The Use of Digital Object Identifiers (DOIs)

The MedSEA CKP experience and other examples

Simona Simoncelli (INGV, Bologna)
simona.simoncelli@ingv.it

OUTLINE

- What is a DOI
- Motivations
- Effects
- Examples:

1. MedSea Ckeckpoint
2. SeaDataNet
3. CMEMS MED-MFC

- Some points of discussion

What is a DOI

- It is a persistent identifier or handle used to uniquely identify objects, standardized by the International Organization for Standardization (ISO)
- **Each DOI name permanently and unambiguously identifies the object to which it is associated**
- It also associates metadata with objects, allowing it to provide users with relevant pieces of information about the objects and their relationships
- Included as part of this metadata are network actions that allow DOI names to be resolved to web locations where the objects they describe can be found
- DOIs are in wide use mainly to identify academic, professional, and government information, such as journal articles, **research reports and data sets**, official publications and other types of information resources
- A DOI aims to be "resolvable", usually to some form of access to the information object to which the DOI refers. This is achieved by binding the DOI to metadata about the object, such as a URL, indicating where the object can be found.
- It is used to **locate, identify, and cite research data**

Motivations

Researchers and scientists are dealing with many activities and products like:

- Monitoring activities → sampling/technical developments
- Data management → formatting/QC/archiving/disseminating
- Project Management → reporting/disseminating/implementing services
- Software development → new applications
- Model development → operational oceanography, …
- Operational activities → management software, products development, products validation, reporting, product dissemination
- Service development → web applications

The use of DOI would give credit to all these activities and their creators

Effects

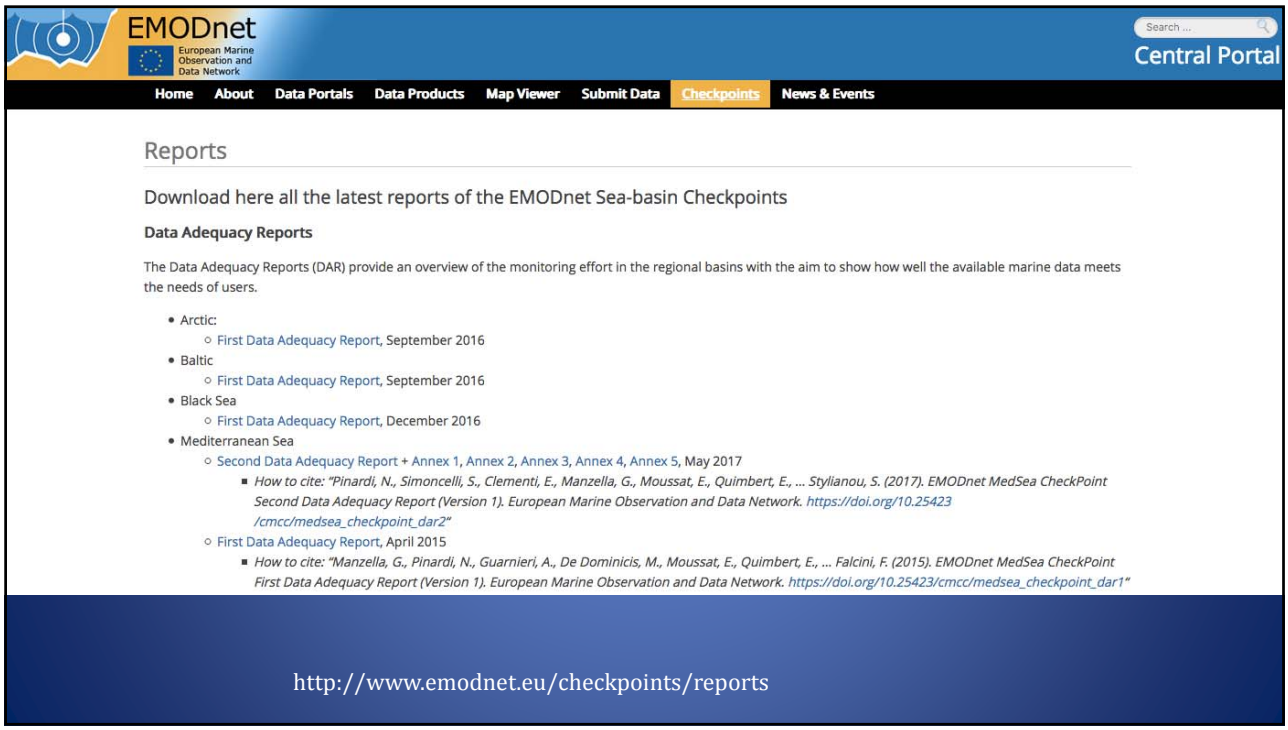
- Increase data sharing stimulating **data providers** to unlock earlier their data
- Increase data re-use, stimulating downstream applications (i.e. MPA connectivity)
- Encourage the creation/population of web catalogues (ISO)
- Stimulate/improve communication and dissemination activities
- Improve the overall quality of science ”products” through transparency (i.e. project reports, project results) and sharing
- Benefit the scientist career proving all his skills
- Increase science efficiency (cost/benefits) and reduce duplication of efforts
- Increase knowledge and pushes forward it
- Improve scientific papers production and transparency
- Provide statistics on the usage of the “object”→ its impact

The MedSea Checkpoint case

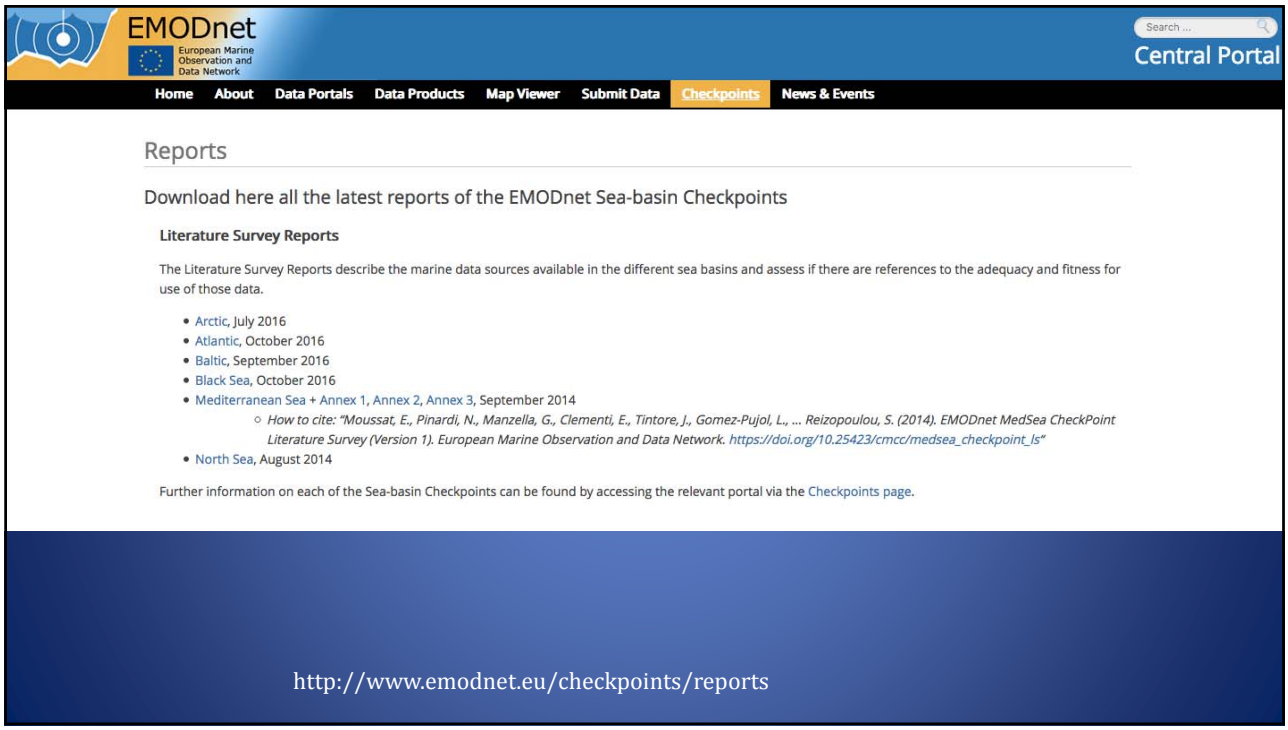
The main object is to **locate, identify, and cite project results**

DOI have been assigned to :

1. Main reports: DAR 1, DAR 2 and Literature Survey (by CMCC)
2. Targeted Products and the relative quality measures (by Ifremer)
3. Challenge reports describing the Targeted Products, their methodology and their expert evaluation (on going) (by CMCC)



http://www.emodnet.eu/checkpoints/reports



http://www.emodnet.eu/checkpoints/reports

Project Reports

http://www.emodnet-mediterranean.eu/reports_news/

Final Data Adequacy Report for Mediterranean checkpoint (with Annexes)

MEDSEA_D11.4 – Data Adequacy Report May 2017

How to cite: "Pinardi, N., Simoncelli, S., Clementi, E., Manzella, G., Moussat, E., Quimbert, E., ... Stylianou, S. (2017). EMODnet MedSea CheckPoint Second Data Adequacy Report (Version 1). European Marine Observation and Data Network. https://doi.org/10.25423/cmcc/medsea_checkpoint_dar2"

- Annex1_D11.4
- Annex2_D11.4
- Annex3_D11.4
- Annex4_D11.4
- Annex5_D11.4

First Data Adequacy Report for Mediterranean checkpoint

MEDSEA_D11.2 – Data Adequacy Report 18 June 2015

How to cite: "Manzella, G., Pinardi, N., Guarnieri, A., De Dominicis, M., Moussat, E., Quimbert, E., ... Falcini, F. (2015). EMODnet MedSea CheckPoint First Data Adequacy Report (Version 1). European Marine Observation and Data Network. https://doi.org/10.25423/cmcc/medsea_checkpoint_dar1"

Literature survey for Mediterranean checkpoint

The MedSea Literature Survey summarizes the first main attempt to classify the input data sets on the Mediterranean Sea by selecting a number of applications in order to determine how easy it is to access and use the data needed for these applications and their fitness for purpose.

MEDSEA_D1.3 – Literature Survey EMODNET MedSea Checkpoint 26 September 2014

How to cite: "Moussat, E., Pinardi, N., Manzella, G., Clementi, E., Tintore, J., Gomez-Pujol, L., ... Reizopoulou, S. (2014). EMODnet MedSea CheckPoint Literature Survey (Version 1). European Marine Observation and Data Network. https://doi.org/10.25423/cmcc/medsea_checkpoint_ls"

ANNEX1 – Methodology for classifying the existing upstream data according to the Literature'Survey 26 September 2014

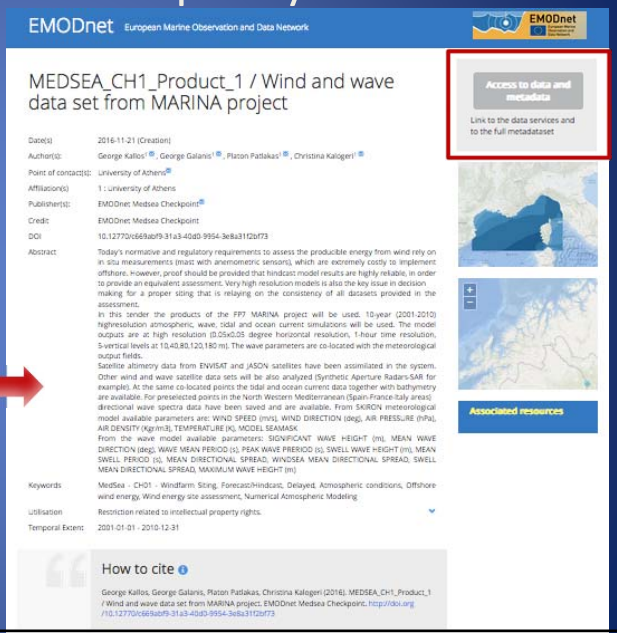
ANNEX2 – SeaDataNet characteristic categories and definitions (P02 list) 26 September 2014

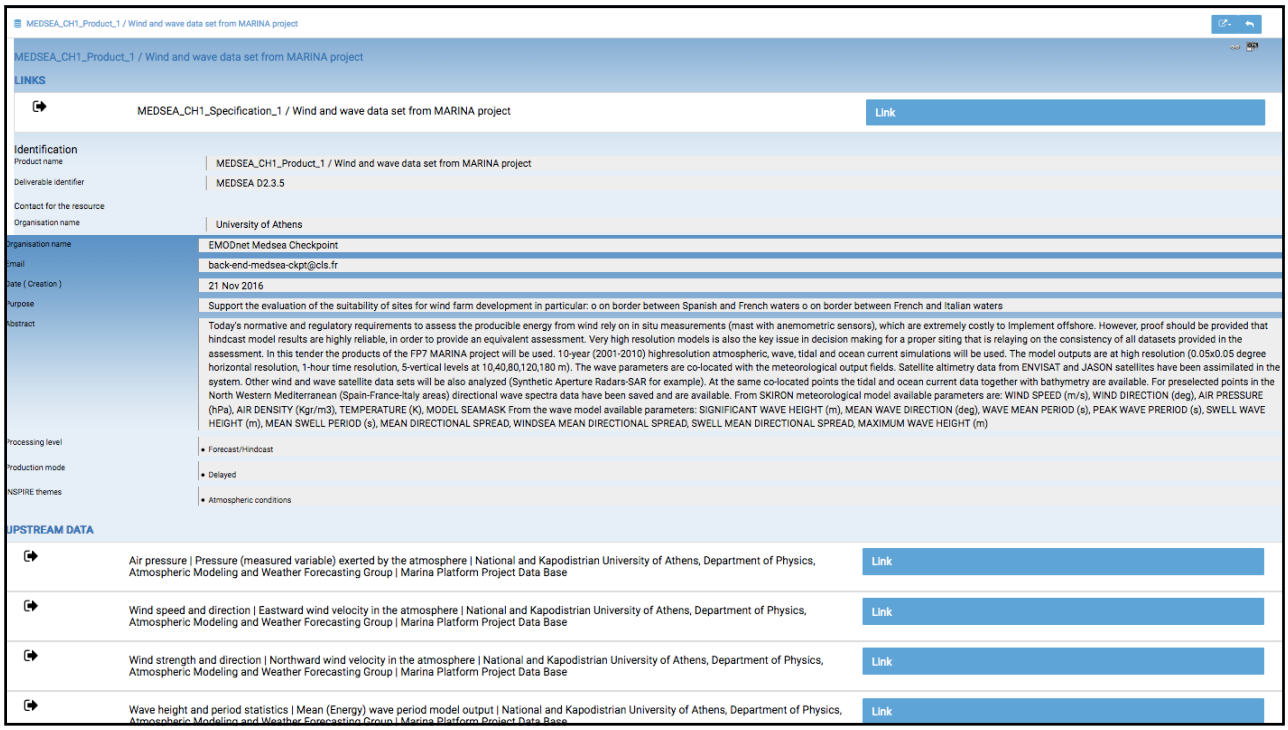
ANNEX3 – List of the upstream data sources and characteristics needed by the seven challenges of the 'MEDSEA Checkpoint project' using the SeaDataNet Common Vocabularies 26 September 2014

Targeted Products and the relative quality measures

Targeted Products

Name of Targeted product	Short description	Format
MEDSEA_CH1_Product_1	A wind – wave data set	shapefile
Full description and quality information: http://dx.doi.org/10.12770/c669abf9-31a3-40d0-9954-3e8a31f2bf73		
MEDSEA_CH1_Product_2	A suitability index of a wind farm in the NW Mediterranean concerning the environmental resources	shapefile
Full description and quality information: http://dx.doi.org/10.12770/225600af-6d9d-480e-b14f-bc26387d9bcd		
MEDSEA_CH1_Product_3	A suitability index of a wind farm in the NW Mediterranean concerning the environmental resources, the natural barriers, human activities, MPA and fisheries	shapefile
Full description and quality information: http://dx.doi.org/10.12770/d7b457aa-1339-482d-8b98-fedae1bc3ce2		





MEDSEA_CH1_Product_1 / Wind and wave data set from MARINA project

MEDSEA_CH1_Product_1 / Wind and wave data set from MARINA project

LINKS

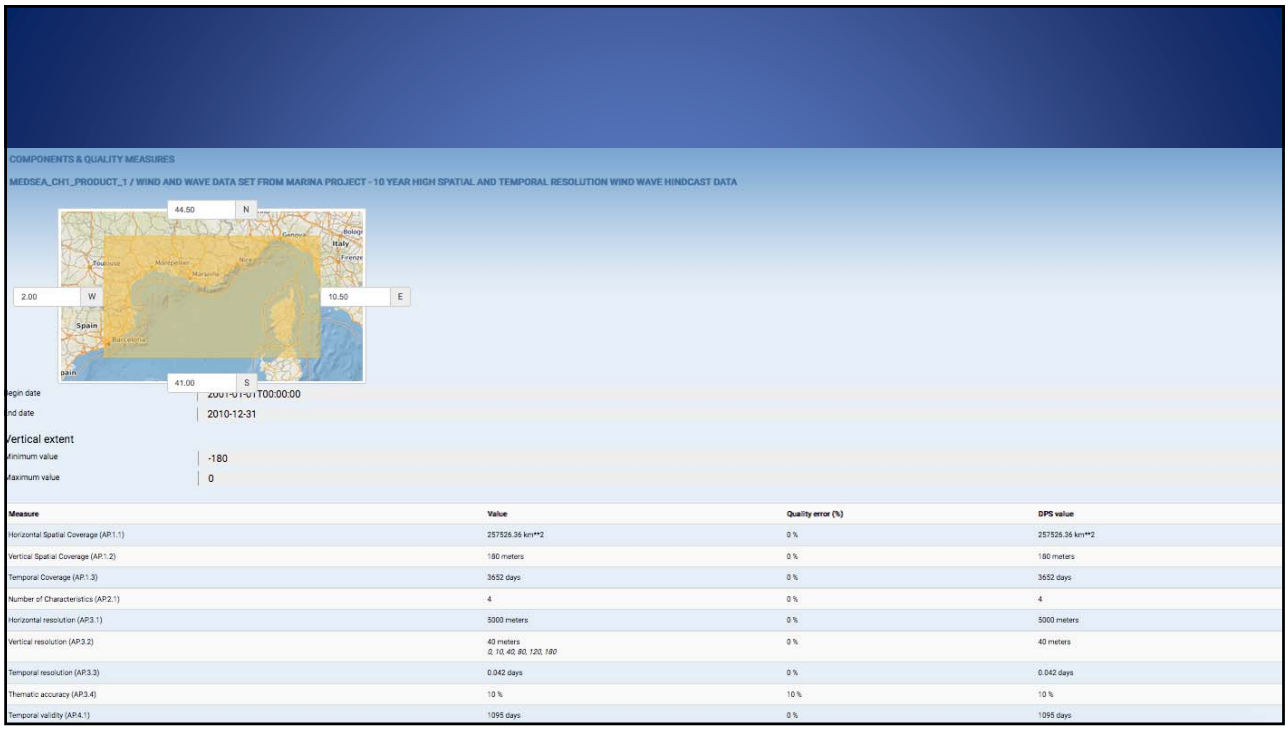
	MEDSEA_CH1_Specification_1 / Wind and wave data set from MARINA project	Link

Identification

Product name	MEDSEA_CH1_Product_1 / Wind and wave data set from MARINA project
Deliverable identifier	MEDSEA D2.3.5
Contact for the resource	
Organisation name	University of Athens
Organisation name	EMODnet Medsea Checkpoint
Email	back-end-medsea-ckpt@cls.fr
Date (Creation)	21 Nov 2016
Purpose	Support the evaluation of the suitability of sites for wind farm development in particular: o on border between Spanish and French waters o on border between French and Italian waters
Abstract	Today's normative and regulatory requirements to assess the producible energy from wind rely on in situ measurements (mast with anemometric sensors), which are extremely costly to Implement offshore. However, proof should be provided that hindcast model results are highly reliable, in order to provide an equivalent assessment. Very high resolution models is also the key issue in decision making for a proper siting that is relaying on the consistency of all datasets provided in the assessment. In this tender the products of the FP7 MARINA project will be used. 10-year (2001-2010) highresolution atmospheric, wave, tidal and ocean current simulations will be used. The model outputs are at high resolution (0.05x0.05 degree horizontal resolution, 1-hour time resolution, 5-vertical levels at 10,40,80,120,180 m). The wave parameters are co-located with the meteorological output fields. Satellite altimetry data from ENVISAT and JASON satellites have been assimilated in the system. Other wind and wave satellite data sets will be also analyzed (Synthetic Aperture Radars-SAR for example). At the same co-located points the tidal and ocean current data together with bathymetry are available. For preselected points in the North Western Mediterranean (Spain-France-Italy areas) directional wave spectra data have been saved and are available. From SKIRON meteorological model available parameters are: WIND SPEED (m/s), WIND DIRECTION (deg), AIR PRESSURE (hPa), AIR DENSITY (Kgr/m3), TEMPERATURE (K), MODEL SEAMASK From the wave model available parameters: SIGNIFICANT WAVE HEIGHT (m), MEAN WAVE DIRECTION (deg), WAVE MEAN PERIOD (s), PEAK WAVE PRERIOD (s), SWELL WAVE HEIGHT (m), MEAN SWELL PERIOD (s), MEAN DIRECTIONAL SPREAD, WINDSEA MEAN DIRECTIONAL SPREAD, SWELL MEAN DIRECTIONAL SPREAD, MAXIMUM WAVE HEIGHT (m)
Processing level	• Forecast/Hindcast
Production mode	• Delayed
INSPIRE themes	• Atmospheric conditions

UPSTREAM DATA

	Air pressure | Pressure (measured variable) exerted by the atmosphere | National and Kapodistrian University of Athens, Department of Physics, Atmospheric Modeling and Weather Forecasting Group | Marina Platform Project Data Base	Link
	Wind speed and direction | Eastward wind velocity in the atmosphere | National and Kapodistrian University of Athens, Department of Physics, Atmospheric Modeling and Weather Forecasting Group | Marina Platform Project Data Base	Link
	Wind strength and direction | Northward wind velocity in the atmosphere | National and Kapodistrian University of Athens, Department of Physics, Atmospheric Modeling and Weather Forecasting Group | Marina Platform Project Data Base	Link
	Wave height and period statistics | Mean (Energy) wave period model output | National and Kapodistrian University of Athens, Department of Physics, Atmospheric Modeling and Weather Forecasting Group | Marina Platform Project Data Base	Link



COMPONENTS & QUALITY MEASURES

MEDSEA_CH1_PRODUCT_1 / WIND AND WAVE DATA SET FROM MARINA PROJECT - 10 YEAR HIGH SPATIAL AND TEMPORAL RESOLUTION WIND WAVE HINDCAST DATA

Begin date 2001-01-01T00:00:00

End date 2010-12-31

Vertical extent

Minimum value -180

Maximum value 0

Measure	Value	Quality error (%)	DPS value
Horizontal Spatial Coverage (AP.1.1)	257526.36 km**2	0 %	257526.36 km**2
Vertical Spatial Coverage (AP.1.2)	180 meters	0 %	180 meters
Temporal Coverage (AP.1.3)	3652 days	0 %	3652 days
Number of Characteristics (AP.2.1)	4	0 %	4
Horizontal resolution (AP.3.1)	5000 meters	0 %	5000 meters
Vertical resolution (AP.3.2)	40 meters *0, 10, 40, 80, 120, 180*	0 %	40 meters
Temporal resolution (AP.3.3)	0.042 days	0 %	0.042 days
Thematic accuracy (AP.3.4)	10 %	10 %	10 %
Temporal validity (AP.4.1)	1095 days	0 %	1095 days

COMPONENTS & QUALITY MEASURES

MEDSEA_CH1_PRODUCT_1 / WIND AND WAVE DATA SET FROM MARINA PROJECT - 10 YEAR HIGH SPATIAL AND TEMPORAL RESOLUTION WIND WAVE HINDCAST DATA

44.50 N | 2.00 W | 10.50 E | 41.00 S

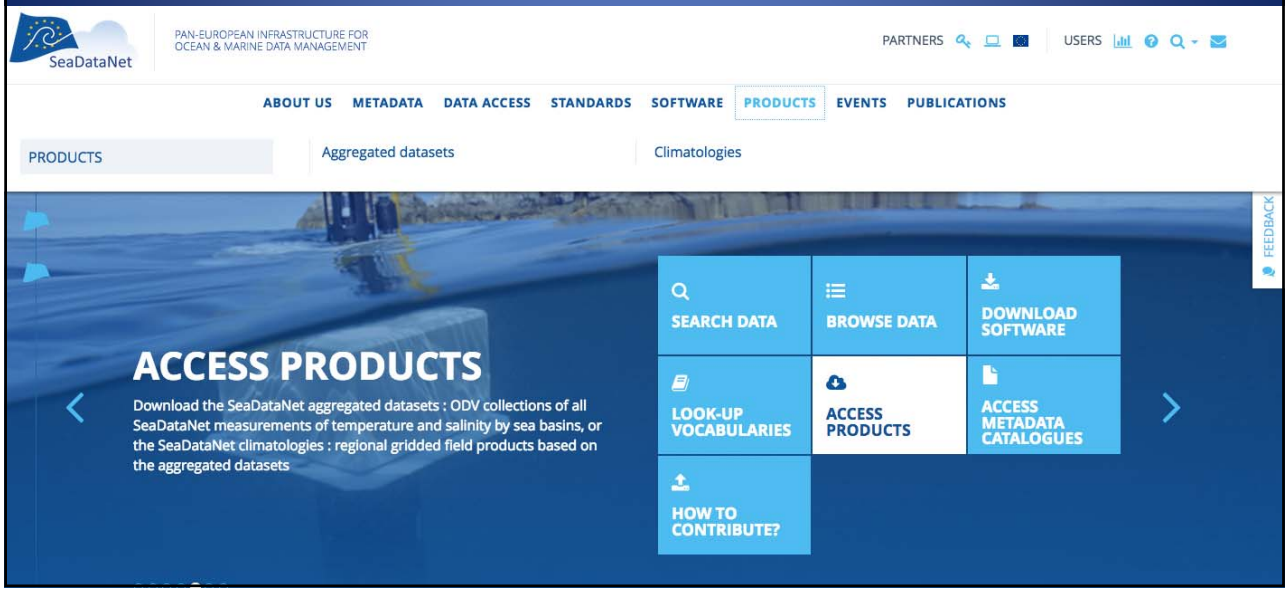
egin date	2001-01-01T00:00:00
nd date	2010-12-31

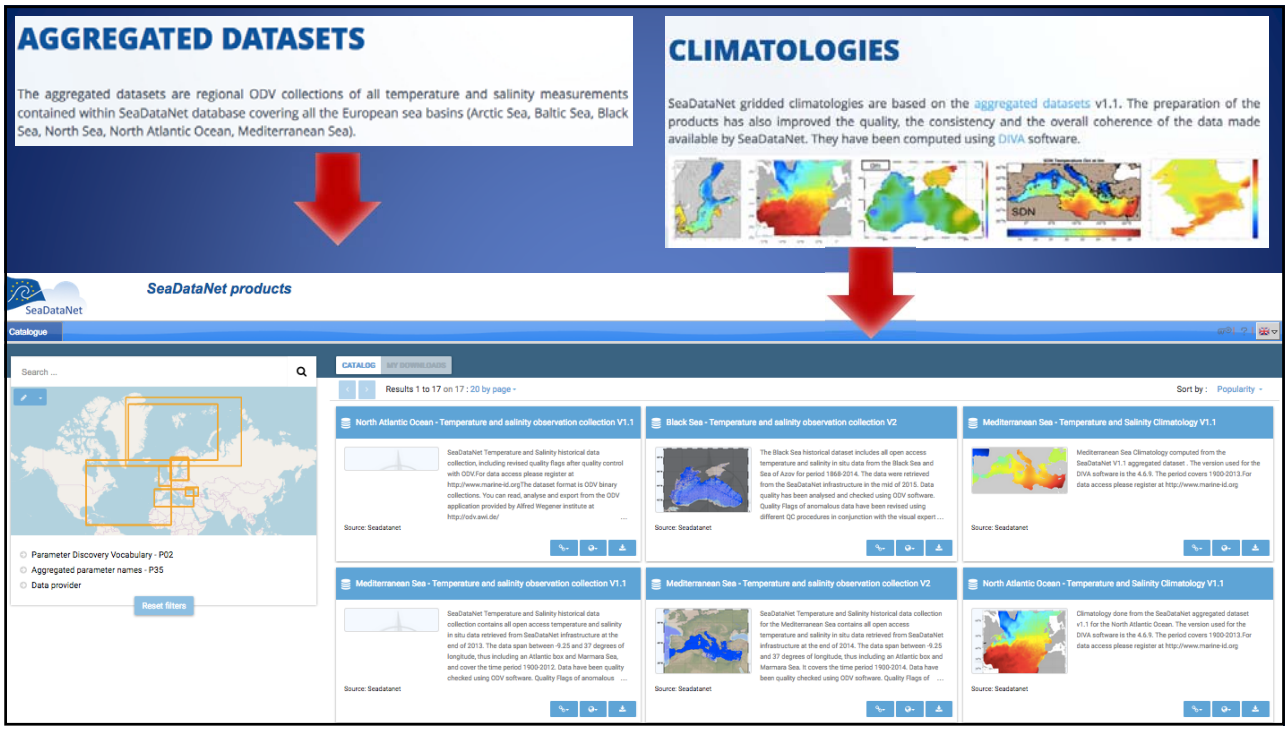
Vertical extent

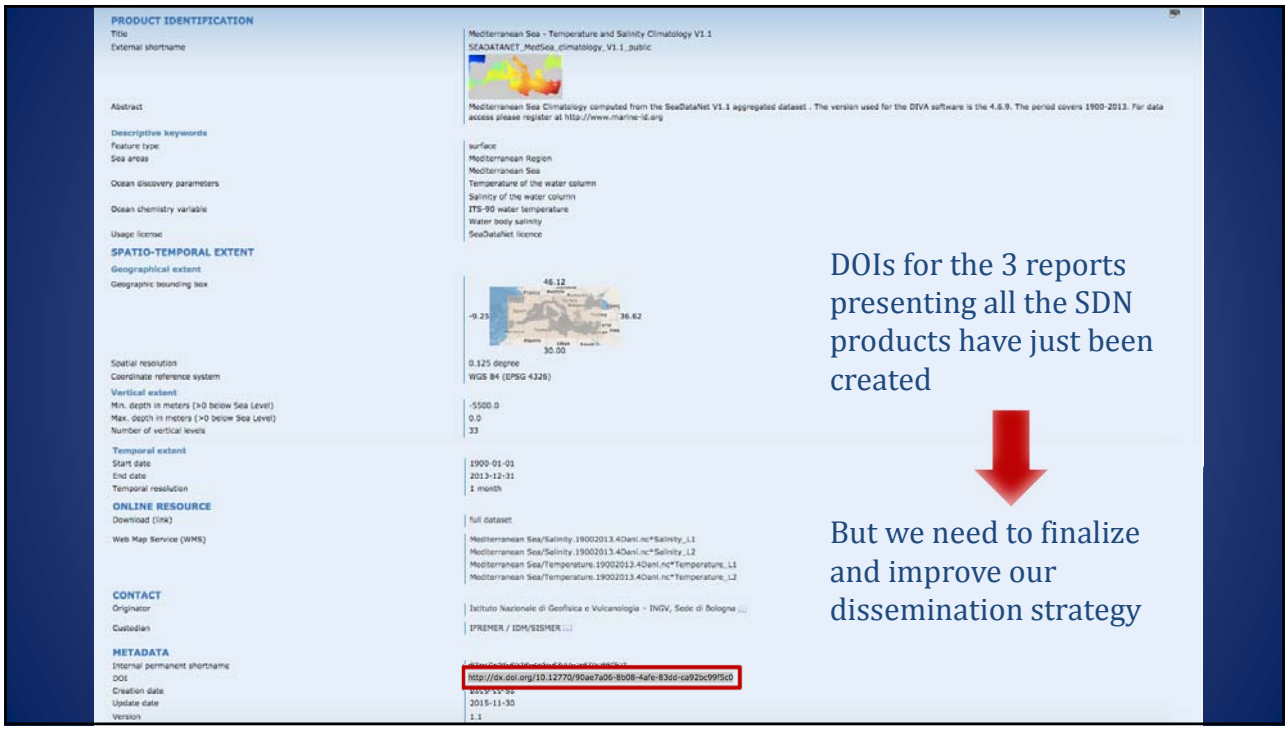
Minimum value	-180
Maximum value	0

Measure	Value	Quality error (%)	DPS value
Horizontal Spatial Coverage (AP1.1)	257526.36 km**2	0 %	257526.36 km**2
Vertical Spatial Coverage (AP1.2)	180 meters	0 %	180 meters
Temporal Coverage (AP1.3)	3652 days	0 %	3652 days
Number of Characteristics (AP2.1)	4	0 %	4
Horizontal resolution (AP3.1)	5000 meters	0 %	5000 meters
Vertical resolution (AP3.2)	40 meters 0, 10, 40, 80, 120, 180	0 %	40 meters
Temporal resolution (AP3.3)	0.042 days	0 %	0.042 days
Thematic accuracy (AP3.4)	10 %	10 %	10 %
Temporal validity (AP4.1)	1095 days	0 %	1095 days

SeaDataNet Example of data products



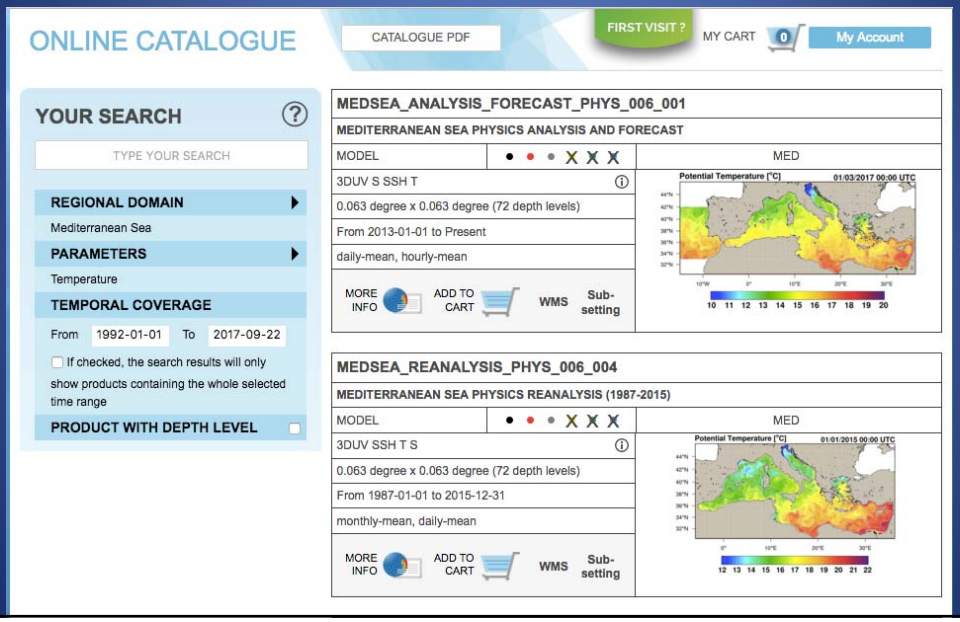




DOIs for the 3 reports presenting all the SDN products have just been created

But we need to finalize and improve our dissemination strategy

MED-MFC products



Conclusions

- We are still learning how to manage information in order to communicate it simply
- The researchers are happy to get acknowledgment of their work
- DOI allows to identify datasets uniquely
- DOI should increase visibility and trust
- The users find all the information related to the resource
- Researchers get credits from their citation especially on their data usage
- **Scientific Data journals** allow also to publish datasets, i.e. Earth System Science Data (IF 6.696) and Nature Scientific Data
- It should change the way to write papers→ 1) more focus on the main subject allowing to cite additional resources (i.e. project reports); 2) leaving the data analysis and validation to data journals