

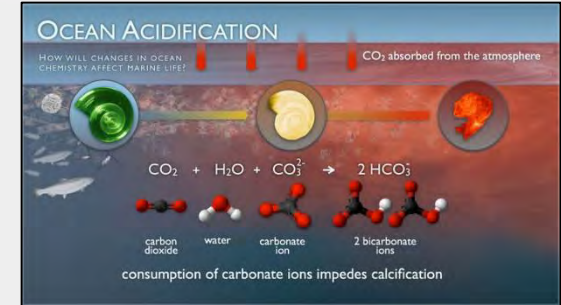
BlueBRIDGE



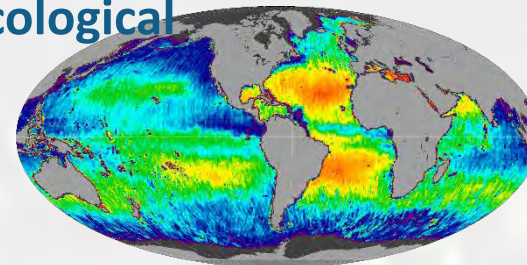
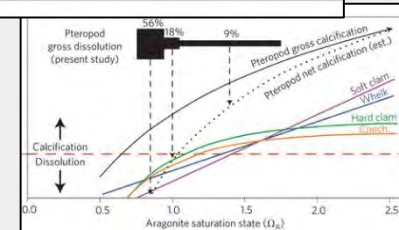
**Bridging environmental data providers and
SeaDataNet DIVA service
within a collaborative and distributed
e-Infrastructure**

Gianpaolo Coro, Pasquale Pagano, Umberto Napolitano
ISTI-CNR, Italy

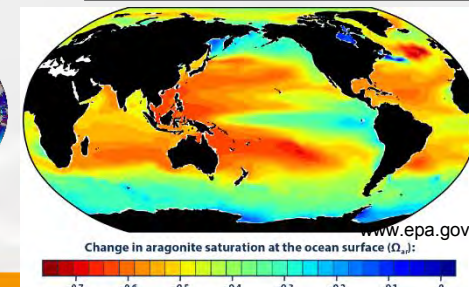




From Pmel-NOAA - <http://www.pmel.noaa.gov/co2/story/Ocean+Acidification>



<http://earthobservatory.nasa.gov/>



www.epa.gov

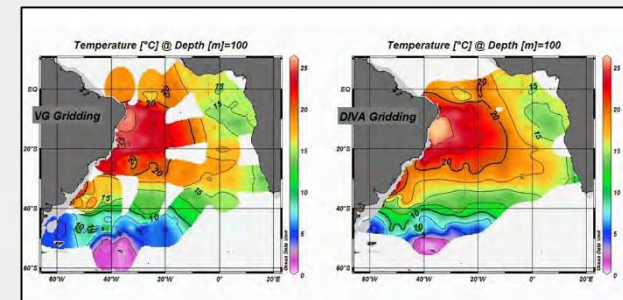
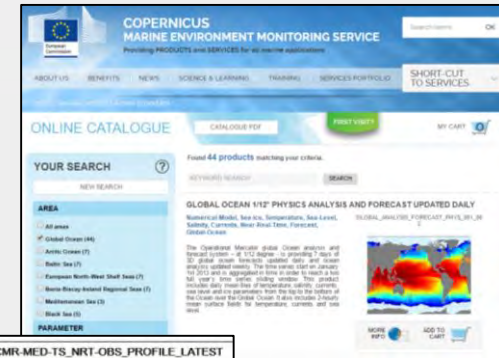
- Among its devastating effects, **ocean acidification harms** organisms whose life depends on **shells** and on **coral reefs**. It affects important marine ecosystems hosting high biodiversity and food availability

- Ries, J. B. (2012). Oceanography: A sea butterfly flaps its wings. *Nature Geoscience*, 5(12), 845-846.

Highlighted the **decrease of calcification** of a planktonic mollusc, the sea butterfly (*Limacina helicina*). The shell-building capacity of this organism declines with **decreasing aragonite saturation**, which is due to the decrease of average pH in the seas.

- Monitoring and containing ocean acidification** helps preventing coral reefs and shell-building organisms to dissolve, and thus helps **preventing ecological disasters**.

- **Environmental observations** of parameters like aragonite saturation and pH are usually available as **scattered *in situ* data**, published on restricted-access data e-Infrastructures (e.g. the Copernicus Marine Environment Monitoring Service)
- **Interpolation services exist** (e.g. the SeaDataNet Data-Interpolating Variational Analysis service, DIVA) to estimate global, uniform distributions of environmental parameters from scattered observations.
- However, these services usually **require data to be compliant with a non-standard format** and cannot accept *in situ* data formats directly.
- Furthermore, they may **benefit from facilities to communicate and to publish** their results for a large public.



e-Infrastructures enable researchers at different locations across the world to **collaborate** in the context of their home institutions or in national or multinational scientific initiatives.

- People can work together having shared access to unique or distributed scientific facilities (including data, instruments, computing and communications).

Examples:

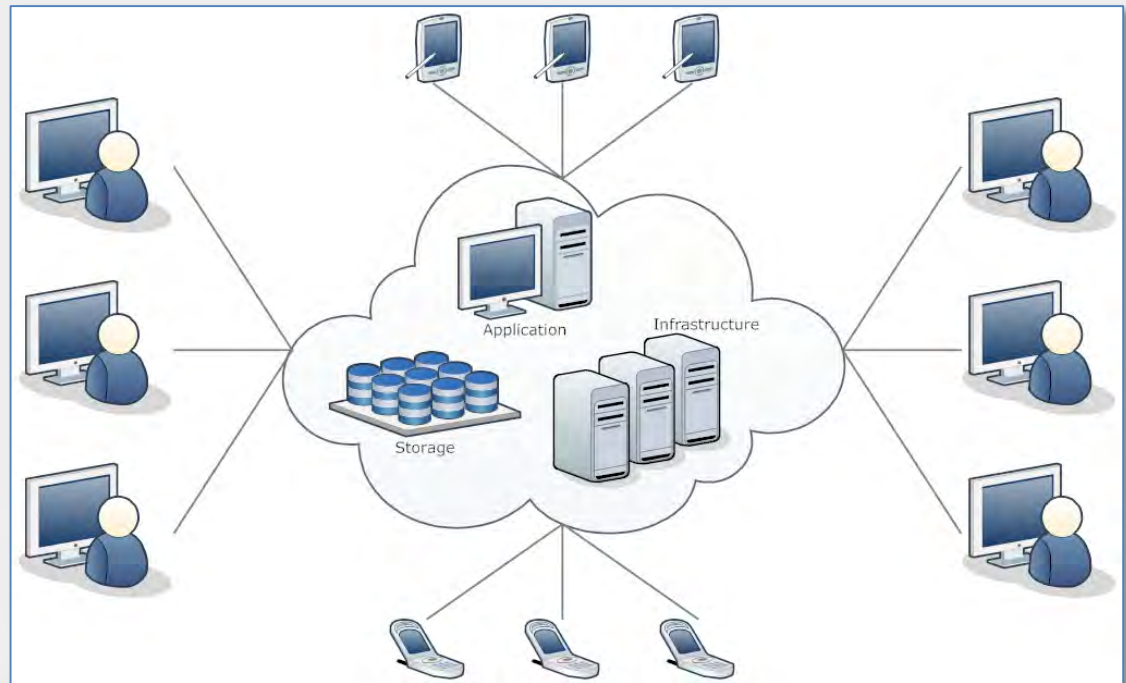
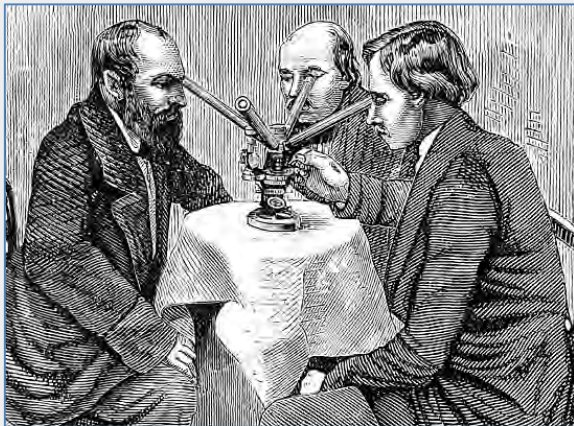
Belief, <http://www.beliefproject.org/>

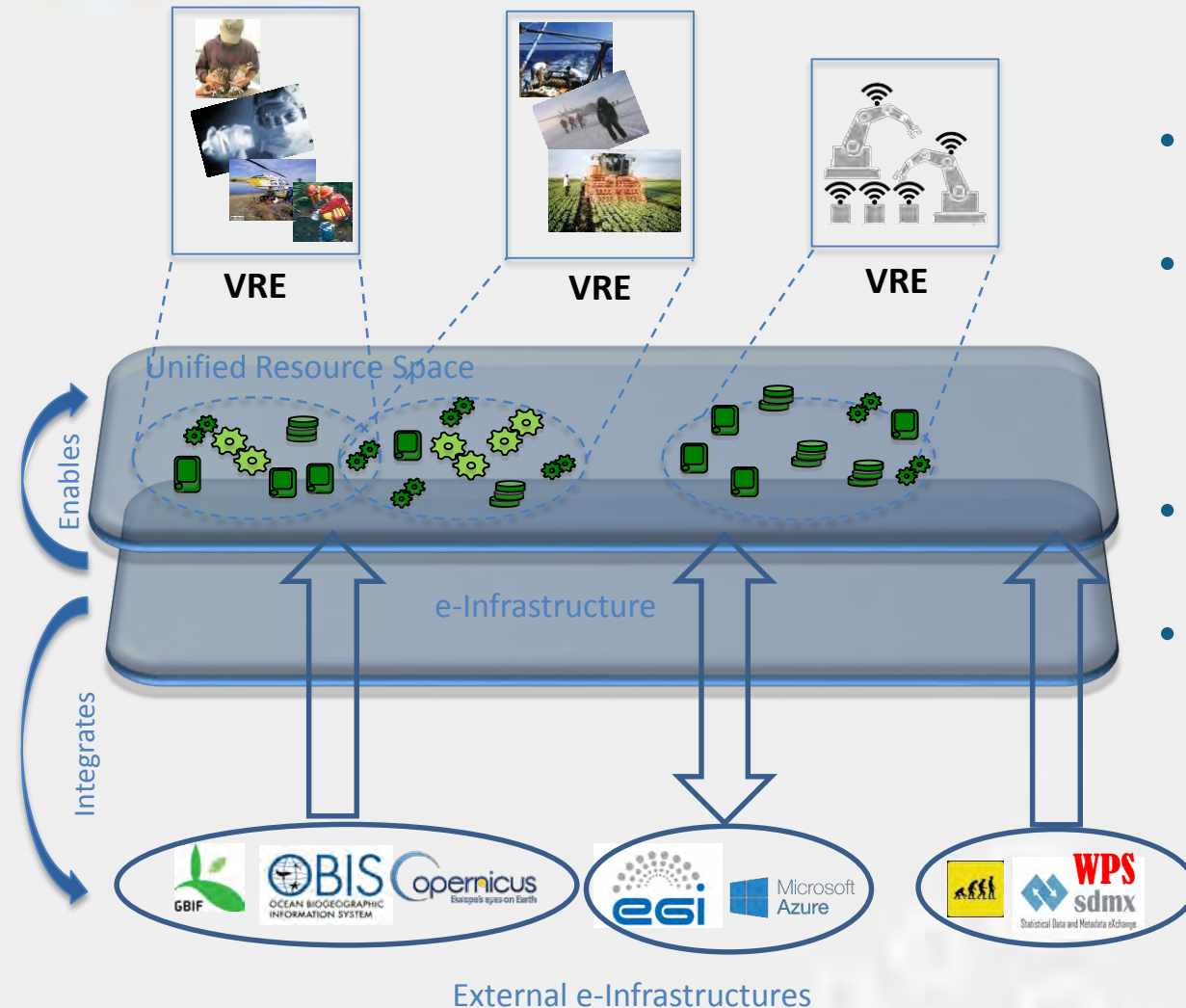
OpenAire, <http://www.openaire.eu/>

i-Marine, <http://www.i-marine.eu/>

EU-Brazil OpenBio,

<http://www.eubrazilopenbio.eu/>





- Define sub-communities
- Allow temporary dedicated assignment of computational, storage, and data resources
- Manage policies
- Support data and information sharing

New Science Paradigms

- ❑ **Open Science:** make scientific research, data and dissemination accessible to all levels of an inquiring society, amateur or professional.

Keywords: Open Access, Open research, Open Notebook Science

- ❑ **E-Science:** computationally intensive science is carried out in highly distributed network environments that use large data sets and require distributed computing and collaborative tools.

Keywords: Provenance of the scientific process, Scientific workflows

- ❑ **Science 2.0:** process and publish large data sets using a collaborative approach. Share from raw data to experimental results and processes. Support collaborative experiments and Reproducibility-Repeatability-Reusability (R-R-R) of Science.

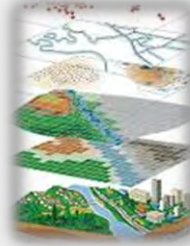
Keywords: collaborative and repeatable Science



Databases



Cloud storage



Geospatial data

Storage

Data management



Metadata generation
and management



Harmonisation



Sharing



Cloud computing



Elastic resources
assignment



Multi-platform: R,
Java, Fortran

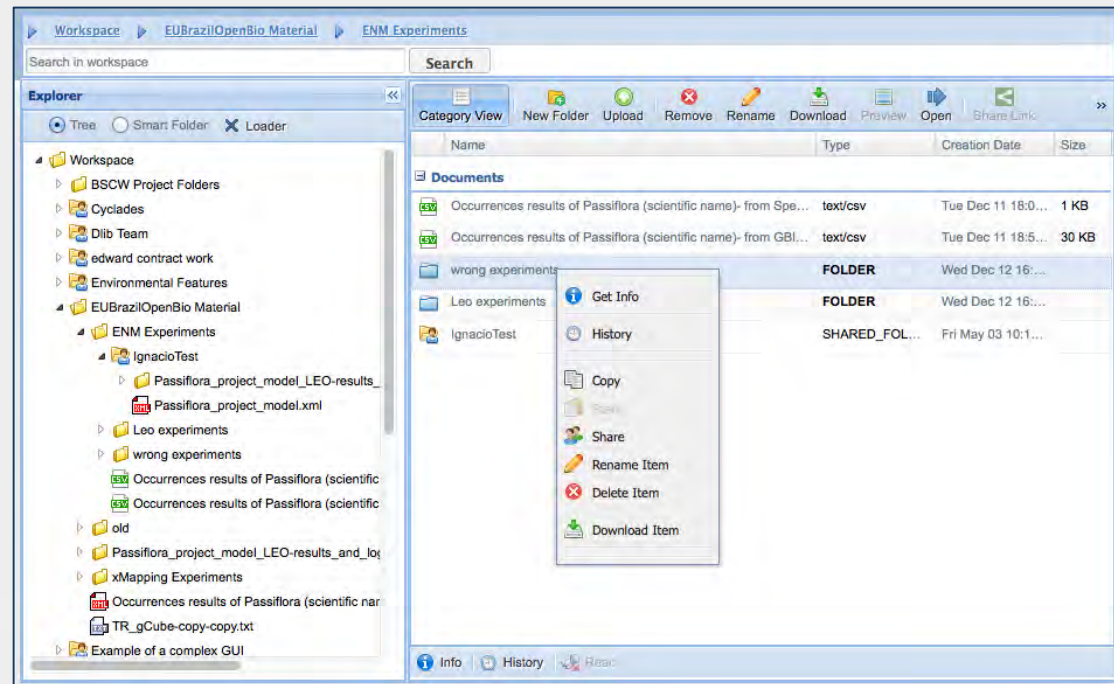
Processing

The Workspace – an online files storage system

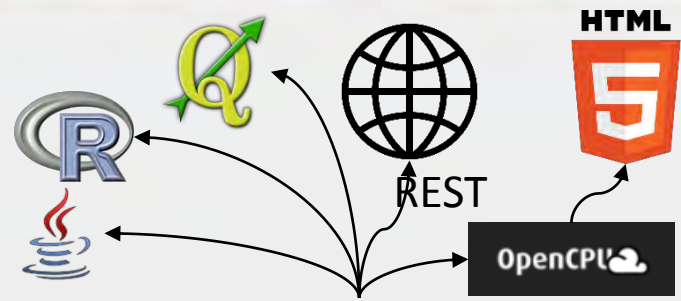
A free-of-use folder-based file system allows managing and sharing information objects.

Information objects can be

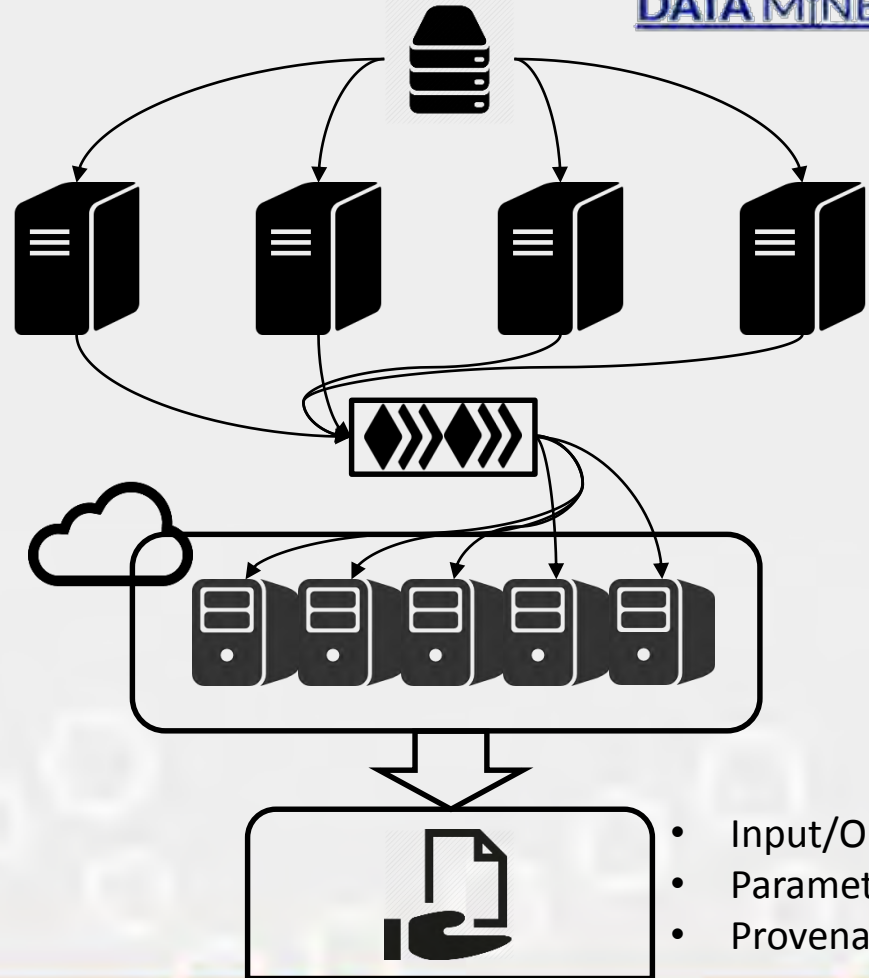
- files, dataset, workflows, experiments, etc.
- organized into folders
- shared
- disseminated via public URLs



- Experiments on Big Data
- Sharing inputs and results
- Save the *provenance* of experiments
- Supports R-R-R of experiments



DATA MINER

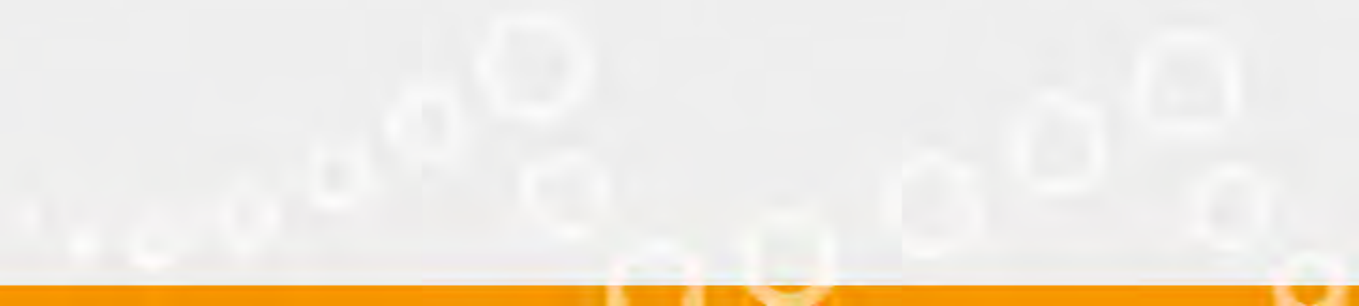


- Input/Out
- Parameters
- Provenance

The screenshot shows the DATA MINER web interface. The left sidebar lists various analysis categories like Bayesian Methods, Charts, Classification, etc. The main area displays the configuration for the 'DbSCAN' operator. The 'OccurrencePointsTable' parameter is set to 'Temp Bari.csv'. The 'FeaturesColumnNames' section shows a list of columns with checkboxes for selection. A 'Start Computation' button is visible at the bottom.



Proposed solution

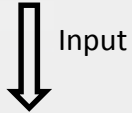
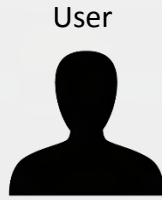




The SeaDataNet-BB Connector Architecture

Excel window: v - Excel

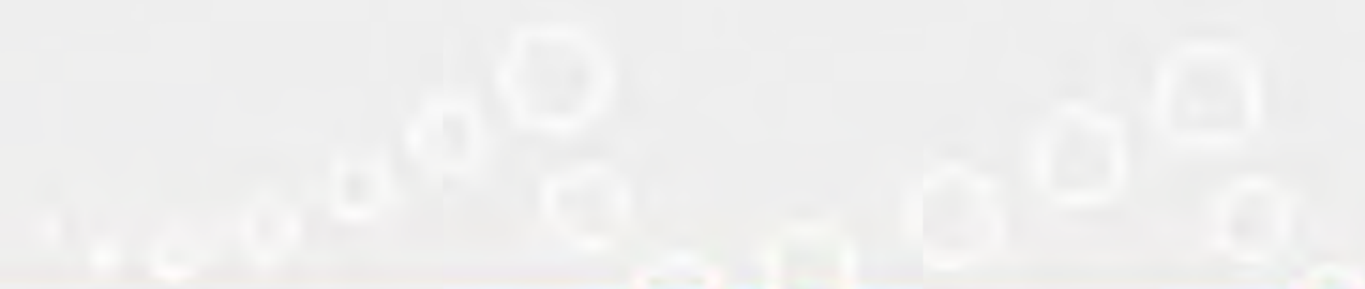
	A	B	C
1	latitude	longitude	value
2	20.5391	58.3623	3.585
3	20.5365	58.3605	3.66
4	20.5339	58.3586	3.435
5	20.5312	58.3568	3.51
6	20.5285	58.355	3.51



Sharing



Other user





The SeaDataNet-BB Connector Architecture

Excel window: v - Excel

A	B	C
1	latitude	longitude value
2	20.5391	58.3623 3.585
3	20.5365	58.3605 3.66
4	20.5339	58.3586 3.435
5	20.5312	58.3568 3.51
6	20.5285	58.355 3.51

Input



Sharing

Other user

User



VRE



WPS



DATA MINER





The SeaDataNet-BB Connector Architecture

A1	A	B	C
1	latitude	longitude	value
2	20.5391	58.3623	3.585
3	20.5365	58.3605	3.66
4	20.5339	58.3586	3.435
5	20.5312	58.3568	3.51
6	20.5285	58.355	3.51

Input



Sharing

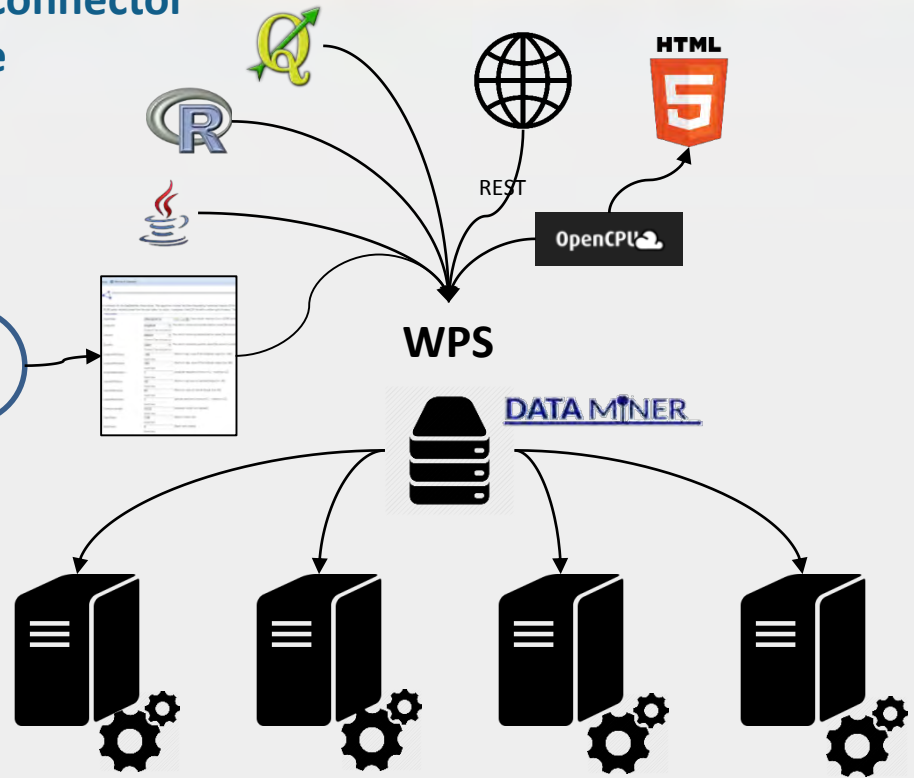
Work-space

Other user

User

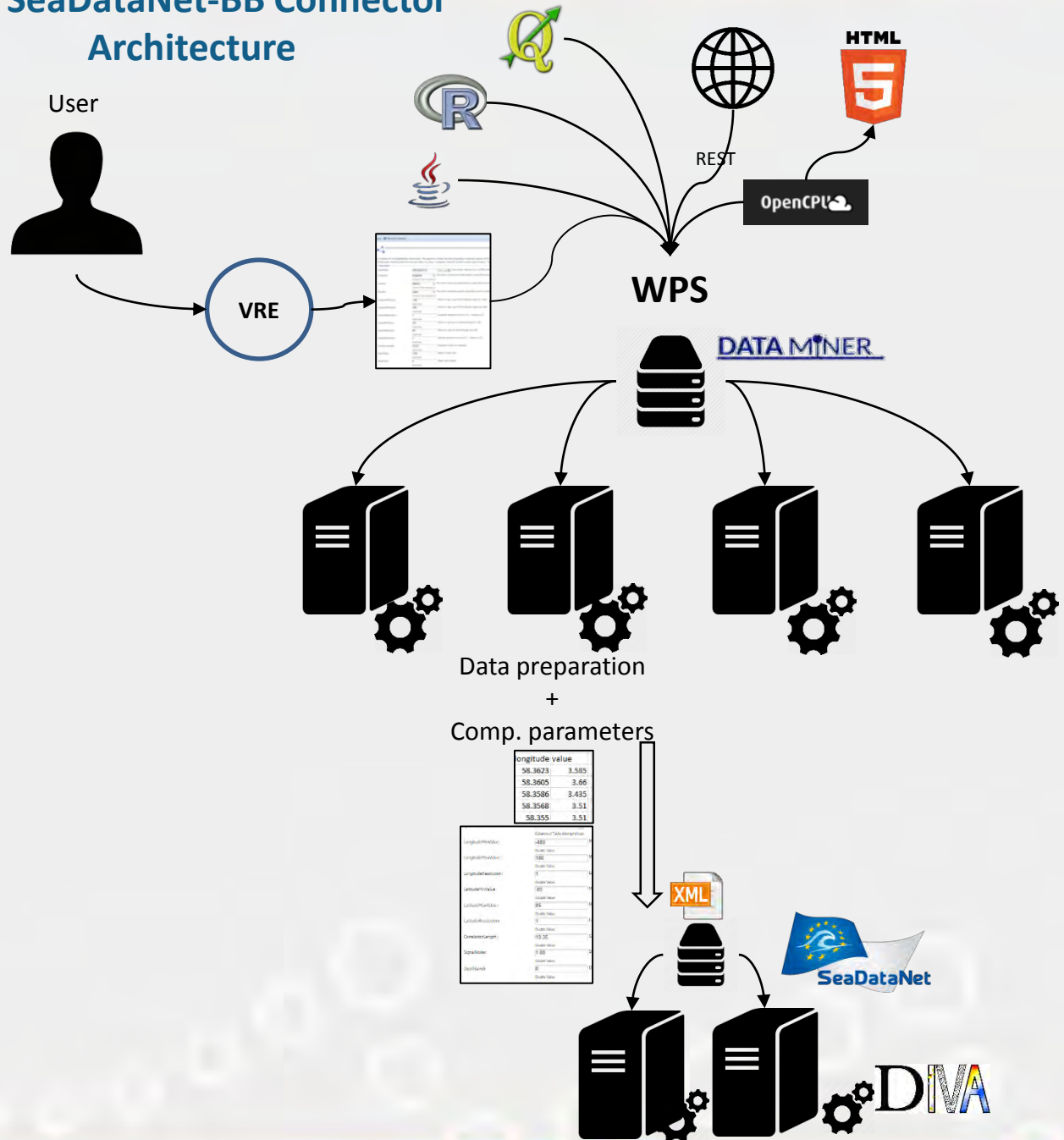
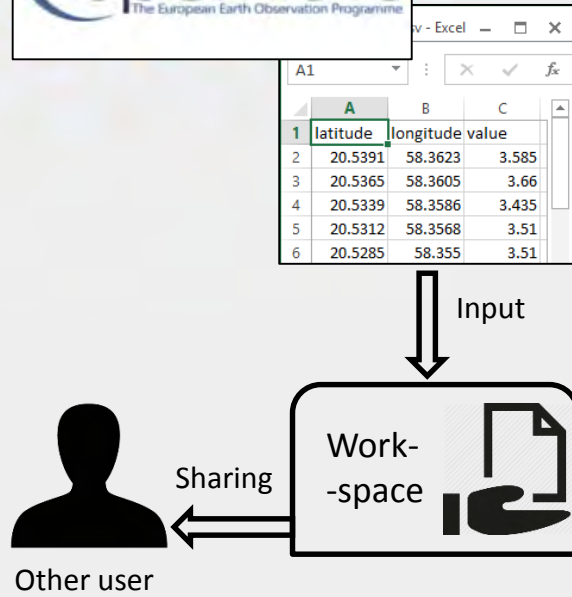


VRE





The SeaDataNet-BB Connector Architecture

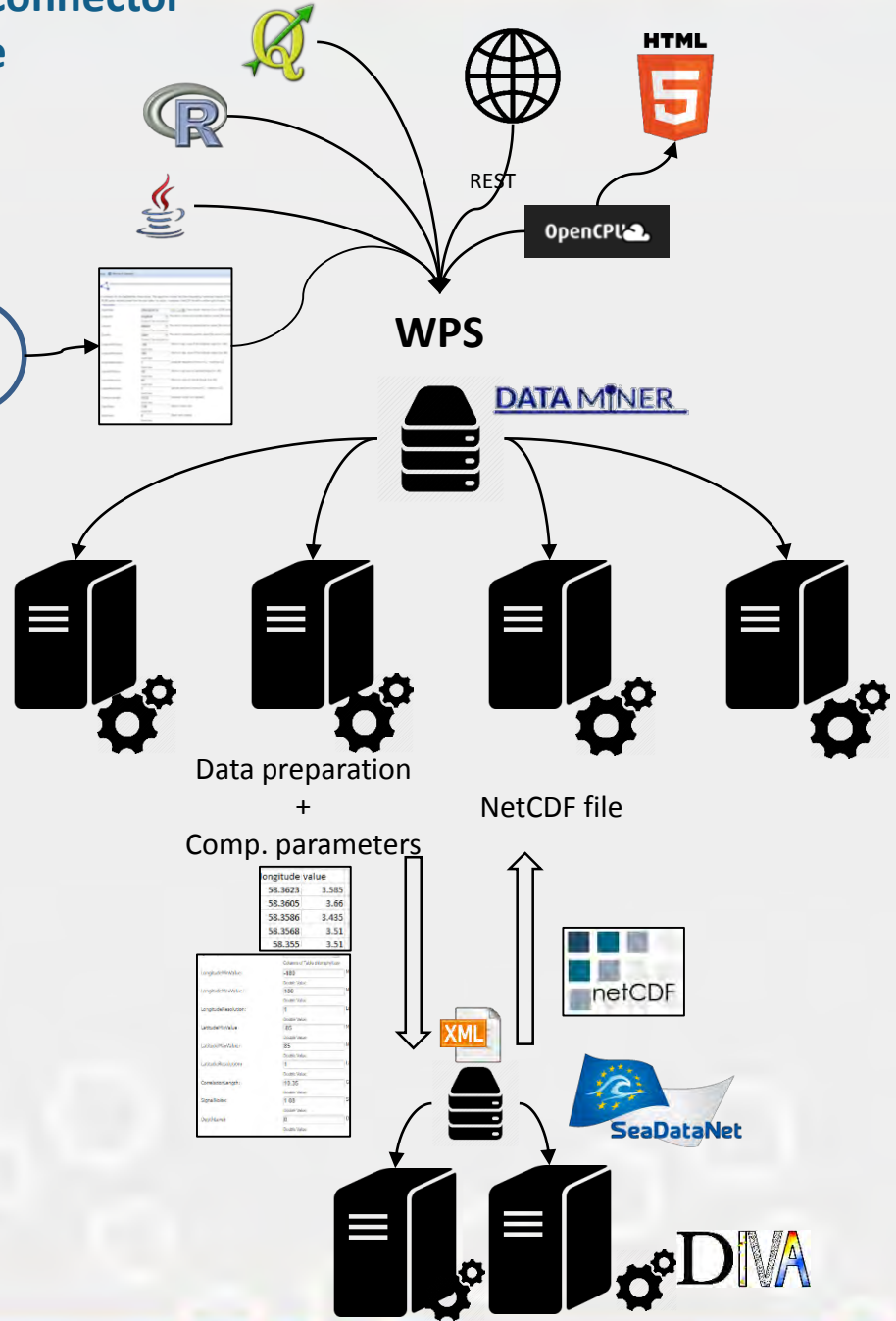
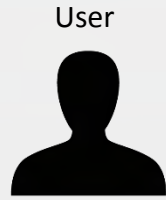
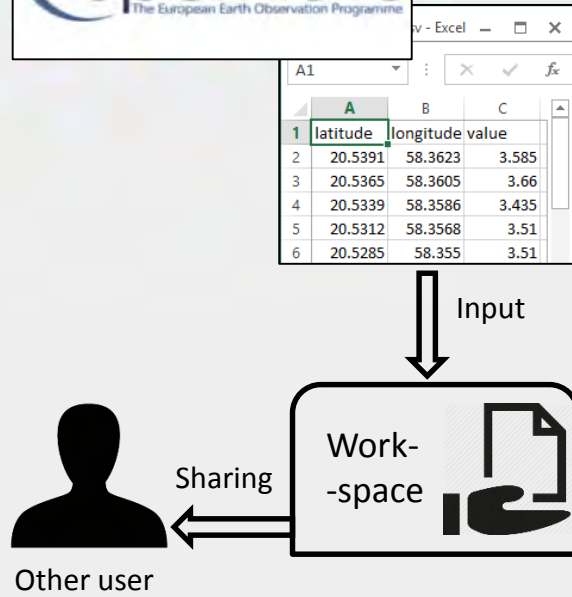


longitude value	3.585
58.3623	3.585
58.3605	3.66
58.3586	3.435
58.3568	3.51
58.355	3.51

Class	Value
lat	20.5391
lon	58.3623
val	3.585
lat	20.5365
lon	58.3605
val	3.66
lat	20.5339
lon	58.3586
val	3.435
lat	20.5312
lon	58.3568
val	3.51
lat	20.5285
lon	58.355
val	3.51



The SeaDataNet-BB Connector Architecture

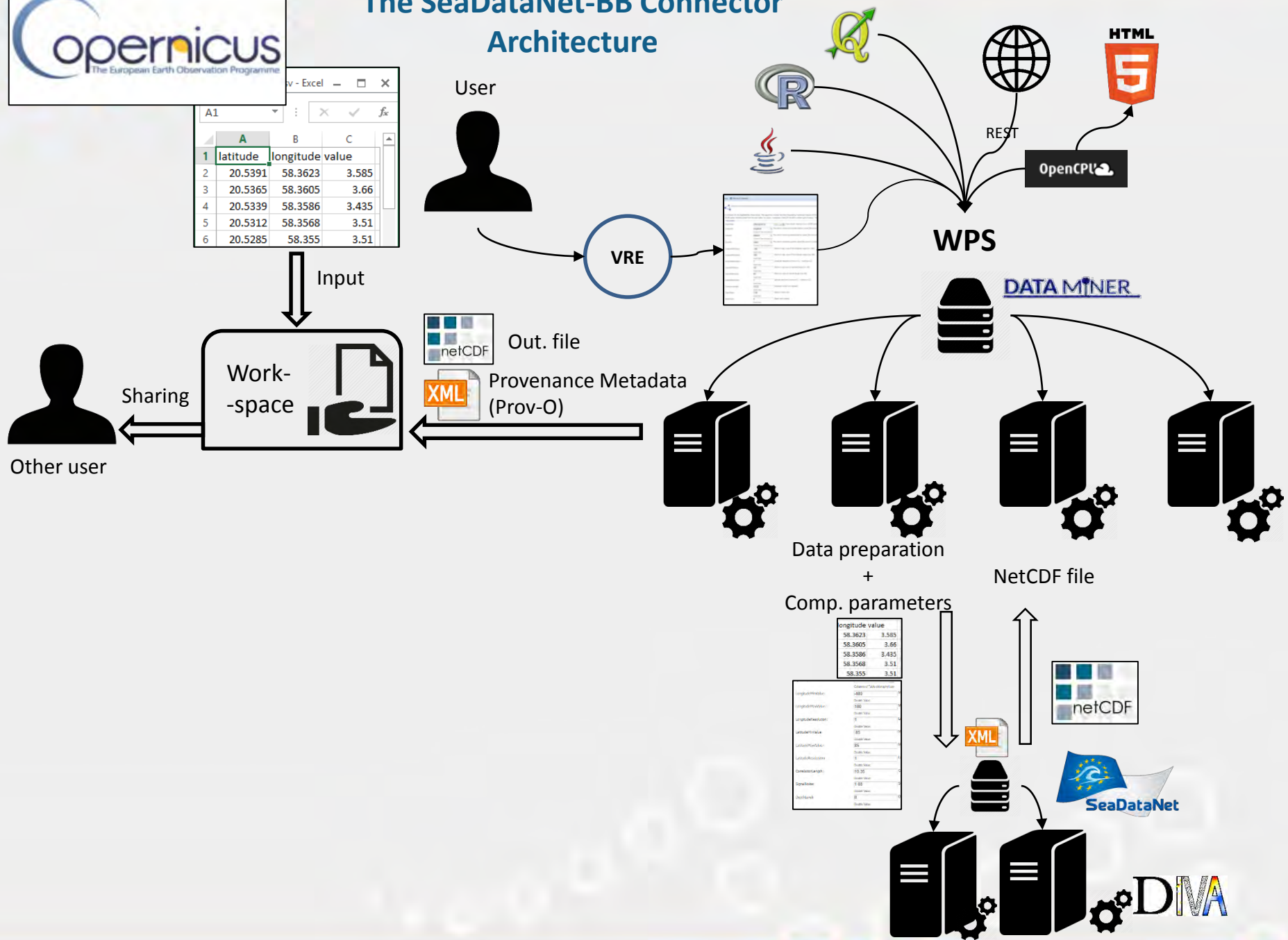


longitude Value	3.585
58.3623	3.585
58.3605	3.66
58.3586	3.435
58.3568	3.51
58.355	3.51



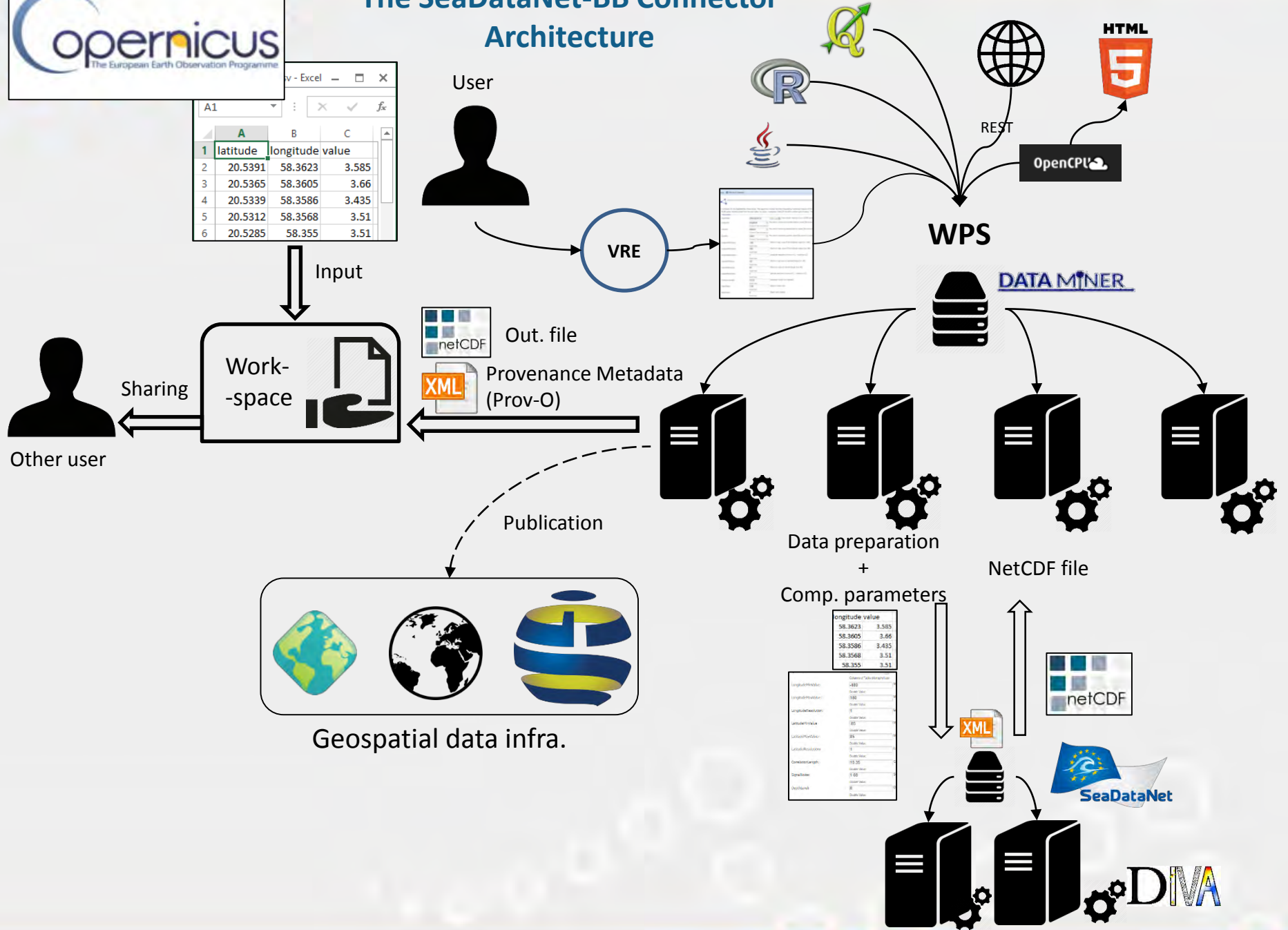


The SeaDataNet-BB Connector Architecture



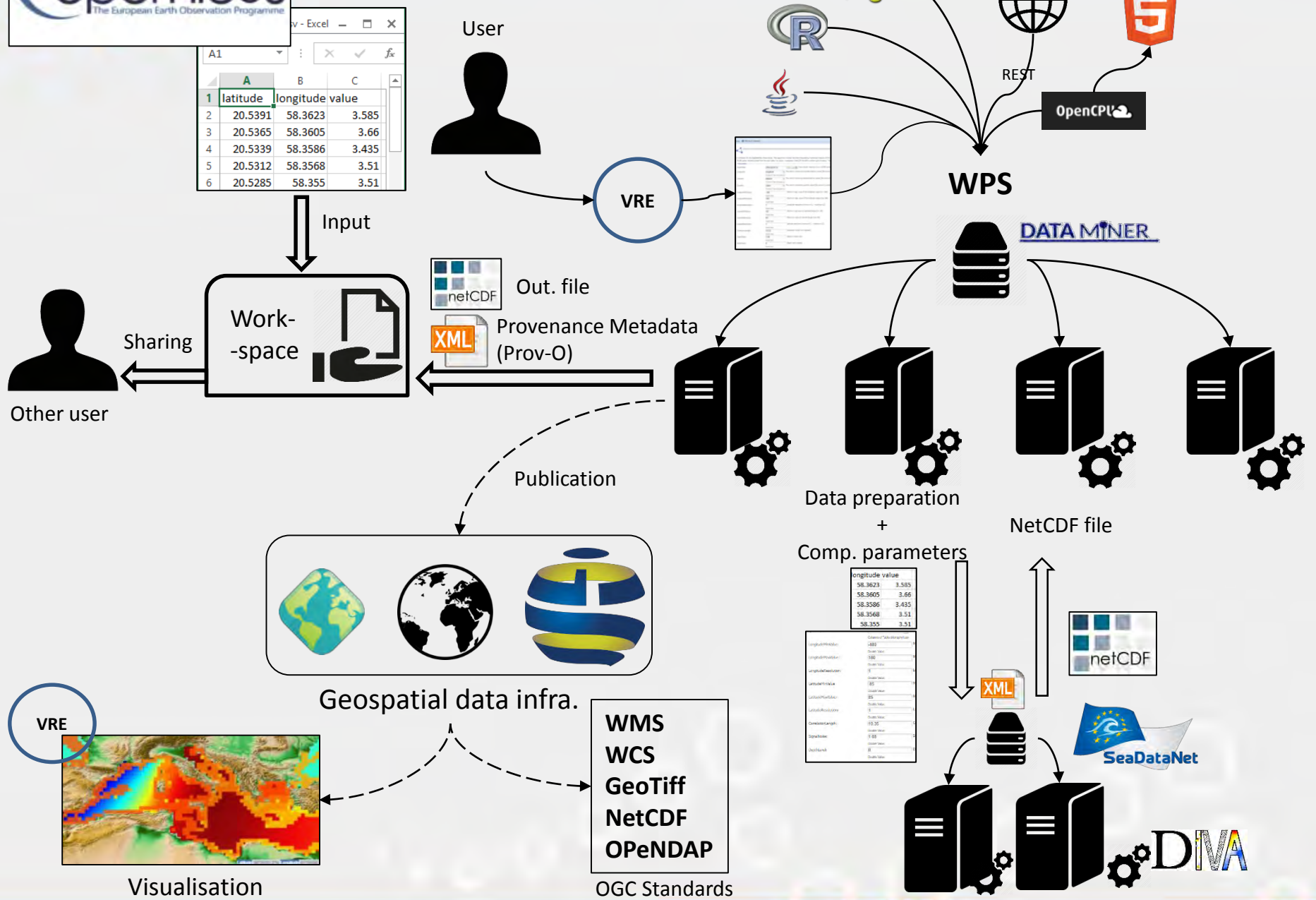


The SeaDataNet-BB Connector Architecture





The SeaDataNet-BB Connector Architecture





Example

Demo Video : <http://goo.gl/yX3kww>

YOUR SEARCH ?

NEW SEARCH

AREA

- All areas
- Global Ocean (41)
- Arctic Ocean (12)
- Baltic Sea (8)
- European North-West Shelf Seas (9)
- Iberia-Biscay-Ireland Regional Seas (9)
- Mediterranean Sea (5)
- Black Sea (6)

PARAMETER

- All parameters
- Ocean Temperature (18)
- Ocean Salinity (13)
- Ocean Currents (10)
- Sea Ice (14)
- Sea Level (17)
- Winds (4)
- Ocean Optics (2)
- Ocean Chemistry (4)
- Ocean Biology (2)
- Ocean Chlorophyll (7)

TIME COVERAGE

- All time coverages

Found 41 products matching your criteria.

KEYWORD SEARCH

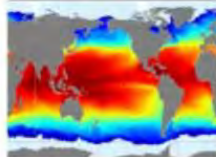
SEARCH

GLOBAL OCEAN 1/12° PHYSICS ANALYSIS AND FORECAST UPDATED DAILY

Numerical model, Sea-ice, Sea-level, Temperature, Salinity, Currents, Forecast, Near-real-time, Global-ocean

GLOBAL_ANALYSIS_FORECAST_PHY5_001_002

The Operational Mercator global Ocean analysis and forecast system – at 1/12 degree - is providing 7 days of 3D global ocean forecasts updated daily and ocean analysis updated weekly. The time series start on January 1st 2013 and is aggregated in time in order to reach a two full year's time series sliding window. This product includes daily mean files of temperature, salinity, currents, sea level and ice parameters from the top to the bottom of the Ocean over the Global Ocean. It also includes 2-hourly mean surface fields for temperature, salinity and currents.



MORE INFO

ADD TO CART

GLOBAL OCEAN BIOGEOCHEMISTRY ANALYSIS AND WEEKLY FORECAST

Numerical model, Ocean-chlorophyll, Ocean-chemistry, Ocean-biology, Near-real-time, Global-ocean

GLOBAL_ANALYSIS_FORECAST_BIO_001_014

Product GLOBAL_ANALYSIS_FORECAST_BIO_001_014 produced by Mercator Ocean in Toulouse, France, is a global Ocean Biogeochemical analysis product at 1/2°. It is providing a 7-days mean global forecast updated weekly as well as 3D global ocean biogeochemical weekly mean analysis for the past 2 years updated every week. This product includes weekly mean files of dissolved iron, nitrate, phosphate, silicate, dissolved oxygen, chlorophyll concentration, phytoplankton concentration and primary production parameters from the top to the bottom of the Global Ocean on a 1/2° regular grid projection interpolated from the 1/4° ERA-Interim-C reanalysis. Vertical coverage includes 50

STEP 1: get *in situ* or satellite data

MY CART

INSITU_MED_NRT_OBSERVATIONS_013_035

Mediterranean Sea- In-Situ Near Real Time Observations

Product id: INSITU_MED_NRT_OBSERVATIONS_013_035

Dataset:

Oceanotron-INS-HCMR-MED-TS_NRT-OBS_PROFILE_LATEST

Variable:

Oceanotron-INS-HCMR-MED-TS_NRT-OBS_PROFILE_LATEST > TEMP

Units: degree_Celsius Time: 2015-12-09 00:00:00.000Z +/- 1 day Depth (m): 0



Example: Copernicus Marine Environment Monitoring Service

STEP 2: extract data in tabular format and upload them on the Workspace

chlorophyll.csv - Excel

	A	B	C
1	latitude	longitude	value
2	20.5391	58.3623	3.585
3	20.5365	58.3605	3.66
4	20.5339	58.3586	3.435
5	20.5312	58.3568	3.51
6	20.5285	58.355	3.51
7	20.5258	58.3531	3.51
8	20.5232	58.3513	3.585
9	20.5204	58.3494	3.66
10	20.5177	58.3476	3.66
11	20.5151	58.3458	3.735
12	20.5124	58.344	3.81
13	20.5097	58.3421	3.585
14	20.5071	58.3403	3.585
15	20.5044	58.3385	3.36
16	20.5017	58.3367	3.36
17	20.499	58.3349	3.285
18	20.4964	58.333	3.285
19	20.4936	58.3312	3.36
20	20.491	58.3294	3.435
21	20.4883	58.3276	3.36
22	20.4856	58.3257	3.585
23	20.483	58.3239	3.51
24	20.4803	58.3221	3.435
25	20.4775	58.3203	3.585
26	20.4748	58.3184	3.51
27	20.4721	58.3166	3.585
28	20.4695	58.3148	3.585
29	20.4668	58.313	3.585
30	20.4641	58.3111	3.585
31	20.4614	58.3093	3.735
32	20.4587	58.3074	3.435
33	20.4561	58.3056	3.51
34	20.4534	58.3037	3.435
35	20.4508	58.3019	3.585
36	20.4481	58.3	3.435
37	20.4455	58.2982	3.51
38	20.4429	58.2964	3.585
39	20.4402	58.2945	3.435
40	20.4375	58.2927	3.66
41	20.4349	58.2908	3.435
42	20.4323	58.2889	3.51
43	20.4297	58.2871	3.435
44	20.4272	58.2852	3.435
45	20.4245	58.2834	3.585

Drag and drop

Search by name

Search

Tree SmartFolder

undefined's workspace

- VRE Folders
- 2015_Sept_21-22-RDP_Workshop
- algorithm_shark
- Articles
- data hypoheic experiment
- DataMiner
- DataMinerAlgorithms
- DataMinerInputs
- Datasets
- ensembleModel
- esperimento
- Experiments
- grida_intersect
- Groups
- Ichthyop
- Ichthyop_multiple_runs
- Ichthyop_netCDF_CF_compliant
- Ichthyop_netCDF_to_shapefile
- ICOS_test
- ICOS_test_v2
- input files
- KMEANS_ID_d5da7b73-696a-4090-9f97-7af5
- Management presentation
- Meetings and presentations
- mpa_intersect_v2
- NewNewCaros
- Other Shares
- Output Data Sets
- Output Data Sets
- Parall_STEPS_ICCAT_BFTE_StockAssesme
- Projects
- Prova CondiVisa
- SAI
- SAI Examples
- SAIExperiment
- shark_abundancy
- shark_tests
- statistical-manager-figs-algorithms
- StatisticalAlgorithmsImporter
- STEPS_ICCAT_BFTE_StockAssessment
- SupplierFolder
- test

Documents

Name	Owner	Type	Last Update	Size
chlorophyll.csv	Gianpaolo Coro	textcsv	04 Oct 11:49 A...	10 KB
WKLIFE4ID.csv	Gianpaolo Coro	textcsv	28 Sep 10:43 P...	2 KB
WKLIFE45stocks.csv	Gianpaolo Coro	textcsv	28 Sep 10:43 P...	15 KB
hspen21.csv	Gianpaolo Coro	textcsv	23 Sep 01:16 A...	5 KB
hspen.csv	Gianpaolo Coro	textcsv	23 Sep 01:12 A...	2,583...
DAPSTOM.csv	Gianpaolo Coro	textcsv	28 Jul 02:20 P...	2 KB
NS_DIET.csv	Gianpaolo Coro	textcsv	28 Jul 02:20 P...	1 KB
FISHING_PARAM_MSY.csv	Gianpaolo Coro	textcsv	28 Jul 02:19 P...	1 KB
FISHING_PARAM_HIST.csv	Gianpaolo Coro	textcsv	28 Jul 02:19 P...	1 KB
IVfile.csv	Gianpaolo Coro	textcsv	28 Jul 02:19 P...	3 KB
NS_PARAMS.csv	Gianpaolo Coro	textcsv	28 Jul 02:18 P...	1 KB
waa.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	12 KB
test taxa_sarda.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
temperature_argo_small.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
Temp Bari1.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
Temp Bari1.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	5 KB
taxatest_gadius.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
taxa1024.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	40 KB
taxa1.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
tacsatmin13.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
tacsatmin12.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	67 KB
tacsatmin1.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	108 KB
tacsatmin1 (1).csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	108 KB
TACSAT2_example.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	69 KB
speciesTaxa.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB
speciesforabsence.txt	Gianpaolo Coro	textpl...	24 Jul 01:21 A...	1 KB
simpleWebsite (1).zip	Gianpaolo Coro	applic...	24 Jul 01:21 A...	3,722...
shapefiletest.zip	Gianpaolo Coro	applic...	24 Jul 01:21 A...	3 KB
run_config.xml	Gianpaolo Coro	applic...	24 Jul 01:21 A...	1 KB
RawTaxaNamesTable.txt	Gianpaolo Coro	textpl...	24 Jul 01:21 A...	1 KB
PCAA_Age1_25_Run3.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	32 KB
OccurrencePointsTable.txt	Gianpaolo Coro	textpl...	24 Jul 01:21 A...	1 KB
OccurrencePointsTable.csv	Gianpaolo Coro	textcsv	24 Jul 01:21 A...	1 KB

STEP 3: setup the experiment through the DataMiner interface

Seadatanet Interpolator

A connector for the SeaDataNet infrastructure. This algorithm invokes the Data-Interpolating Variational Analysis (DIVA) SeaDataNet service to interpolate spatial data. The model uses GEBCO bathymetry data and requires an estimate of the maximum spatial span of the correlation between points and the signal-to-noise ratio, among the other parameters. It can interpolate up to 10,000 points randomly taken from the input table. As output, it produces a NetCDF file with a uniform grid of values. This powerful interpolation model is described in Troupin et al. 2012, 'Generation of analysis and consistent error fields using the Data Interpolating Variational Analysis (Divav)', Ocean Modelling, 62-53, 90-101.

Parameters

InputTable: **chlorophyll.csv** Input tabular resource. Up to 10,000 points will be randomly taken from this table. [a http link to a table in UTF-8 encoding following this template: ((GENERIC)) A generic comma separated csv file in UTF-8 encoding]

Longitude: **longitude** The column containing longitude decimal values [the name of a column from InputTable]
Columns of Table chlorophyll.csv

Latitude: **latitude** The column containing latitude decimal values [the name of a column from InputTable]
Columns of Table chlorophyll.csv

Quantity: **value** The column containing quantity values [the name of a column from InputTable]
Columns of Table chlorophyll.csv

LongitudeMinValue: **-180** Minimum deg. value of the longitude range (min -180)
Double Value

LongitudeMaxValue: **180** Maximum deg. value of the longitude range (max 180)
Double Value

LongitudeResolution: **1** Longitude resolution (minimum 0.1 - maximum 10)
Double Value

LatitudeMinValue: **-85** Minimum deg value of Latitude Range (min -85)
Double Value

LatitudeMaxValue: **85** Maximum value of Latitude Range (max 85)
Double Value

LatitudeResolution: **1** Latitude resolution (minimum 0.1 - maximum 10)
Double Value

CorrelationLength: **10.35** Correlation length (arc degrees)
Double Value

SignalNoise: **1.08** Signal to noise ratio
Double Value




DepthLevel: **0** Depth level (meters)
Double Value


Select CSV


Name	Owner	Created
1_giorgiole_core_neurax	Giorgiole Caro	24 Jul 01:20 AM 2016
atom_gem_giorgiole_test1	Giorgiole Caro	24 Jul 01:20 AM 2016
ANDport.csv	Giorgiole Caro	24 Jul 01:20 AM 2016
CAA_Aqut_25.csv	Giorgiole Caro	24 Jul 01:20 AM 2016
interpolated.csv	Giorgiole Caro	16 Oct 11:29 AM 2016
CPUE_Phar3.csv	Giorgiole Caro	24 Jul 01:20 AM 2016
EZSst1.csv	Giorgiole Caro	24 Jul 01:20 AM 2016
DAPSTOM.csv	Giorgiole Caro	28 Jul 02:18 PM 2016
DAPSTOM.txt	Giorgiole Caro	24 Jul 01:20 AM 2016
Necab.csv	Giorgiole Caro	24 Jul 01:20 AM 2016
FISHING_PARAM_HST1	Giorgiole Caro	28 Jul 02:18 PM 2016
FISHING_PARAM_MST1	Giorgiole Caro	28 Jul 02:18 PM 2016


A connector for the SeaDataNet infrastructure. This algorithm invokes the Data-Interpolating Variational Analysis (DIVA) SeaDataNet service to interpolate 10,000 points randomly taken from the input table. As output, it produces a NetCDF file with a uniform grid of values. This powerful interpolation model is

Parameters

InputTable:    Input tabular resource. Up to 10,000 points will be randomly taken from the

Longitude:  The column containing longitude decimal values [the name of a column from InputTable]
Columns of Table chlorophyll.csv

Latitude:  The column containing latitude decimal values [the name of a column from InputTable]
Columns of Table chlorophyll.csv

Quantity:  The column containing quantity values [the name of a column from InputTable]
Columns of Table chlorophyll.csv

LongitudeMinValue: Minimum deg. value of the longitude range (min -180)
Double Value

LongitudeMaxValue: Maximum deg. value of the longitude range (max 180)
Double Value

LongitudeResolution: Longitude resolution (minimum 0,1 - maximum 10)
Double Value

LatitudeMinValue: Minimum deg value of Latitude Range (min -85)
Double Value

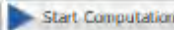
LatitudeMaxValue: Maximum value of Latitude Range (max 85)
Double Value



LatitudeResolution: Latitude resolution (minimum 0,1 - maximum 10)
Double Value

CorrelationLength: Correlation length (arc degrees)
Double Value

SignalNoise: Signal to noise ratio
Double Value

DepthLevel: Depth level (meters)
Double Value

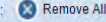



Operators   User Perspective


FILTERED RESULTS (1 ITEM FOUND)

Seadatanet Interpolator
A connector for the SeaDataNet infrastructure. This algorithm invokes the Data-Interpolating Variational Analysis (DIVA) SeaDataNet service to interpolate s...

Computations Execution

Tools: 

Computation of **Seadatanet Interpolator**
Created, the id is 48619e63-7fc9-4d82-aa49-e7065e6b24ff [link]
Equivalent Get Request: 

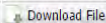


The computation **Seadatanet Interpolator** finished.

Output Values

Minimum value estimated by the model	3.60355
Maximum value estimated by the model	3.77208
Number of observations used	496
A posteriori estimate of signal-to-noise ratio	0.044272

Files

Output file in NetCDF format


STEP 5: get the results and inspect the metadata

Name	Created	operator_name	start_date	end_date	status	execution_platform	VRE
SEADATANET_INTERPOLAT 7fc9-4d82-aa49-e7065e6b24ff	04 Oct 11:55 AM 2016	SEADATANET_INTERPOLAT	04/10/2016 11:55:24	04/10/2016 11:55:47	completed	LOCAL	/d4science.research-infrastructures.eu/gCubeApps
BIONYM_LOCAL_ID_8e8c07 a834-4923-873f-ff5c59e2a857	30 Sep 05:08 PM 2016	BIONYM_LOCAL	30/09/2016 17:08:40	30/09/2016 17:08:45	completed	LOCAL	
BIONYM_ID_bf9b9ac7-4380-4278-a90e-70c9a81aeb06	29 Sep 05:48 PM 2016	BIONYM	29/09/2016 17:42:58	29/09/2016 17:48:23	completed	D4SCIENCE	
BIONYM_ID_58c2db0f16fa-49fd-abd7-337e02e6cf14	29 Sep 05:12 PM 2016	BIONYM	29/09/2016 17:09:06	29/09/2016 17:12:10	completed	D4SCIENCE	
BIONYM_ID_fcd0a7de-3ebf-4716-9467-be8b7e966e0d	29 Sep 05:06 PM 2016	BIONYM	29/09/2016 16:58:38	29/09/2016 17:06:40	completed	D4SCIENCE	
WEB_APP_PUBLISHER_ID_b4be-44e9-9e6d-9696280c2692	29 Sep 04:59 PM 2016	WEB_APP_PUBLISHER	29/09/2016 16:54:39	29/09/2016 16:59:24	completed	LOCAL	
HCAF_FILTER_ID_88ac1ebf-543a-4c0d-b649-70c45c70c70c	29 Sep 04:55 PM 2016	HCAF_FILTER	29/09/2016 16:55:49	29/09/2016 16:55:53	completed	LOCAL	

Output Result

- Number of observations used: 466
- Minimum value estimated by the model: 3.60355
- NetCDF Output File: NetCDFOutputFile [SEADATANET_INTERPOLATOR_ID_48619e03-7fc9-4d82-aa49-e7065e6b24ff].csv
- A posteriori estimate of signal-to-noise ratio: 0.044272
- Maximum value estimated by the model: 3.77208

Input Parameters

- LongitudeResolution: 1.0
- SignalNoise: 1.08
- Longitude: longitude
- Latitude: latitude
- LatitudeResolution: 1.0
- Quantity: value
- CorrelationLength: 10.35
- LongitudeMaxValue: 180.0
- LatitudeMinValue: -85.0
- DepthLevel: 0.0
- LatitudeMaxValue: 85.0
- LongitudeMinValue: -180.0
- InputTable: InputTable [SEADATANET_INTERPOLATOR_ID_48619e03-7fc9-4d82-aa49-e7065e6b24ff].csv

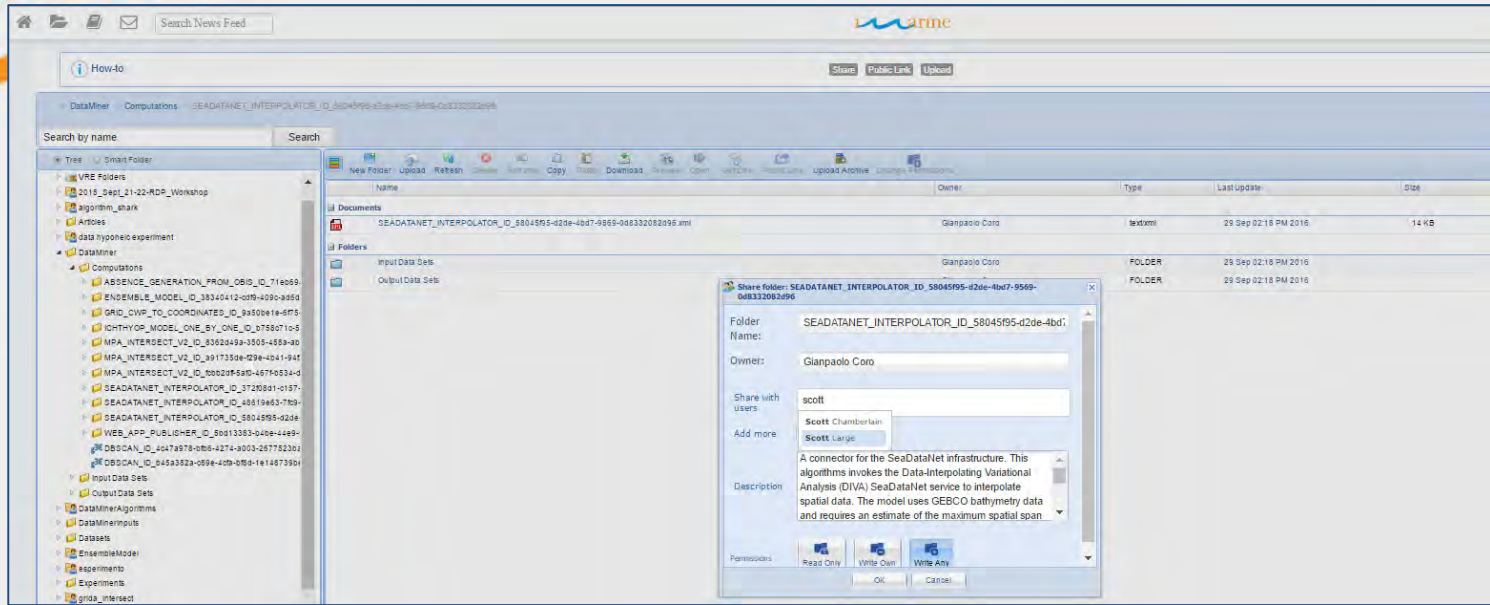
Computation Details

- Start Date: 04/10/2016 11:55:24
- End Date: 04/10/2016 11:55:47
- Status: completed
- VRE: /d4science.research-infrastructures.eu/gCubeApps/BiodiversityLab

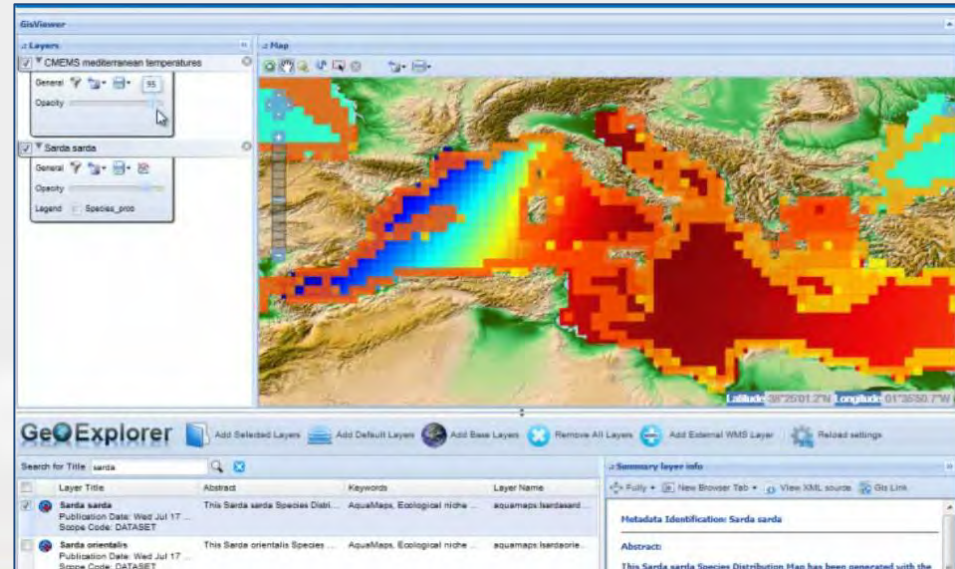
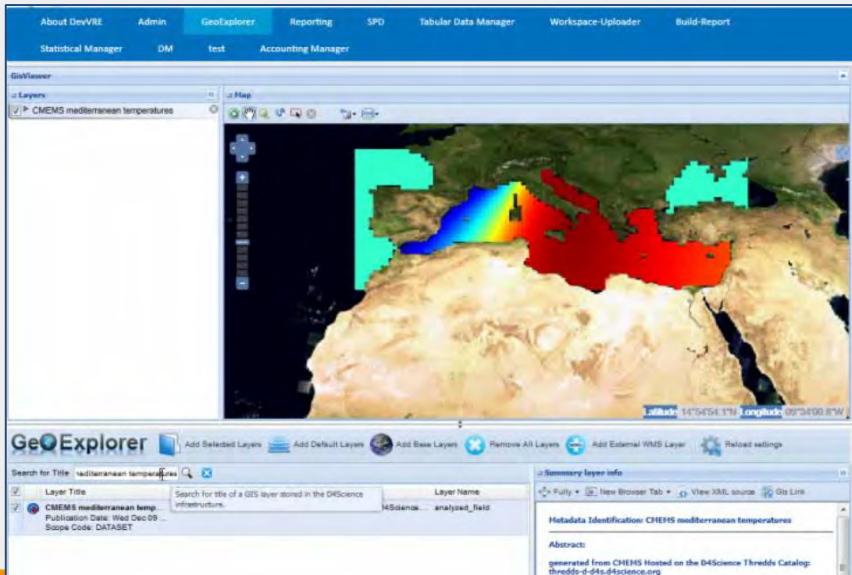
Operator Details

- Operator Name: SEADATANET_INTERPOLATOR
- Operator Description: A connector for the SeaDataNet infrastructure. This algorithm invokes the Data-Interp taken from the input table. As output, it produces a NetCDF file with a uniform grid of v

The PROV-O ontology is used to represent the metadata of the computation



STEP 7: publish the output as a map – overlay with other maps in the infrastructure





Benefits of the integration



- ✓ The process can be accessed through VREs



- ✓ Invoked via communication standards



- ✓ Available for the BlueBRIDGE community
(stock assessment, MPAs, aquaculture)



- ✓ Automatic Web interface for the process



- ✓ Provenance management



- ✓ Storage of results on a high-availability system



- ✓ Collaboration and sharing

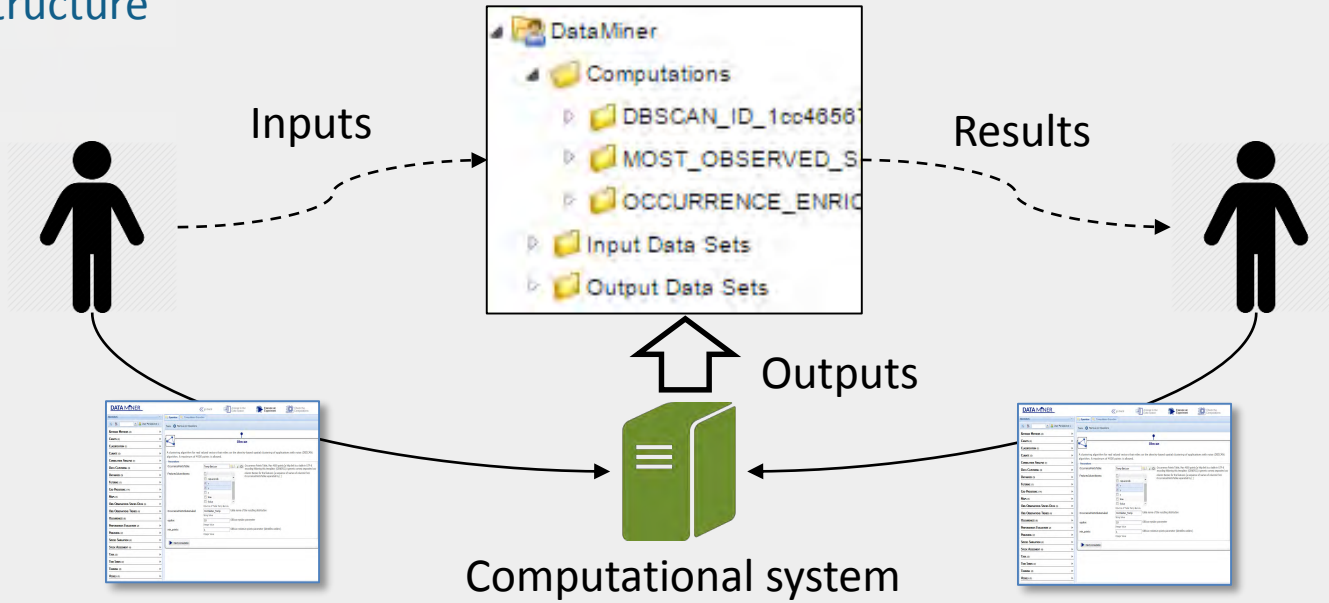


- ✓ R-R-R also from other software, e.g. QGIS

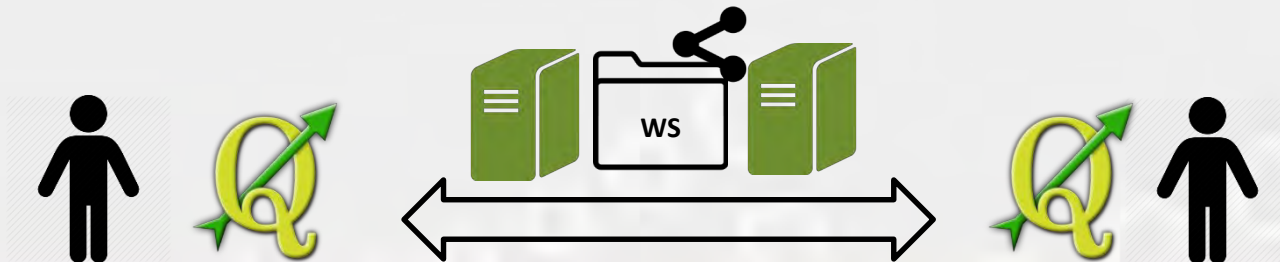
Collaborative experiments

Shared online folders

In the e-Infrastructure



Through third party software



QGIS 2.8.2-Wien

Project Edit View Layer Settings Plugins Vector Raster Database Web Processing Help

QgsWPSClient-2.0.16

Process Documentation

org.gaube.dataanalysis.wps.statisticalmanager.synchserver.mappedclasses.transducerers.ZEXTRACTION

ZEXTRACTION

[OutputTableLabel]
The name of the table to produce:

X coordinate:

Y coordinate:

[Resolution]
Step for Z values:

[Layer]
Layer Title or UUID or HTTP link. E.g. the title or the UUID (preferred) of a layer indexed in the e-Infrastructure on GeoNetwork - You can retrieve it from GeoExplorer. Otherwise you can supply the direct HTTP link of the layer. The format will be guessed from the link. The default is GeoTIFF. Supports several standards (NETCDF, WFS, WCS ASC, GeoTIFF):

[TimeIndex]
Time Index. The default is the first time indexed dataset:

Add Bookmark Back Run

QgsWPSClient-2.0.16

connect kill process

0%

QgsWPSClient-2.0.16 Service Chooser

Server Connections

DataMiner

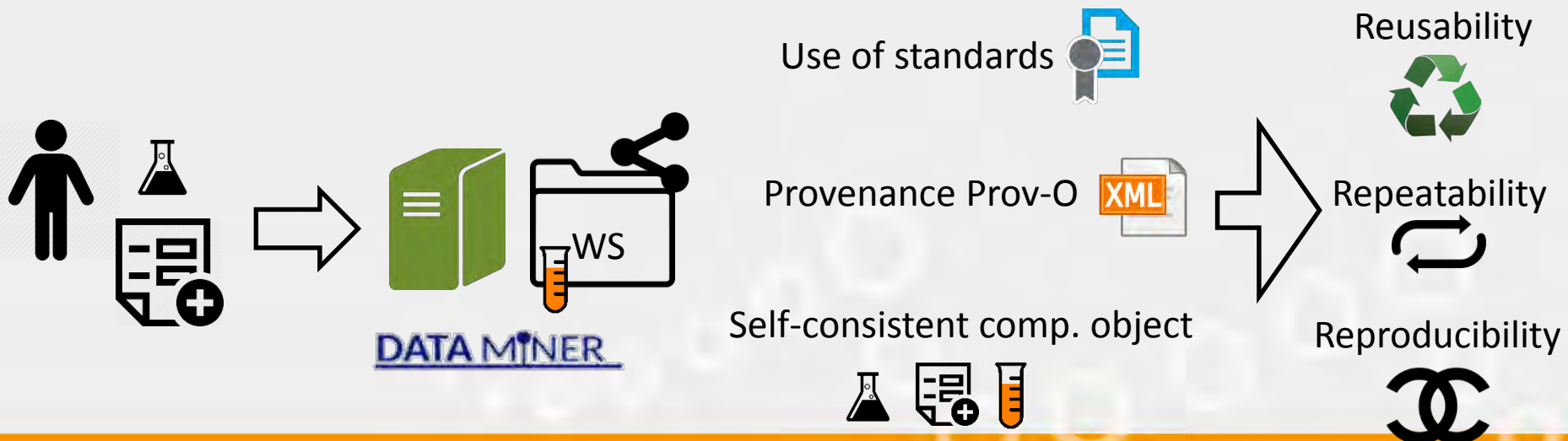
Connect New Edit Delete Bookmarks Add default server

Identifier	Title	Abstract
org.gaube.d...	ZEXTRACTION_TABLE	*
org.gaube.d...	ZEXTRACTION	*
org.gaube.d...	XYEXTRACTOR_TABLE	*
org.gaube.d...	XYEXTRACTOR	*
org.gaube.d...	TIMEEXTRACTION_TABLE	*
org.gaube.d...	TIMEEXTRACTION	*
org.gaube.d...	TIME_SERIES_CHARTS	*
org.gaube.d...	TIME_SERIES_ANALYSIS	*
org.gaube.d...	TIME_GEO_CHART	*
org.gaube.d...	SGVM_INTERPOLATION	*
org.gaube.d...	PRESENCE_CELLS_GENERATION	*
org.gaube.d...	OCCURRENCE_ENRICHMENT	*
org.gaube.d...	OBIS_TAXA_OBSERVATIONS_PER_YEAR	*
org.gaube.d...	OBIS_SPECIES_OBSERVATIONS_PER_YEAR	*
org.gaube.d...	OBIS_SPECIES_OBSERVATIONS_PER_MEOW_A...	*
org.gaube.d...	OBIS_SPECIES_OBSERVATIONS_PER_LINE_AREA	*
org.gaube.d...	OBIS_SINGLE_SPECIES_DISTRIBUTION_PER_A...	*
org.gaube.d...	OBIS_MOST_OBSERVED_TAXA	*
org.gaube.d...	OBIS_MOST_OBSERVED_SPECIES	*
org.gaube.d...	MAX_ENT_NICHE_MODELLING	*
org.gaube.d...	HCAF_FILTER	*
org.gaube.d...	GEO_CHART	*
org.gaube.d...	GENERIC_CHARTS	*
org.gaube.d...	FAO_OCEAN_AREA_COLUMN_CREATOR_FROM...	*
org.gaube.d...	FAO_OCEAN_AREA_COLUMN_CREATOR	*
org.gaube.d...	CSQUARE_COLUMN_CREATOR	*
org.gaube.d...	BIONYM_LOCAL	*
org.gaube.d...	ABSENCE_CELLS_FROM_AQUAMAPS	*
org.gaube.d...	ICCAT_VPA	*
org.gaube.d...	CMSY	*
org.gaube.d...	BIONYM	*
org.gaube.d...	QUALITY_ANALYSIS	*
org.gaube.d...	MSP...	*

about OK Close

Conclusions

- We have presented the benefits brought by a bridge between infrastructures;
- Our solution adds collaboration facilities and extends the usages of the connected service/data;
- This work goes in the direction of modern Science paradigms.



General links:

D4Science e-Infrastructure website: <https://www.d4science.org>

Portal of the D4Science infrastructure: <http://services.d4science.org>

Portal of the BlueBRIDGE infrastructure: <http://i-marine.d4science.org>

About DataMiner:

DataMiner user's guide: https://wiki.gcube-system.org/gcube/DataMiner_Manager

DataMiner installation guide: https://wiki.gcube-system.org/gcube/DataMiner_Installation

Contributing with new algorithms:

Tutorial on the integration of new algorithms with the DataMiner

http://gcube.wiki.gcube-system.org/gcube/index.php/How-to_Implement_Algorithms_for_the_Statistical_Manager

Related work:

- Candela, L., Castelli, D., Coro, G., Pagano, P., & Sinibaldi, F. (2013). Species distribution modeling in the cloud. *Concurrency and Computation: Practice and Experience*.
- Coro, G., Candela, L., Pagano, P., Italiano, A., & Liccardo, L. (2015). Parallelizing the execution of native data mining algorithms for computational biology. *Concurrency and Computation: Practice and Experience*, 27(17), 4630-4644.
- Candela, L., Castelli, D., Coro, G., Lelii, L., Mangiacrapa, F., Marioli, V., & Pagano, P. (2015). An infrastructure-oriented approach for supporting biodiversity research. *Ecological Informatics*, 26, 162-172.
- Coro, G., Magliozzi, C., Ellenbroek, A., Kaschner, K., & Pagano, P. (2016). Automatic classification of climate change effects on marine species distributions in 2050 using the AquaMaps model. *Environmental and Ecological Statistics*, 23(1), 155-180.
- Coro, G., Large, S., Magliozzi, C., & Pagano, P. (2016). Analysing and forecasting fisheries time series: purse seine in Indian Ocean as a case study. *ICES Journal of Marine Science: Journal du Conseil*, fsw131.
- Coro, G., Webb, T. J., Appeltans, W., Bailly, N., Cattrijsse, A., & Pagano, P. (2015). Classifying degrees of species commonness: North Sea fish as a case study. *Ecological Modelling*, 312, 272-280.



Thank you!