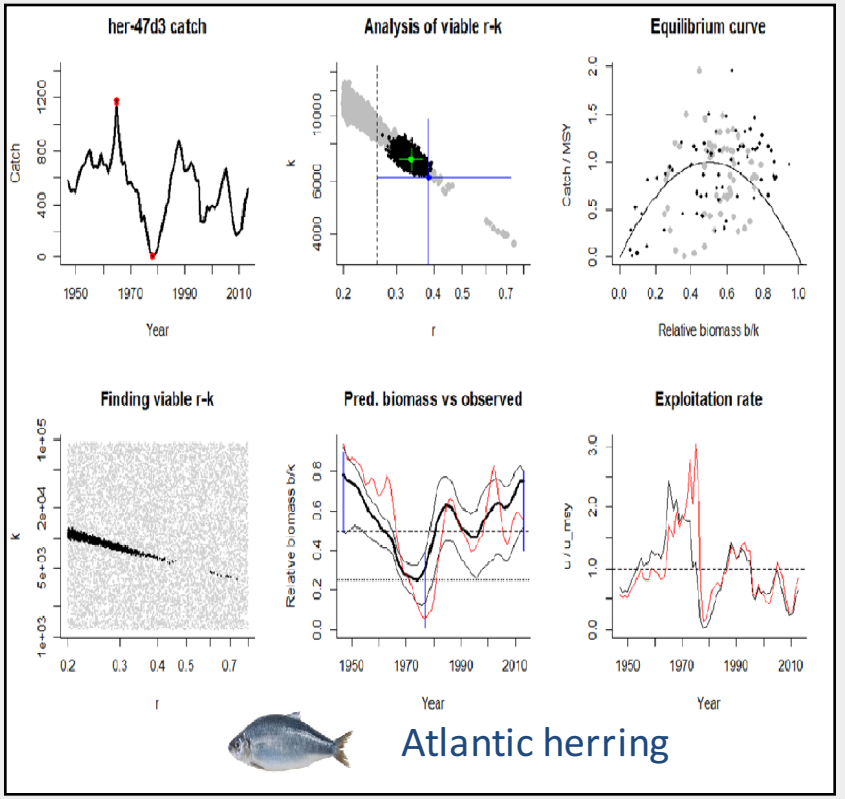


Algorithms categories

- Anomalies Detection (4)
- Classification (1)
- Climate (4)
- Correlation Analysis (1)
- Data Clustering (4)
- Filtering (2)
- Function Simulation (1)
- Occurrences (8)
- Performances Evaluation (2)
- Species Simulation (6)
- Training (2)
- Time Series (2)
- Taxa (5)
- Maps (7)
- Geo Processing (14)
- Bayesian Methods (5)
- Obis Observations Species Data (3)
- Obis Observations Trends (4)
- SPECIES Procedures (2)
- Vessels (3)
- Databases (9)
- Spread (4)
- Charts (4)
- Precipitations (2)
- Stock Assessment (7)

Stock assessment

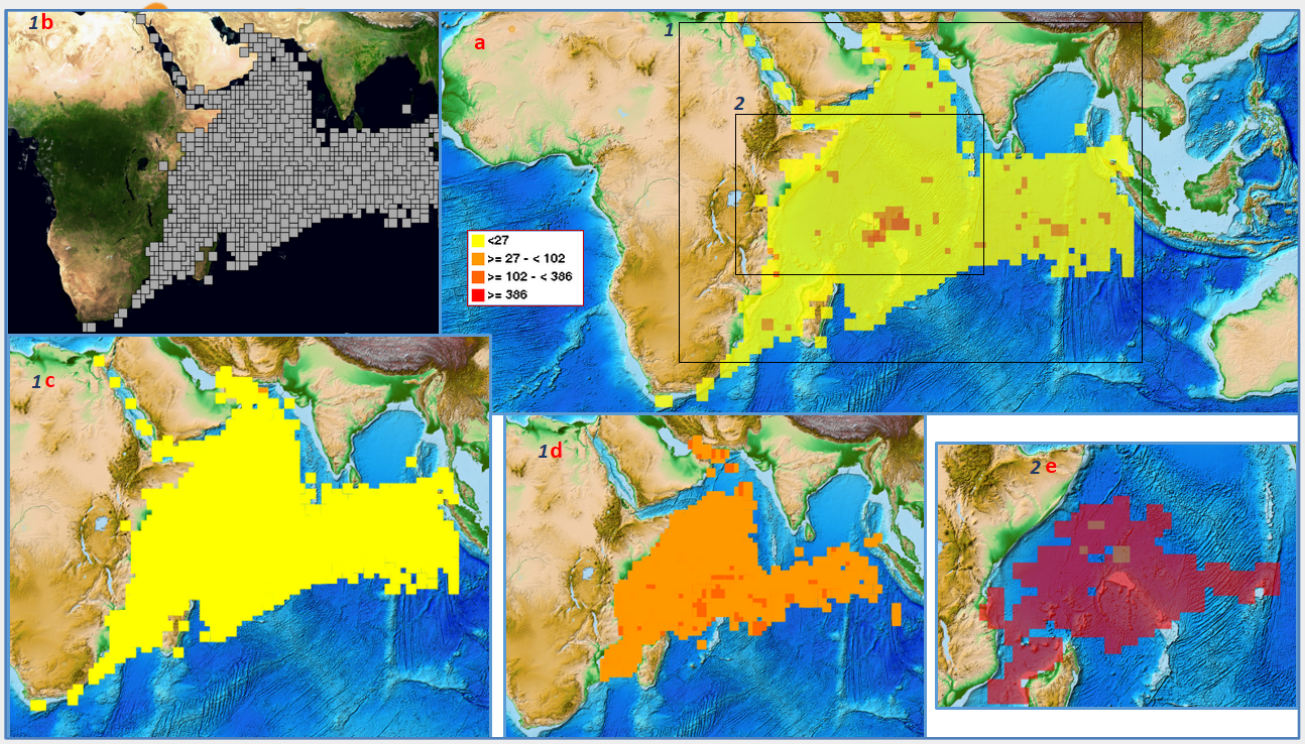
- **Length-Weight Relations**: estimates Length-Weight relation parameters for marine species, using Bayesian methods. Developed by R. Froese, T. Thorson and R. B. Reyes
- **SGVM interpolation**: interpolation of vessels trajectories. Developed by the Study Group on VMS, involving ICES
- **FAO MSY**: stock assessment for FAO catch data. Developed by the Resource Use and Conservation Division of the FAO Fisheries and Aquaculture Department (ref. Y. Ye)
- **ICCAT VPA**: stock assessment method for International Commission for the Conservation of Atlantic Tunas (ICCAT) data. Developed by Ifremer and IRD (ref. S. Bonhommeau, J. Bard)
- **CMSY**:estimates Maximum Sustainable Yield from catch statistics. Prime choice for ICES as main stock assessment tool. Developed by R. Froese, G. Coro, N. Demirel, K. Kleisner and H. Winker

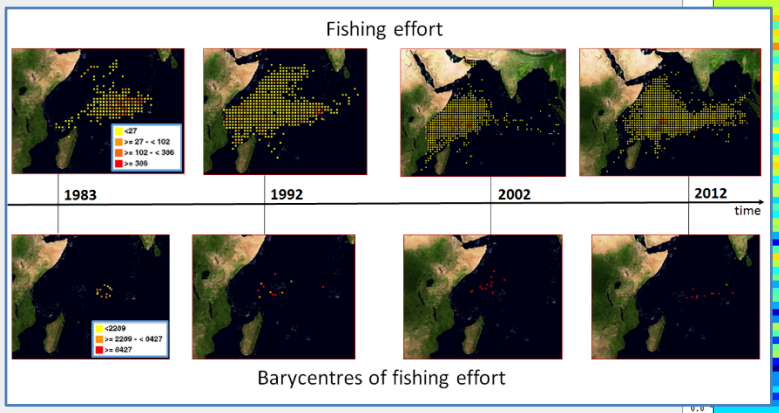


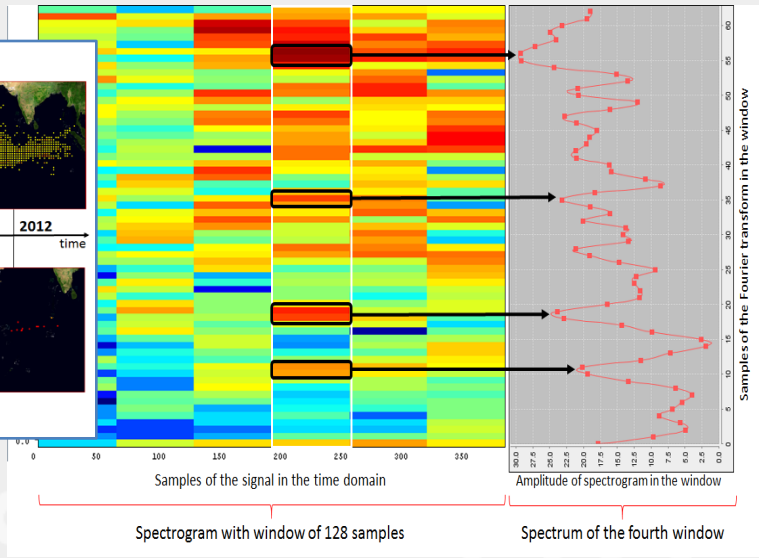
Atlantic herring

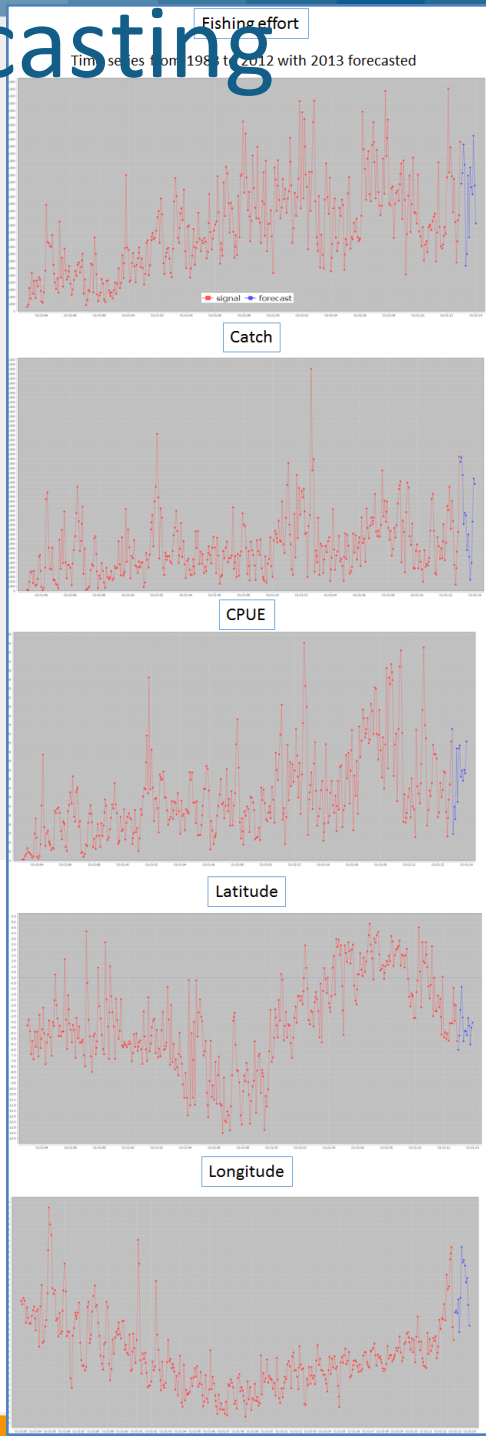
D4Science reduced time-to-market: State-of-the-art models to estimate Maximum Sustainable Yield computational time reduced of 95% in average

Catch statistics forecasting

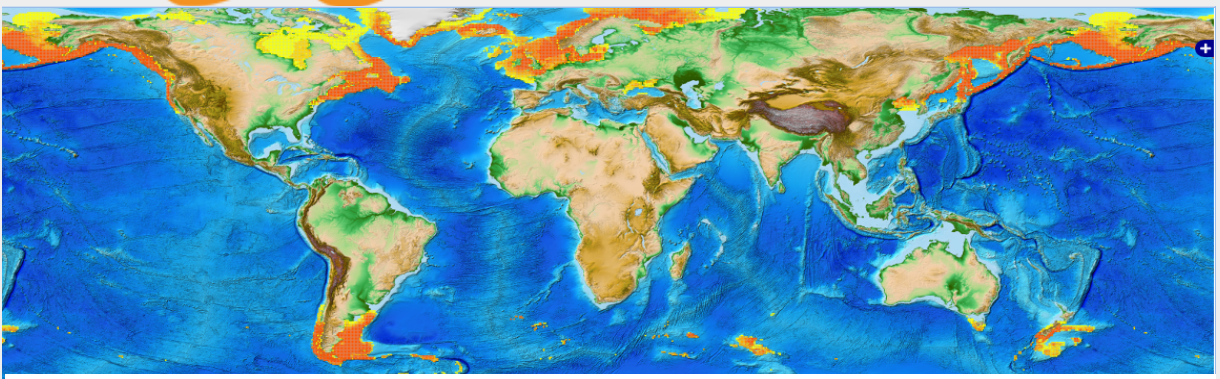




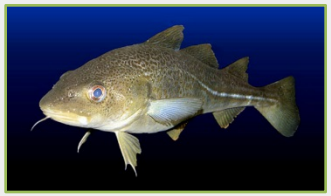




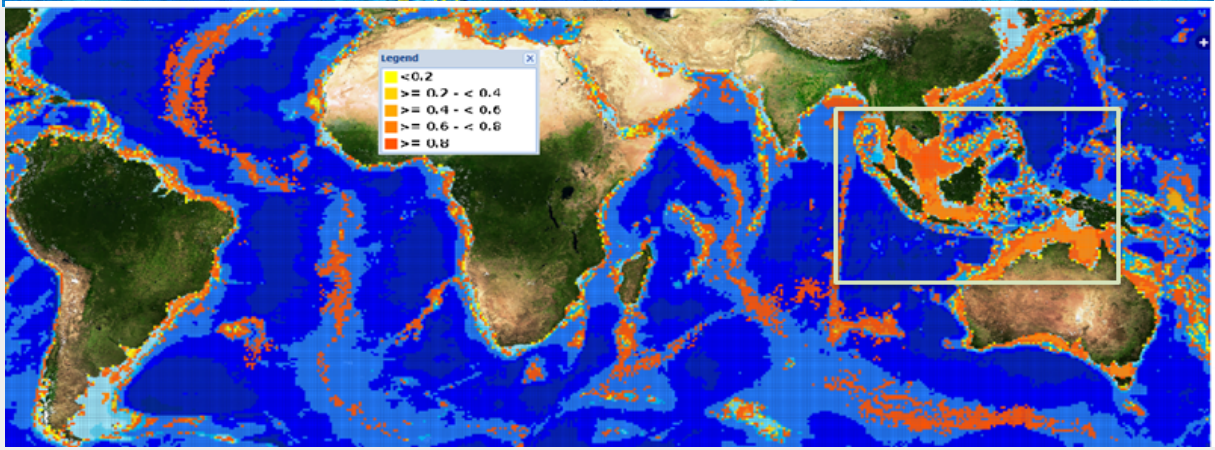
Ecology



AquaMaps

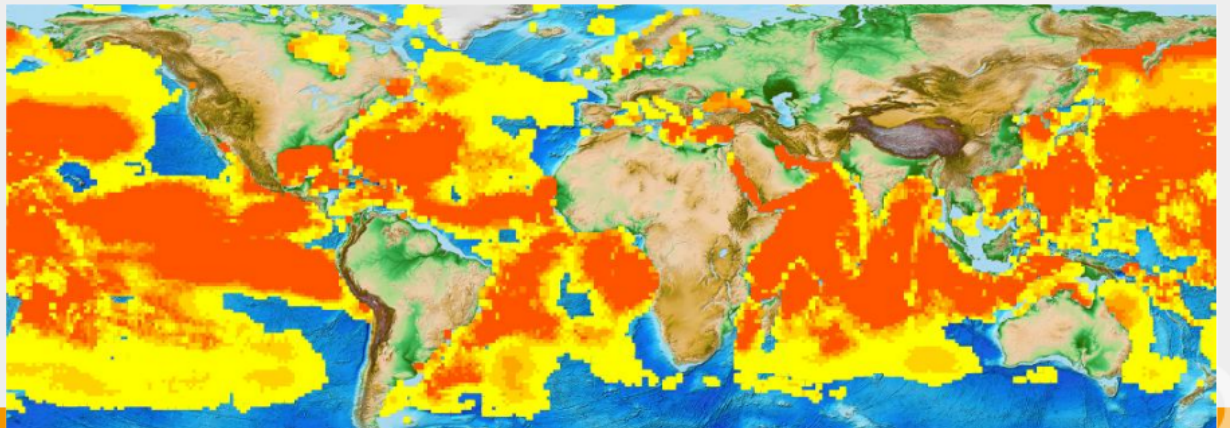


Atlantic cod



Neural Networks

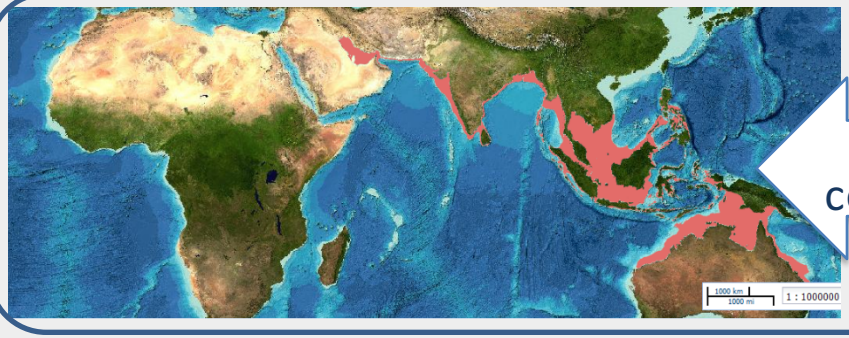
Coelacanth



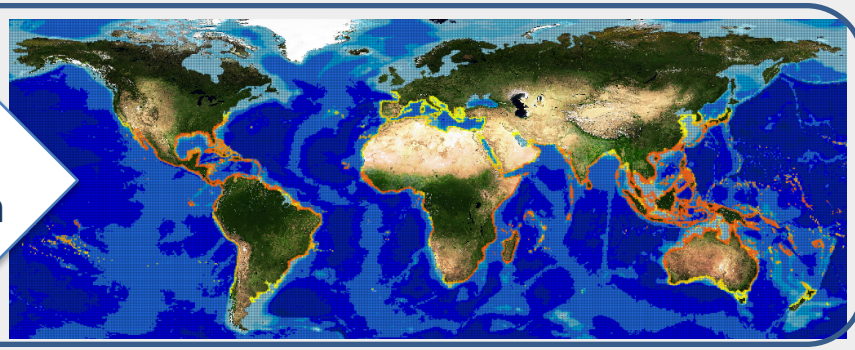
Neural Networks and MaxEnt

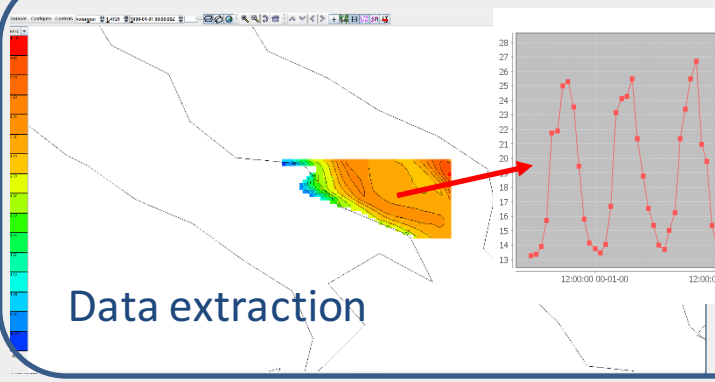
Giant squid

Geospatial data processing

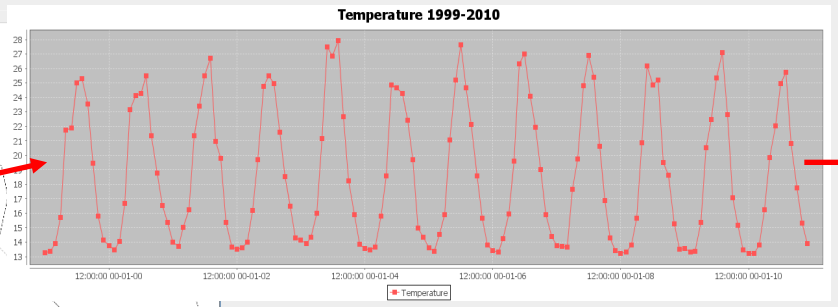


Maps comparison

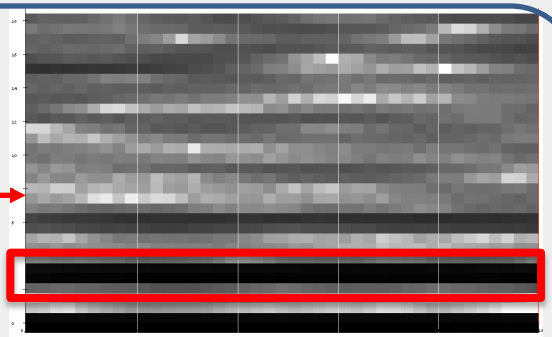




Data extraction



Signal processing

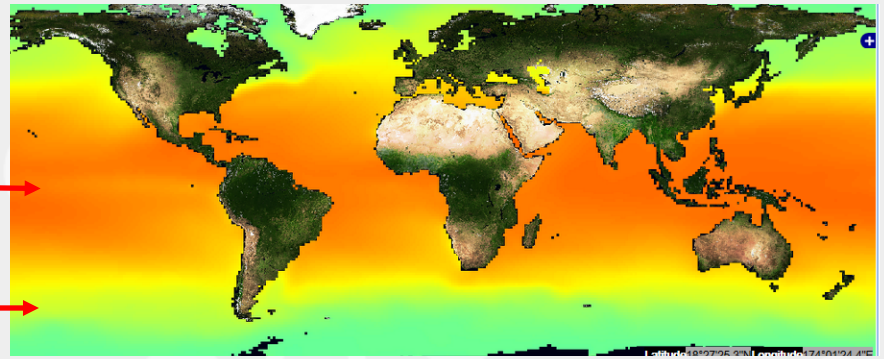


Periodicity detection

CWP_AO_MAJO...	un_country	CL_SPECIES:Ne...	Year	obsvalue	unit
27	56	HER	2008	0	t
27	833	HER	1990	552	t
27	833	HER	1991	629	t
27	833	HER	1992	856	t
27	833	HER	1993	775	t
27	833	HER	1994	716	t
27	833	HER	1995	615	t
27	833	HER	1996	693	t
27	833	HER	1997	821	t
27	833	HER	1998	0	t
27	833	HER	1999	1	t
27	833	HER	2001	35	t
27	833	HER	2002	0	t
27	833	HER	2003	4	t
27	833	HER	2004	3	t
27	833	HER	2007	1	t
27	833	HER	2008	2	t
27	833	HER	2009	1	t
27	833	HER	2010	0	t
27	56	HER	1996	2	t
27	246	HER	1990	67748	t
27	246	HER	1991	53216	t
27	246	HER	1992	74051	t
27	246	HER	1993	79234	t

Maps generation

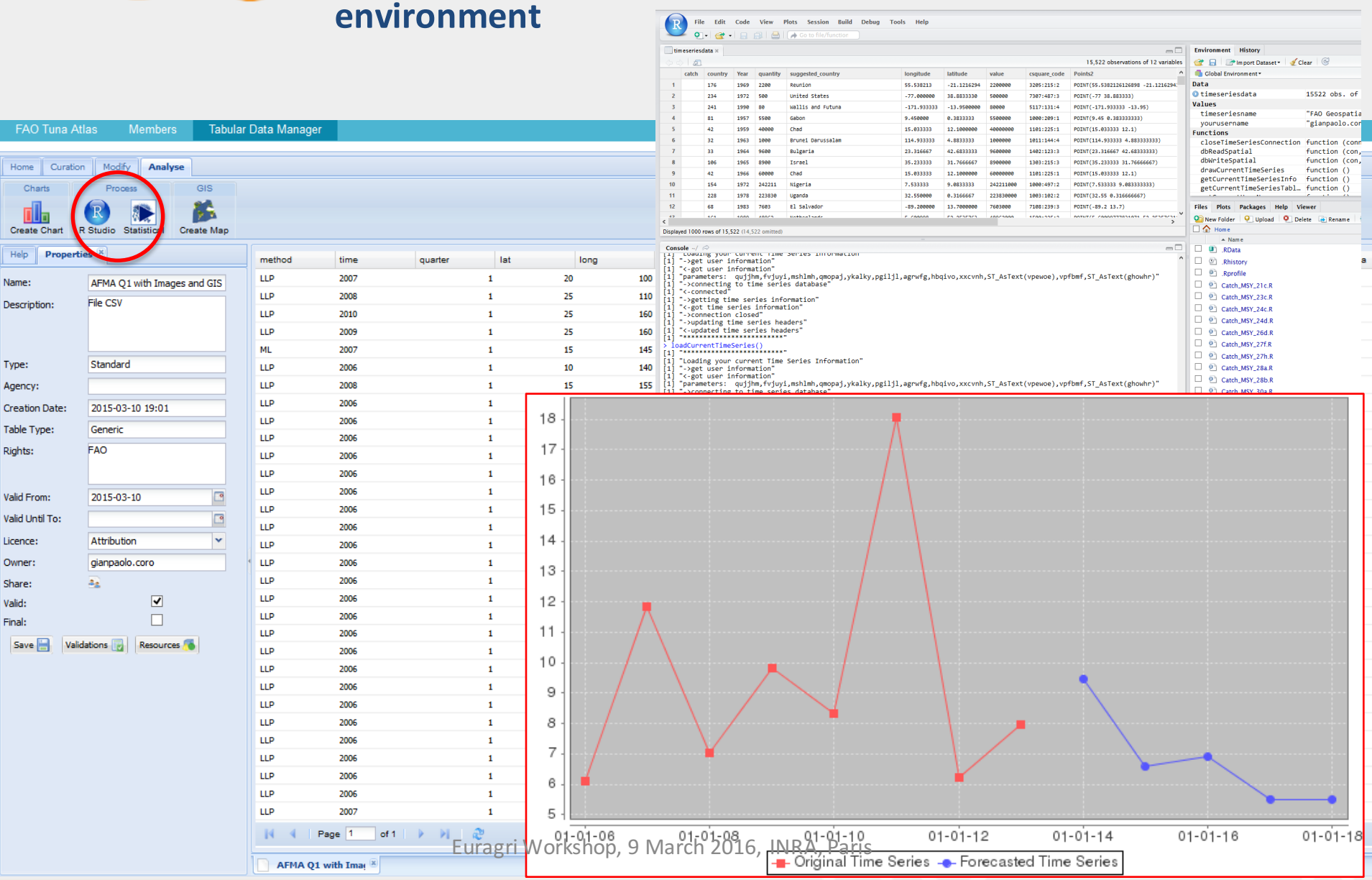
NetCDF file



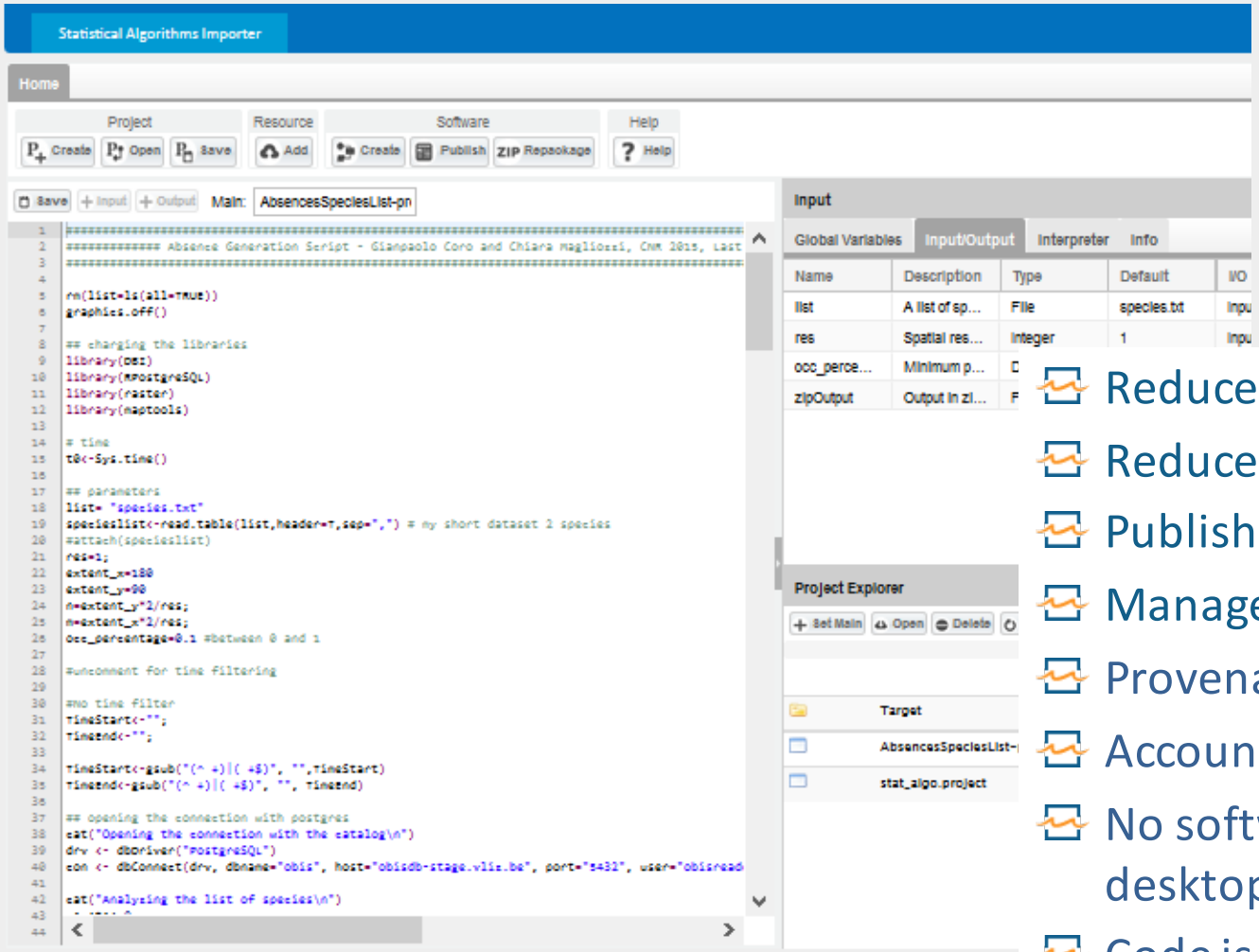
BlueBRIDGE

TabMan

Forecasting and analysis using an online R development environment



Algorithm Importer



- Reduce computation time
- Reduce integration time
- Publish algorithm as-a-Service
- Manage access policies
- Provenance Management
- Accounting
- No software required on desktop machines
- Code is not disclosed

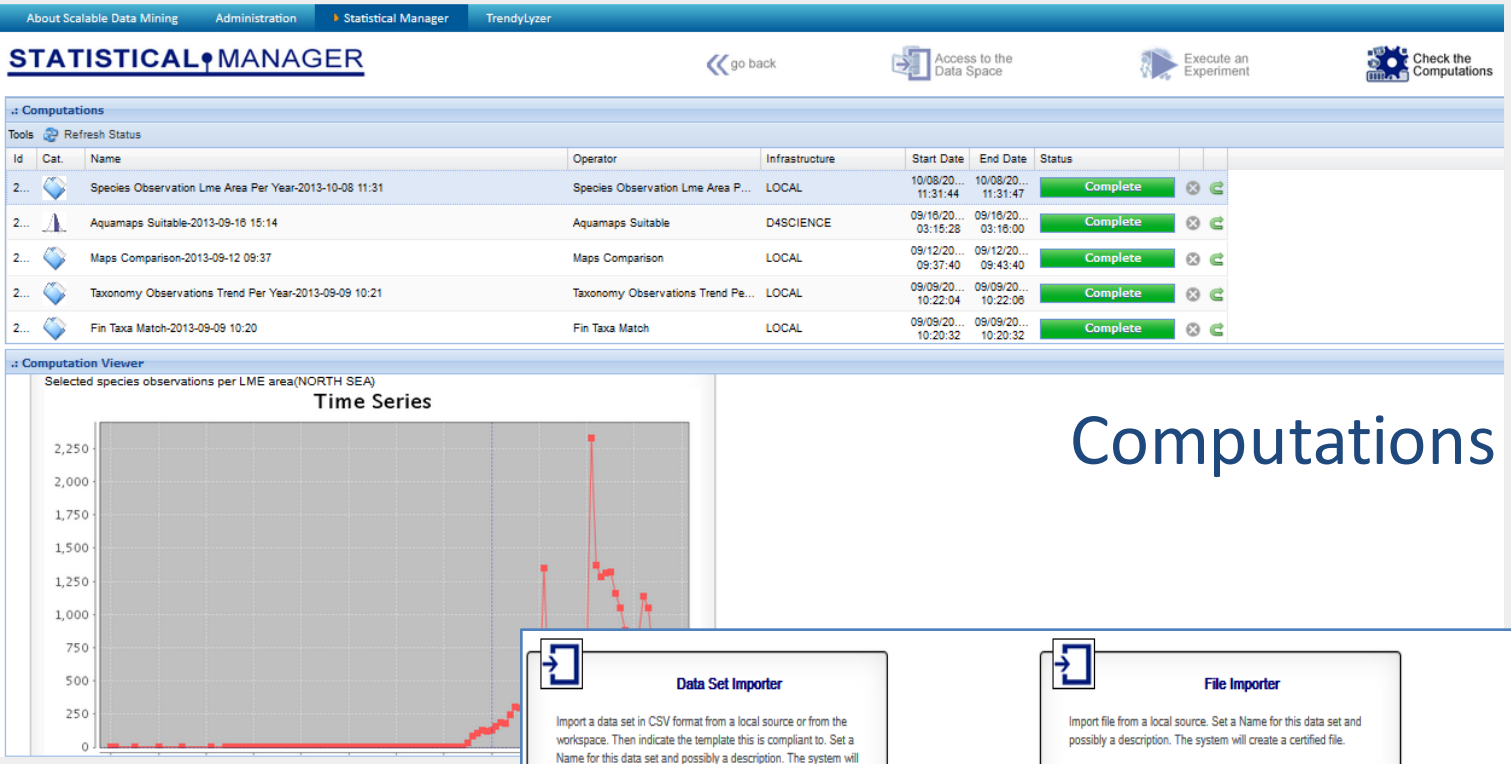
Interfaces

1 - iMarine portal

2- QGIS (via WPS Client)

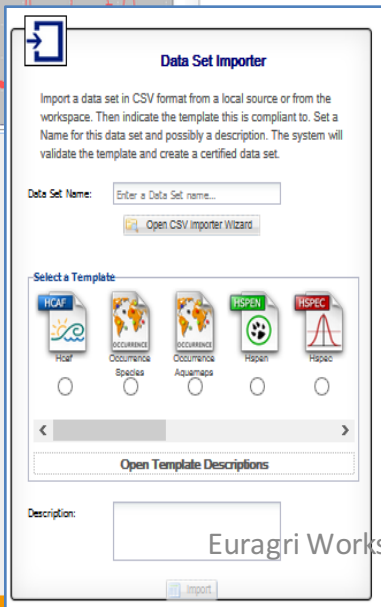
3 -WPS

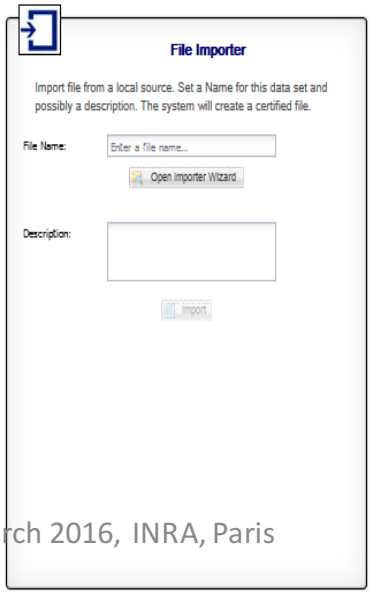
Computations and data

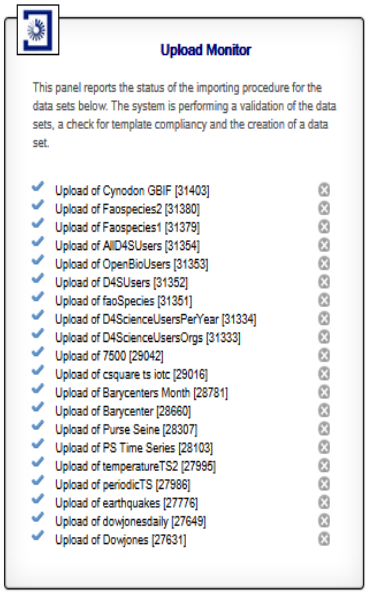


Computations history and summary

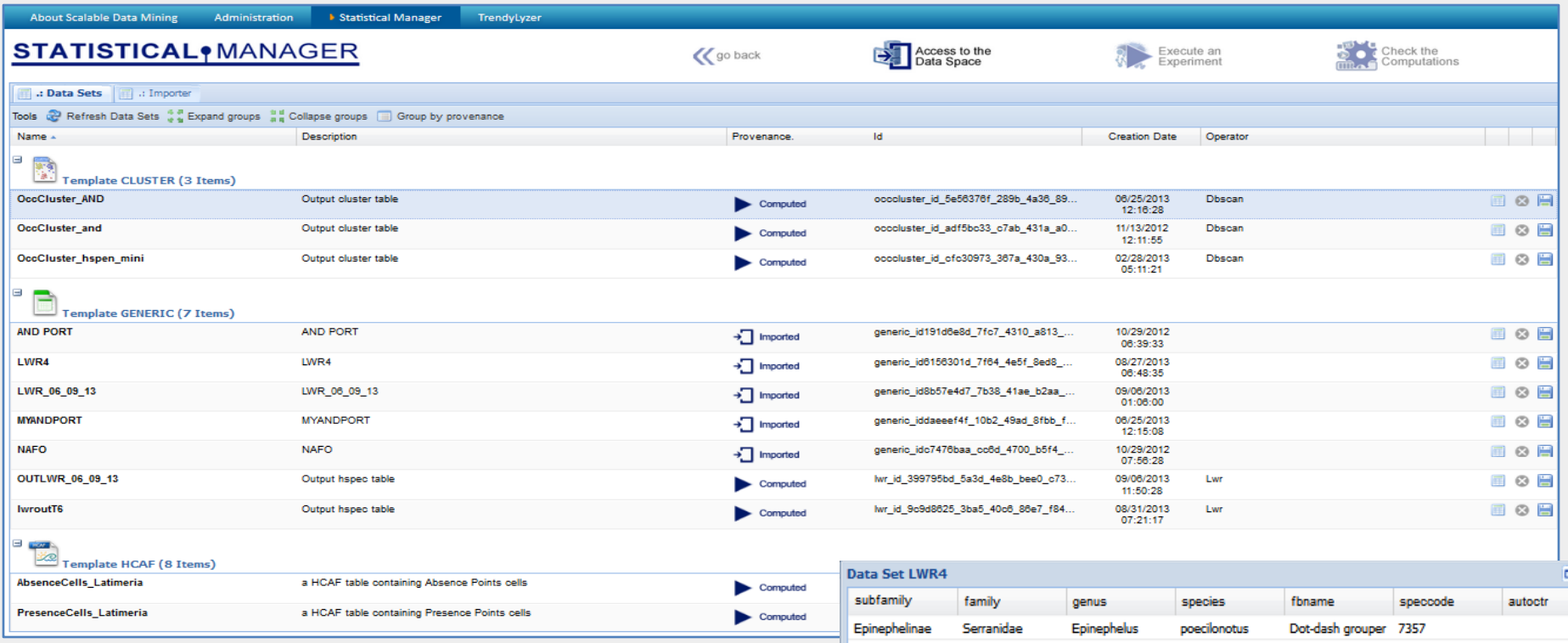
Data importing facilities

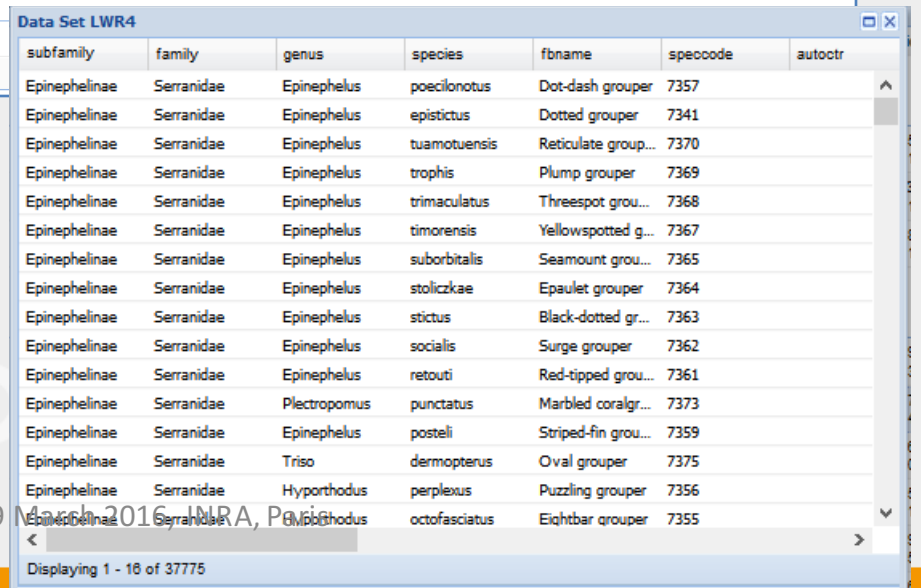






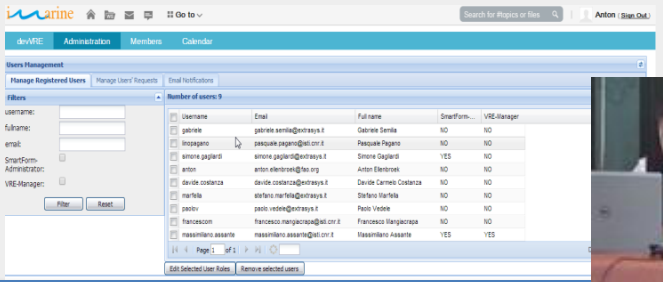
Data Management

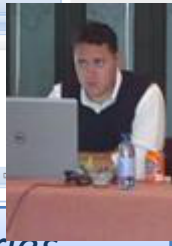




Data exploring and tables visualisation facilities

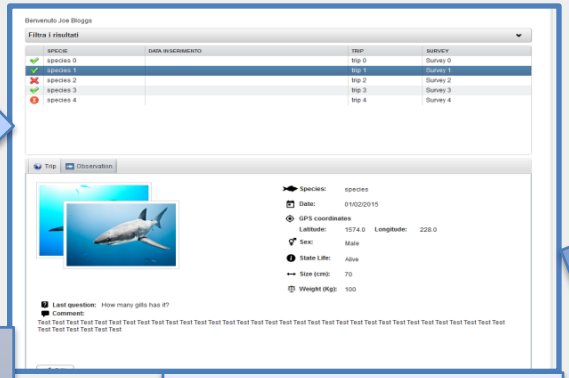
Data collection: SmartForm VRE



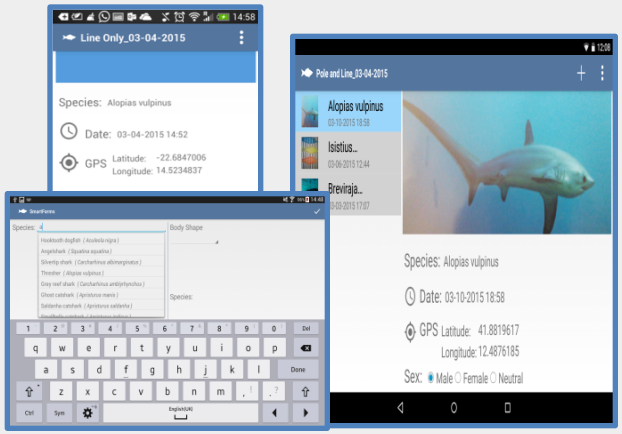


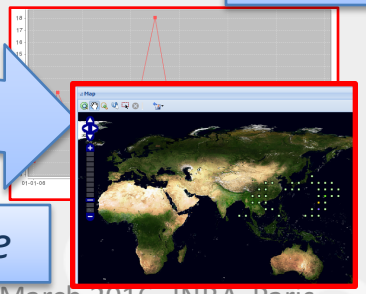
Define forms, controlled vocabularies, ….

Select a Fishery survey



Validate& Enrich

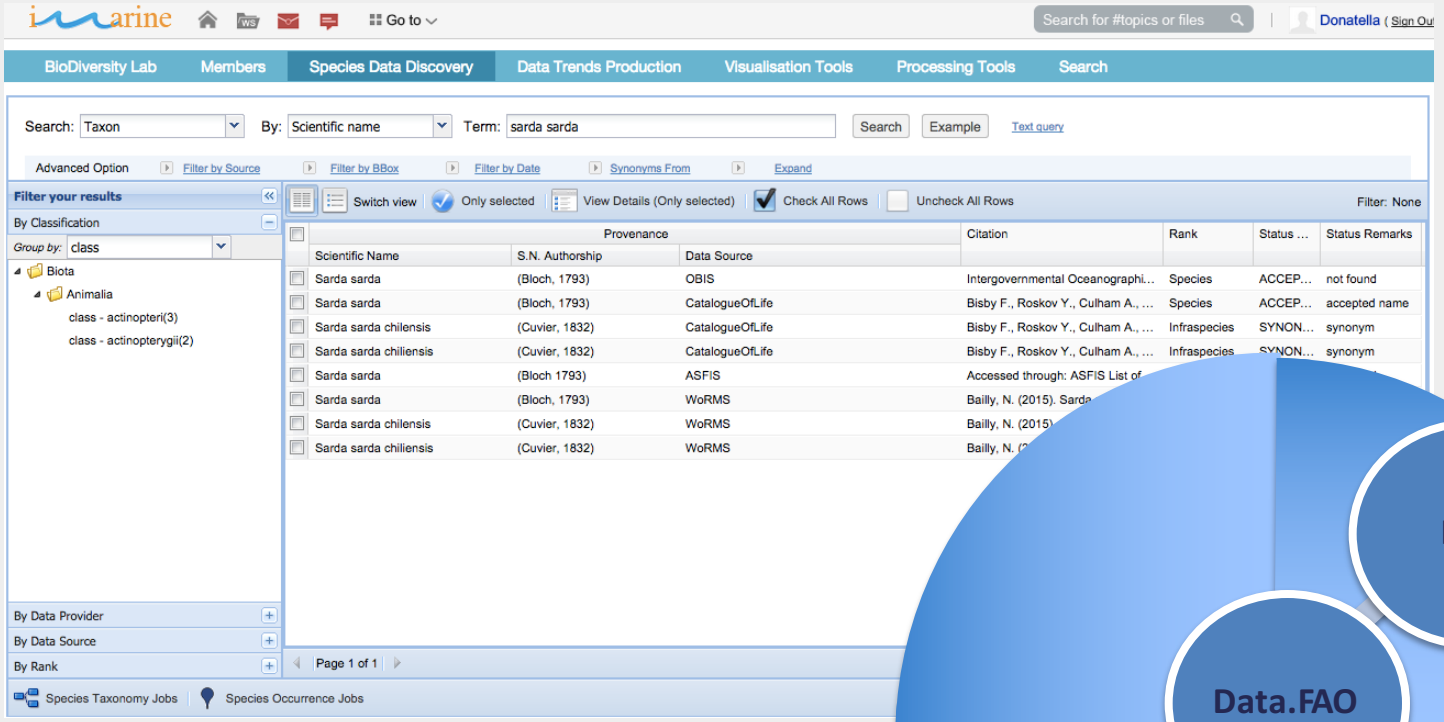


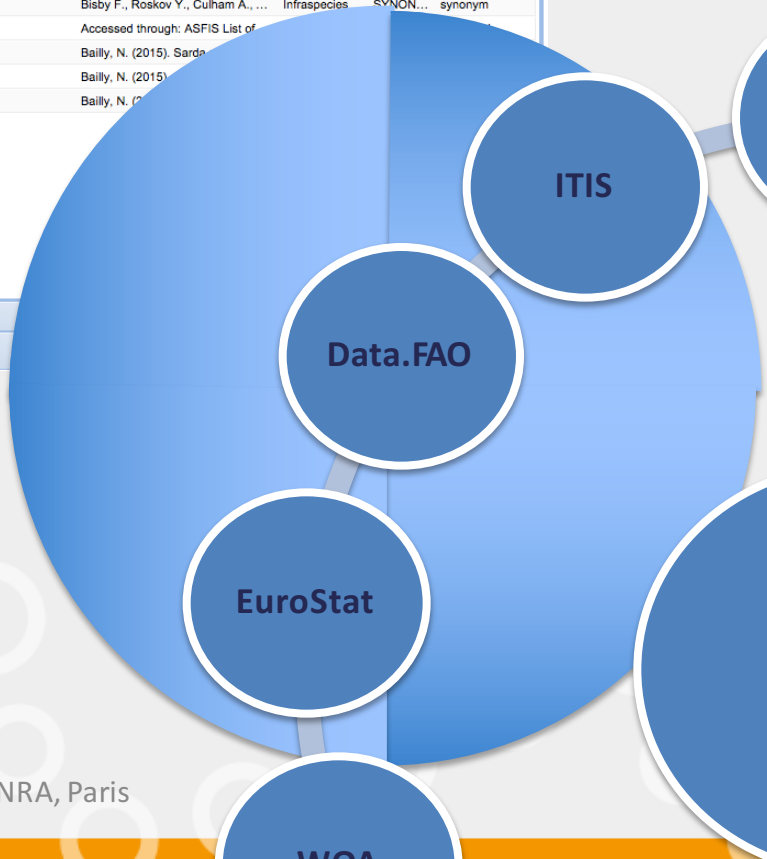


Analyse

Deposit

BioDiversityLab: uniform access to data



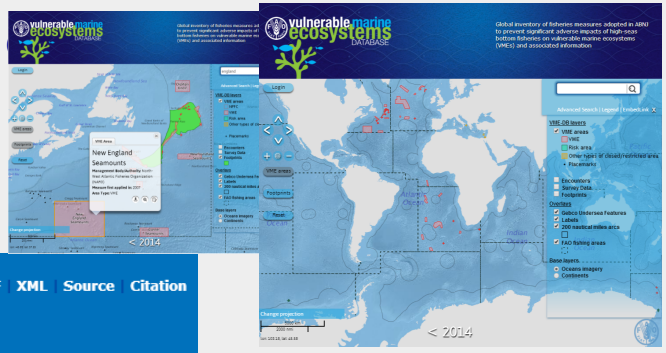


BlueBRIDGE

VME-DB

Report production: VME-DB

The **International Guidelines for the Management of Deep-Sea Fisheries on the High Seas**. **VME database** to assist in informed decision making and the development **sustainability** and reduce impacts.



Time

Corner Seamounts Selected year: 2012

Print | PDF | XML | Source | Citation

Description | Map | Management | Regional overview | Meeting reports | Media

Measures specific to this VME area

Area closed to bottom fishing until 31 Dec 2014, with exploratory fishing option. Vessels fishing this ares shall have a scientific observer onboard. (Art 16.1-3)

Period in force: 2007 – 2014

Source of information
NAFO Conservation and Enforcement Measures 2012 (NAFO FC Doc. 12/1 Serial No. N6001) Fisheries Commission 2012

General measures

Fishing areas
The comprehensive map of existing bottom fishing areas (as delineated by the coordinates shown in Table 1 and illustrated in Figure 4) shall be revised regularly to incorporate any new relevant information. Contracting Parties may, in the future, consider the possibility of refining the comprehensive map on the basis of haul by haul information, if available. (Art.17)

Exploratory fishing protocol
Exploratory fishing covers all bottom fishing activities (a) outside of the existing bottom fishing area and (b) to fisheries within the existing bottom fishing area that show significant change. (Art 15.8). Exploratory fisheries must be conducted according to an exploratory fisheries protocol (Art 18; Annex 1E.I-IV) and are subject to review FC and SC. Exploritory fisheries will be allowed only if there are adequate mitigation measures to prevent SAI to VMEs (Art 19).

Encounter protocols
In existing bottom fishing areas, encounters with VME indicator species above a threshold value are reported to the Executive Secretary and trigger a 2 nmile move on rule. In new fishing areas, such encounters also result in temporary closures of 2 nmile radius and require a more detailed report

Management Body/Authority
Northwest Atlantic Fisheries Organization (NAFO)
The Northwest Atlantic Fisheries Organization's (NAFO) overall objective is to contribute through consultation and cooperation to the optimum utilization, rational management and conservation of the fishery resources of its area of competence, and to ensure the long term conservation and sustainable use of the fishery resources and, in so doing, to safeguard the marine ecosystems in which these resources are found.

Web site
http://www.nafo.int/

Regional Fishery Body fact sheet
http://www.fao.org/fishery/rfb/NAFO/en

VME record

Description (Habitat & Biology)

Specific measures

Historical information on fishing areas and closed areas

Meetings & other Sources of Information

General Measures

RFMO

iMarine experience

BlueBRIDGE

(Sept. 2015 – Feb. 2018)

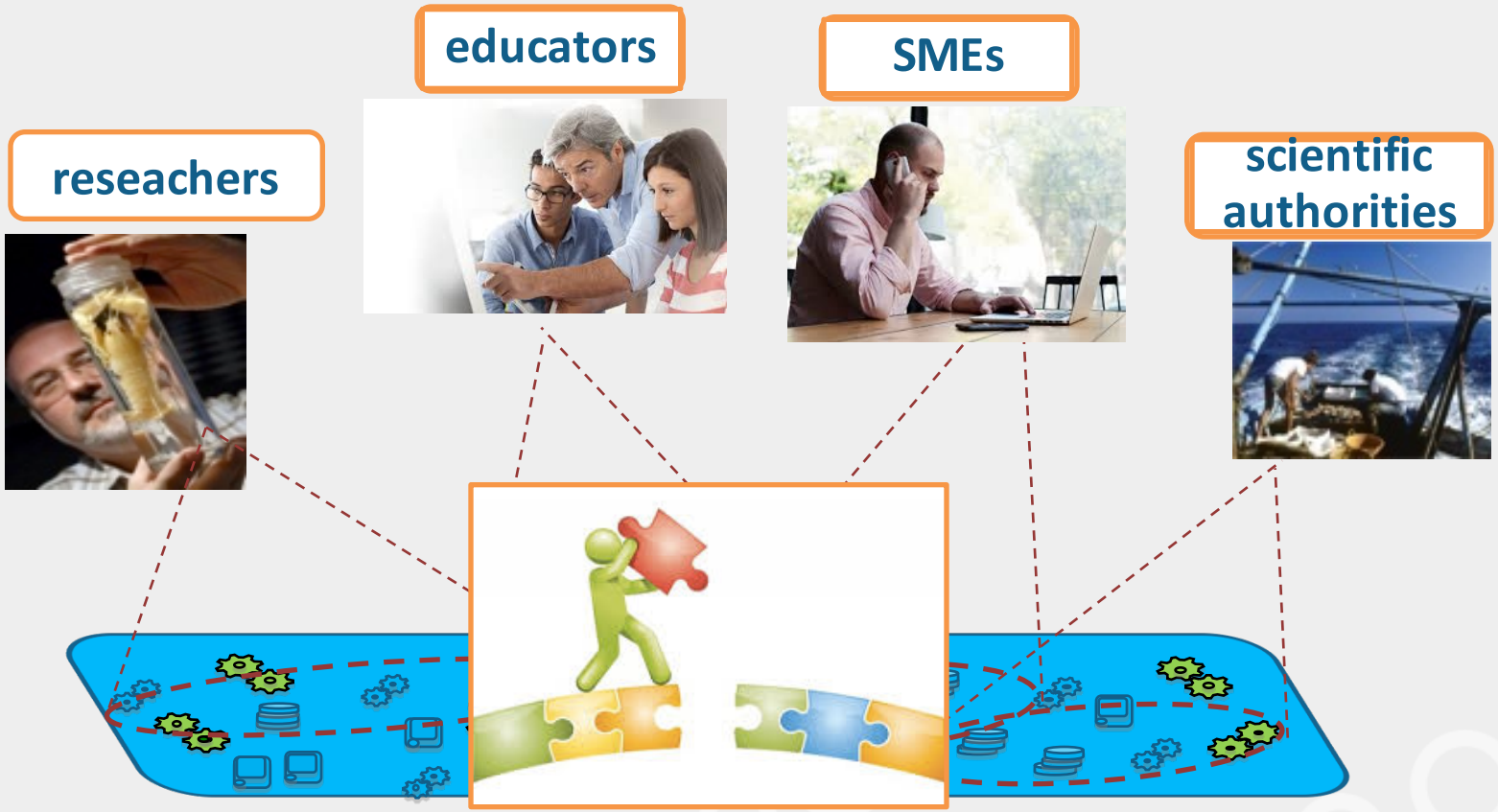
Building **R**esearch environments fostering **I**nnovation, **D**ecision making, **G**overnance and **E**ducation

for Blue growth

With the support of European Commission

The EU Framework Programme for Research & Innovation 2014-2020 HORIZON 2020

Target stakeholders



Expected impact

Researchers

Support capacity building & Innovate current practices of interdisciplinary communities actively involved in increasing scientific knowledge about resource overexploitation, degraded environment and ecosystem

Scientists producing indicators

More solid ground for **informed advice to competent authorities**

Trainers

Cost effective training & **knowledge bridging** between research and innovation

Enterprises

Enlarged spectrum of **growth opportunities**

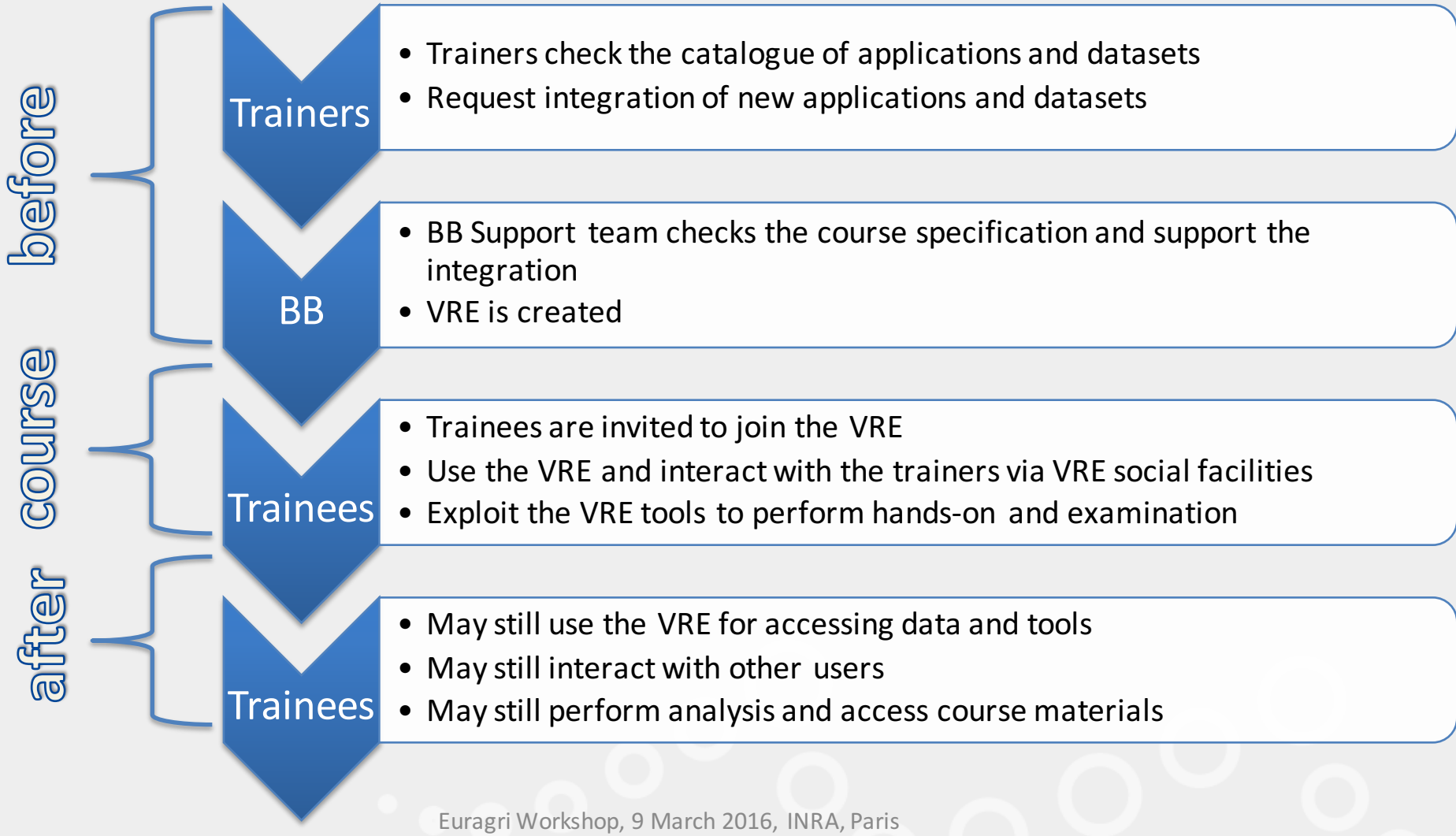
Six interrelated detailed objectives

Blue Assessment	Supporting the collaborative production of scientific knowledge required for **assessing the status of fish stocks** and **producing a global record of stocks and fisheries**
Blue Economy	Supporting the production of scientific knowledge for **analysing socio-economic performance in aquaculture**
Blue Environment	Supporting the production of scientific knowledge **for fisheries & habitat degradation monitoring**
Blue Skill	**Boosting education and knowl. bridging** between research &innovation in the area of protection and mgmt of marine resources
Blue Commons	Developing and deploying **service and resource commons** across VREs to facilitate the exploitation of existing infrastructure resources
Blue Uptake	**Ensuring uptake of the BlueBRIDGE tools and services**, with specifc focus on SMEs, other scientific domains & policy making contexts

Specific Areas

Blue Assessment	• Stock Assessment VRE • Global record of Stocks and Fisheries VRE
Blue Economy	• Performance evaluation, benchmarking and decision making in aquaculture VRE • Strategic Investment analysis and Scient. Planning/Alerting VRE
Blue Environment	• Aquaculture Atlas Generation VRE • Protected Area Impact Maps VRE
Blue Skill	• ICES Knowledge Bridging • IRD Knowledge Bridging • Knowledge Bridging Programmes

Supported Training Workflow



Final Remarks

- The marine and agriculture sectors have many similar needs in term of data management services
- By construction D4Science/iMarine have been built for being able to easy accomodates other needs (e.g. framework, standard protocols, integration of external algorithms, VREs)
- Open to investigate collaborations and synergies

www.i-marine.eu

www.bluebridge-vres.eu

https://www.gcube-system.org/

www.d4science.org