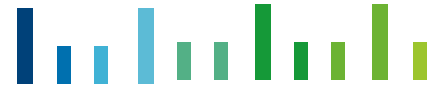


# Joint European Research Infrastructure network for Coastal Observatories



## Real Time Data Management Handbook V1 - D5.2

Grant Agreement n° 262584

Project Acronym: JERICO

Project Title: Towards a Joint European Research Infrastructure network for Coastal Observatories

Coordination: P. Farcy, IFREMER

[jerico@ifremer.fr](mailto:jerico@ifremer.fr), [www.jerico-fp7.eu](http://www.jerico-fp7.eu)

Authors: C Fanara, S Pouliquen, L Petit De La Villeon, G Maudire and A Crise

Involved Institutions: OGS, IFREMER

Version and Date: final, 2012.05.31



# Table of Contents



<b>TABLE OF CONTENTS</b>	<b>3</b>
<b>1. DOCUMENT DESCRIPTION</b>	<b>5</b>
<b>2. SUMMARY AND INTRODUCTION</b>	<b>7</b>
<b>3. MAIN REPORT</b>	<b>9</b>
<b>3.1. Common Standards</b>	<b>9</b>
3.1.1. Common vocabularies for JERICO delayed-mode and real time data formats and meta-databases	9
3.1.2. Common data formats	11
3.1.3. Standard Tools for data access and visualization	14
<b>3.2. JERICO Real Time Data Management Procedures</b>	<b>15</b>
3.2.1. Procedure to provide RT data to JERICO	17
3.2.2. Description of the Common data formats and Real-Time Data portal architecture	19
3.2.3. Data quality control procedure for real-time data	21
3.2.4. Standard Tools for data access and visualization	21
<b>4. CONCLUSIONS</b>	<b>22</b>



# 1. Document description



## REFERENCES

Annex 1 to the Contract: Description of Work (DoW) version 2011-02-22

Document information	
Document Name	RT Data Management Handbook V1
Document ID	<b><i>JERICO-WP5-D5.2-120531-V1</i></b>
Revision	Final
Revision Date	2012.05.31
Author	C Fanara, S Pouliquen, L Petit De La Villeon, G Maudire and A Crise
Security	

History			
Revision	Date	Modification	Author
1.0	2011.12.21	Creation of the document	L Petit De La Villeon, S Pouliquen, G Maudire
2.0	2012.01.02	Integrations	C Fanara, A Crise
3.0	2012.03.28	Integrations	C Fanara, A Crise
4.0	2012.04.02	Integrations	C Fanara, A Crise
5.0	2012.04.30	Integrations	L Petit De La Villeon, S Pouliquen, G Maudire
6.0	2012.05.11	Revision	C Fanara, A Crise

Diffusion list				
Consortium beneficiaries	X			
Third parties				
Associated Partners				



This document contains information, which is proprietary to the JERICO consortium. Neither this document nor the information contained herein shall be used, duplicated or communicated by any means to any third party, in whole or in parts, except with prior written consent of the JERICO Coordinator.

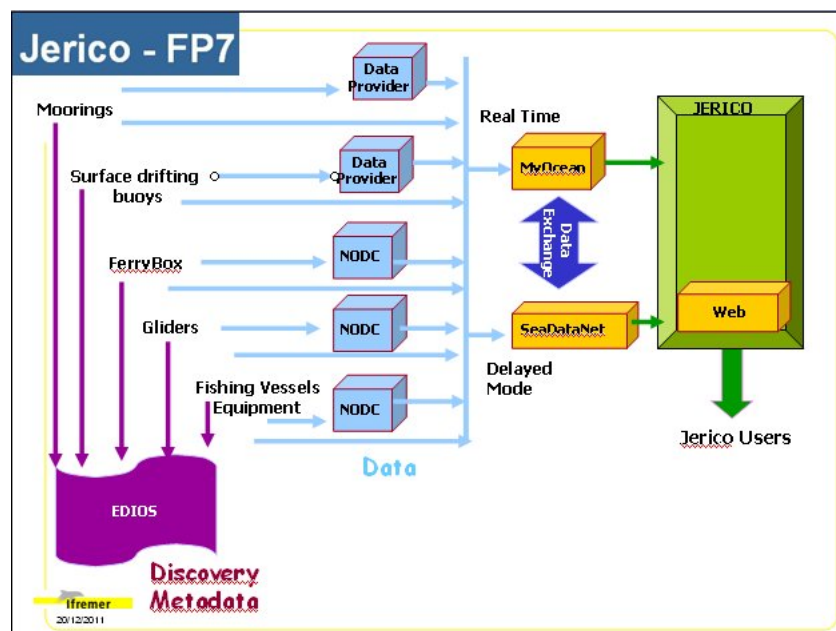
The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

## 2. Summary and Introduction

The success of the JERICO project mainly depends on the goal of assuring that the flow of real-time and delayed mode data coming from the participating observing networks will be reliable, accessible and easy to distribute.

To get this purpose it is necessary the use of common procedures for real time and delayed mode data management within the JERICO community.

This Handbook aims to be the key tool for harmonizing real time and delayed mode data management for in situ data collection and for providing practical advice to JERICO community in data delivery.



General data flow within JERICO.



The structure of the document derives from current experiences largely (but not exclusively) coming from the major European initiatives for the establishment and coordination of infrastructures for the management and distribution of data and products in Europe (SeaDataNet, MyOCEAN and EuroGOOS Regional Centres).

The first part of the document is dedicated to the description of the common standards (vocabularies, data formats and standard tools) for delayed-mode and real time data.

In the second part the procedures for both type of data are shown.

The Handbook contains all the specific references and links to the basic and most important documents (available on WWW sites) that should help each Data Provider to successfully manage JERICO data from acquisition to dissemination.

The paper, here proposed in his first version, can in future be amended in light of specific needs and extended when required.



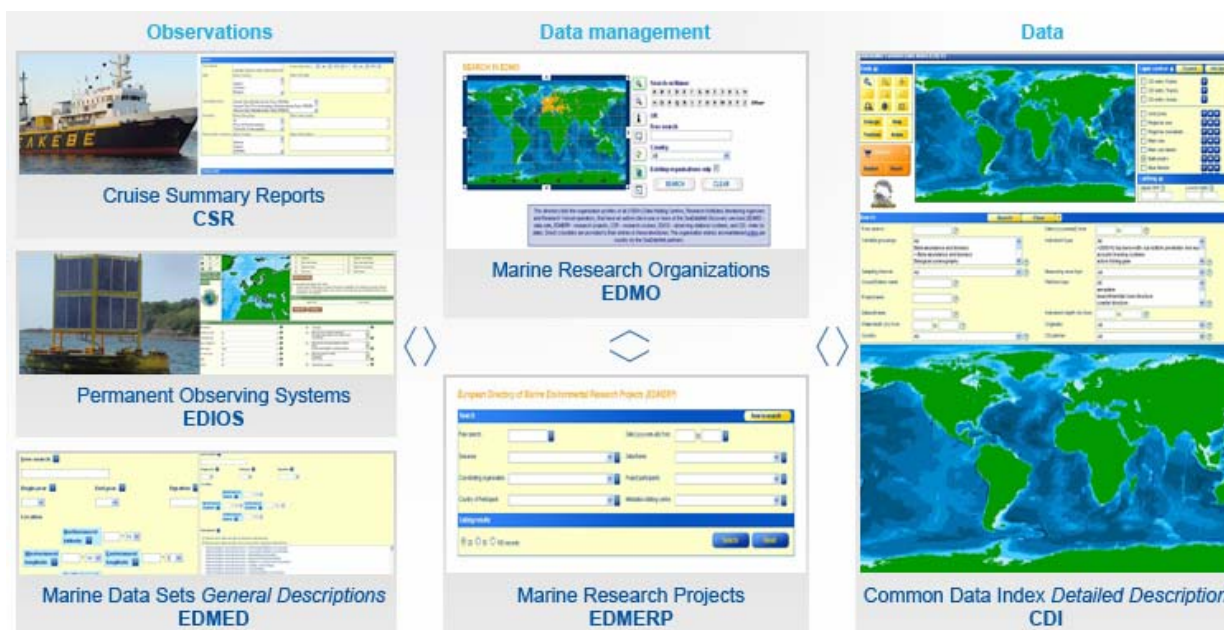
# 3. Main Report



## 3.1. Common Standards

### 3.1.1. Common vocabularies for JERICO delayed-mode and real time data formats and meta-databases

The Metadata Management for JERICO will rely on the metadata directories at European level developed in the framework of SeaDataNet EU-FP7 infrastructure project (<http://www.seadatanet.org>) .



These metadata directories at European level will be used to describe:

- Marine Research projects (EDMERP directory) i.e. JERICO and its components, description done by JERICO coordination team;
- Marine Research Organizations (EDMO directory) involved in JERICO project;



- Cruise Summary Reports for cruises which have been conducted for JERICO deployments (if relevant);
- Permanent Observing Systems (EDIOS directory) to describe JERICO platforms.

**Populating EDIOS directory with the description of JERICO observing system will be mandatory for fixed stations and ferryboxes.**

For gliders it may be possible to populate EDIOS as well.

It will be done using the Mikado tool, an Open Access software developed by SeaDataNet project which is freely downloadable at <http://www.seadatanet.org/Standards-Software/Software>. Mikado user's manual is also available on the same web page. Please, refer to it for further details on managing metadata (information to be typed or extracted from existing meta databases, mandatory and optional fields, ...) and software usage. Helping support may be requested to [sdnuserdesk@seadatanet.org](mailto:sdnuserdesk@seadatanet.org). Mikado can be used in two modes:

Manual typing of information when metadata have not been already digitized;

Automatic mode which retrieves, after appropriate configuration, the metadata from any relational database (i.e. supporting SQL language) or Spreadsheet file (e.g. MS/Excel or extracting information from the ROOSes/MyOcean portal platform index).

After generation, metadata must be sent to SeaDataNet according to guidelines provided at the "[Metadata: How to contribute?](http://www.seadatanet.org/Metadata/How-to-contribute)" link (<http://www.seadatanet.org/Metadata/How-to-contribute>).



### 3.1.2. Common data formats

JERICO data will be distributed using the EUROGOOS recommended formats endorsed by EuroGOOS institutes at the 2010 annual meeting in Sopot. These recommended formats are either ODV (<http://odv.awi.de>) or OceanSites NetCDF (<http://www.coriolis.eu.org/Data-Services-Products/MyOcean-In-Situ-TAC/Documentation>).

Common data formats adopted in JERICO are described in line with data formats adopted within SeaDataNet and MyOcean projects. Data may be provided using alternatively:

- Ocean Data View spread sheet format (ASCII flat file, comma separated values, with a standardized header including parameters and units referring to P011 Common Vocabulary) see <http://odv.awi.de>;
- NetCDF format using CF convention. OceanSites NetCDF used by MyOcean *in-situ* TAC developed jointly with EuroGOOS ROOSes (<http://www.coriolis.eu.org/Data-Services-Products/MyOcean-In-Situ-TAC/Documentation>)
- Optionally, Medatlas format, if already in use at provider's data centre (deprecated).

Description of these common formats may be found at:

<http://www.seadatanet.org/Standards-Software/Data-Transport-Formats>



**Note:** The Nemo software tool has been developed by SeaDataNet project to facilitate formatting of data from ASCII files (e.g. instrument vendor formats) to one of the adopted ASCII common formats: Ocean Data View and Medatlas. This software may also optionally extract metadata information from source files in order to facilitate provision of standardized CDI (ISO 19115 metadata).

The Nemo software is freely downloadable from SeaDataNet web site at:

<http://www.seadatanet.org/Standards-Software/Software>

The NetCDF format is used as an exchange format for Near Real time data distribution especially by the *in-situ* TAC partners and data providers. There is no need of transformation tool. The NetCDF format is planned to be one of the SeaDataNet adopted format.

***Important notice:***

*Since CF convention is restricted to climate change purposes, extensions to this convention may be useful for several type of data (marine chemistry, contaminants and marine biology). These extensions are currently being studied within SeaDataNet project and will be released as soon as possible.*

For QC procedures see 3.3 and 4.3.

For the QC flag scale see the following table.



Code	Meaning	Real-time comment	Post-recovery mode comment
0	No QC was performed	-	-
1	Good data	All real-time QC tests passed	All post-recovery tests passed.
2	Probably good data	-	-
3	Bad data that are potentially correctable These data are not to be used without scientific correction Possible re-calibration needed.	-	-
4	Bad data	Data have failed one or more tests.	Data have been examined and they are too bad to be corrected
5	Value changed .	Data may be recovered after transmission error	-
6	Not used	-	-
7	Nominal value	Data were not observed but reported; i.e. an anchor drop position or an instrument target depth, as opposed to a GPS position time series or a pressure sensor time series.	-
8	Interpolated value	Missing data may be interpolated from neighbouring data in space or time	-
9	Missing value	-	-



### **3.1.3. Standard Tools for data access and visualization**

The Ocean Data View (ODV) software package is a standard tool for data analysis and visualization. Freely available on the SDN portal

<http://www.seadatanet.org/Standards-Software/Software/ODV>,

ODV works with oceanographic and other geo-referenced profiles or sequences of data. It is ideal for visual and automatic quality control, since it is plenty of interactive capabilities and supports a wide range of plot types.



### 3.2. JERICO Real Time Data Management Procedures

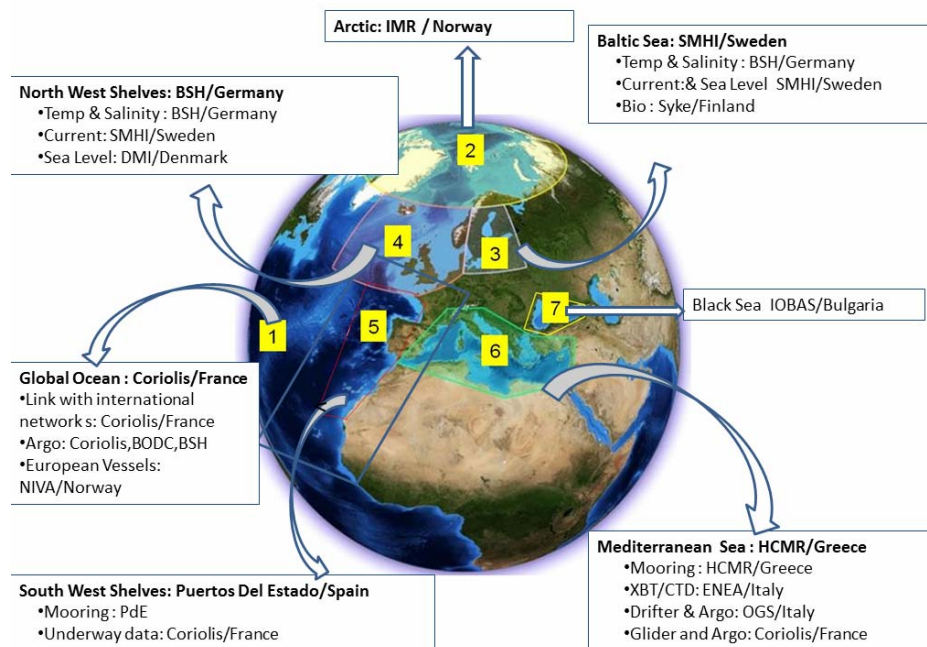
The JERICO near real-time data flow will be based on what has been developed jointly by MyOcean and EuroGOOS/ROOSes.

The *in-situ* Thematic Assembly Centre (*in-situ* TAC) of MyOcean is a distributed service integrating data from different sources for operational oceanography needs. The MyOcean *in-situ* TAC is collecting and carrying out quality control in a homogeneous manner on data from outside MyOcean data providers to fit the needs of internal and external users. It provides access to integrated datasets of core parameters for initialization, forcing, assimilation and validation of ocean numerical models which are used for forecasting, analyses (nowcast) and re-analysis (hindcast) of ocean conditions. Since the primary objective of MyOcean is to forecast ocean state, the initial focus is on observations from automatic observatories at sea (e.g. floats, buoys, gliders, ferrybox, drifters, SOP) which are transmitted in real-time to the shore.

As the structure of the FTP MyOcean portal is based on platforms, it was then feasible to integrate all the parameters they measure and also to include platforms that measure parameters not processed by MyOcean *in-situ* TAC such as wind and waves. This activity has been endorsed by the EuroGOOS ROOSes in collaboration with the regional *in-situ* TAC partners and no additional near real time QC is performed on the new parameters. In collaboration with SeaDataNet, the appropriate vocabularies have been used to integrate the new parameters and SeaDataNet will extend these vocabularies if needed.



The *in-situ* TAC comprises a global *in-situ* centre and 6 regional *in-situ* centres. The *in-situ* TAC has been designed to fulfil the GMES Marine Core Service needs and the EuroGOOS regional systems (ROOS) needs. The focus of the MyOcean *in-situ* TAC is on parameters that are presently necessary for GMES Monitoring and Forecasting Centres namely temperature, salinity, sea level, current, chlorophyll / fluorescence, oxygen and nutrients. Additional parameters such as wind and waves are added by some ROOSes to these regional *in-situ* portals to fulfil additional downstream applications needs.







### **3.2.1. Procedure to provide RT data to JERICO**

JERICO data providers can provide real time data to JERICO data management system through MyOcean using the Global Centre (Coriolis) or the 6 regional centres.

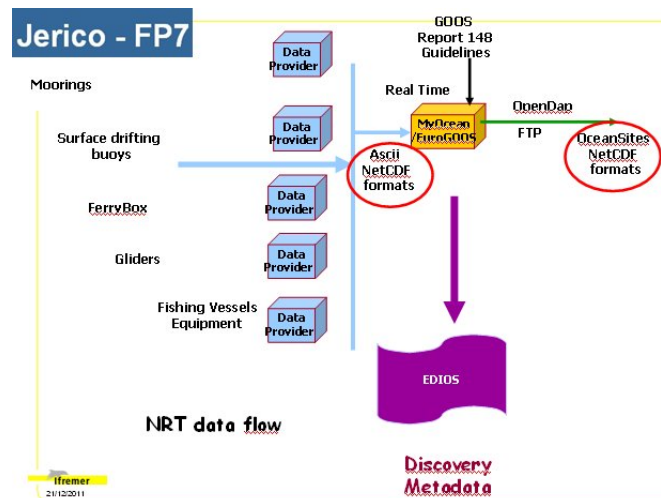
A new data provider has to contact the ROOS portal manager and provide access to his data via an FTP server without changing his in house format as long as it contains **enough metadata information to generate the NetCDF Files: What (platform name, institution , WMO number when available, type of platform, ...), Where (space coverage), When(time coverage), Who (Data provider, contact, Pi, Data Center, ...), How (update interval, QC info).**

The conversion to common NetCDF format is taken on board by the MyOcean *in-situ* TAC partners as well as data integration on the portal.

The ROOS portal managers are described in the following table. They will either handle the management of the JERICO data in their institute or inform the *in-situ* TAC partners to perform it. As defined in the framework of the MyOcean *in-situ* TAC, a unique contact point is defined for each region that organizes the activity and establishes the link with the internal partners.

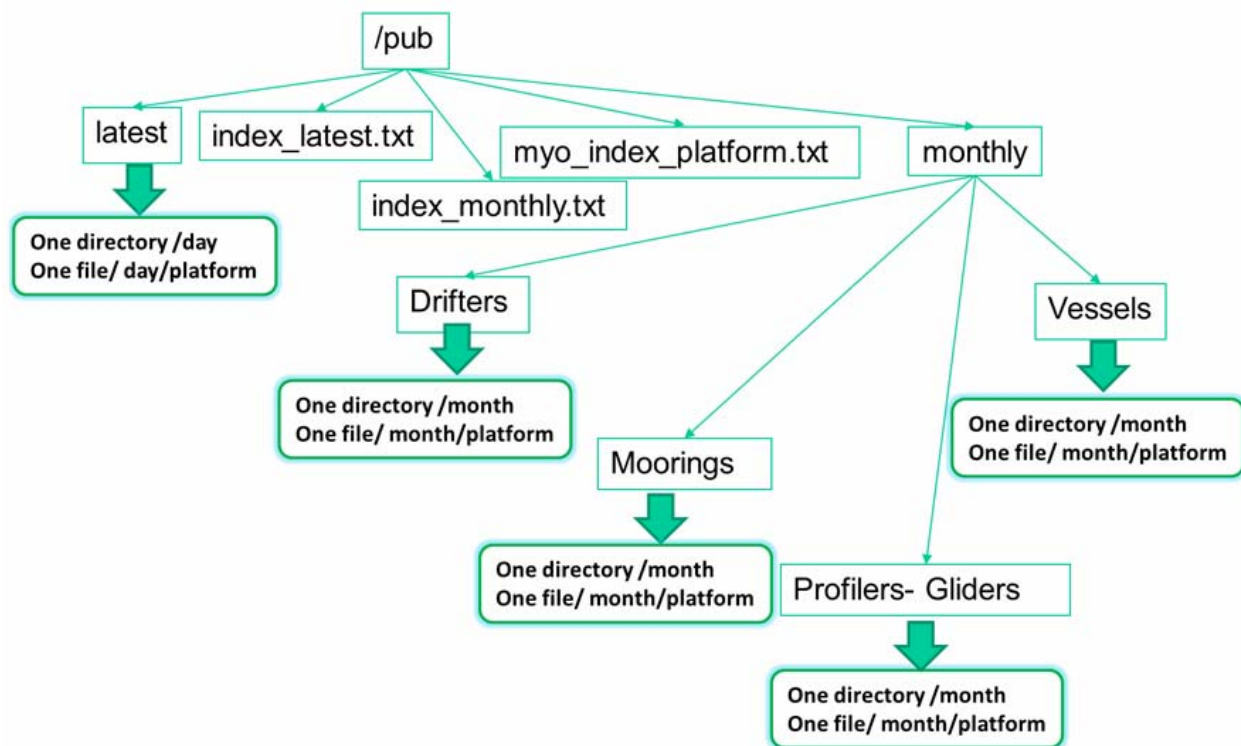


Contact	Email	Region
Coriolis data Center Loic Petit de la Villeon Thierry Carval	codac@ifremer.fr loic.petit.de.la.villeon@ifremer.fr thierry.carval@ifremer.fr	Global and European
Pierre Jaccard	Pierre.Jaccard@niva.no	EU Ferrybox
Leonidas Perivoliotis Dimitris Kassis	lperiv@ath.hcmr.gr dkassis@ath.hcmr.gr	Mediterranean Sea
Henning Wehde Sjur Ringheim Lid	Henning.Wehde@imr.no Sjur.Ringheim.Lid@imr.no	Arctic
Kai Soetje Tobias Gies	kai.soetje@bsh.de tobias.gies@bsh.de	NOOS
Veselka Marinova Asen Stefanov	marinova@io-bas.bg a.stefanov@io-bas.bg	Black Sea
Marta de Alfonso	mar@puertos.es	IBI-ROOS
Thomas Hammarklint Anders Hartman	thomas.hammarklint@smhi.se anders.hartman@smhi.se	BOOS



### 3.2.2. Description of the Common data formats and Real-Time Data portal architecture

There is no requirement on data provider to transform their data into OceanSites Netcdf format as this activity is done by the MyOcean *in-situ* TAC partners. The data will be distributed on the ROOSes portals in OceanSites NetCDF format according to the FTP organisation described in the MyOcean *in-situ* TAC documentation available at <http://www.coriolis.eu.org/Data-Services-Products/MyOcean-In-Situ-TAC/Documentation> and shown in the figure below.



The list of the parameter codes used for Near real-time data circulation is described on <http://www.coriolis.eu.org/Data-Services-Products/MyOcean-In-Situ-TAC/MyOcean-In-Situ-parameter-codes>.



### **3.2.3. Data quality control procedure for real-time data**

The Near Real Time data quality procedures for Temperature, Salinity, Sea Level and Current are available on the EuroGOOS WWW site at [http://www.eurogoos.org/documents/eurogoos/downloads/recommendations\\_for\\_rtqc\\_procedures\\_v1\\_2.pdf](http://www.eurogoos.org/documents/eurogoos/downloads/recommendations_for_rtqc_procedures_v1_2.pdf).

A Near Real Time procedure for core biogeochemical parameters is used within MyOcean and is available on the *in-situ* TAC documentation WWW page [NRT-BIO-QC-Procedure](http://www.coriolis.eu.org/content/download/8202/55503/file/MyOcean_RT_QC_BGC_v1.0.pdf) ([http://www.coriolis.eu.org/content/download/8202/55503/file/MyOcean\\_RT\\_QC\\_BGC\\_v1.0.pdf](http://www.coriolis.eu.org/content/download/8202/55503/file/MyOcean_RT_QC_BGC_v1.0.pdf)). This procedure should be tested and enhanced within the JERICO project with the aim to be endorsed by EuroGOOS as a recommendation to members.

### **3.2.4. Standard Tools for data access and visualization**

The standard tool to access real time data is available from MyOcean web portal, the main entry point to access MyOcean Products & Services.

All Products (Observation or Analysis and Forecasts) are detailed in the MyOcean catalogue at

<http://www.myocean.eu/web/24-catalogue.php> (registration requested).

Active links forward the Users to the product description (“INFO”), allow data download (“DATA ACCESS”) and show the content (“VIEW”).



## 4. Conclusions

A first executive draft of the Data Management Handbook has been attained, including the achievements obtained in SeaDataNet and MyOCEAN initiatives and trying to harmonize the distribution process with the existing tool and procedures.

This document will give accurate indications to JERICO data providers but has also to be disseminated and verified on the field, in order to give relevance to possible lacks or imperfections.

The Handbook collects large number of web references and relevant information that will help data managers in dealing with real-time or delayed-mode oceanographic data.

It represents the summing up of the guidelines for the managing and the dissemination of the JERICO data during the Service Access period.



Real Time In Situ data management system for EuroGOOS: A ROOSes-MyOcean joint effort, S Pouliquen and MyOcean INSTAC partners, October 2011, 6th EuroGOOS Conference proceedings

Recommendations for in-situ data Near Real Time Quality Control, S Pouliquen and DATAMEQ Working group, EG10.19, December 2010 ([http://www.eurogoos.org/documents/eurogoos/downloads/recommendations\\_for\\_rtqc\\_procedures\\_v1\\_2.pdf](http://www.eurogoos.org/documents/eurogoos/downloads/recommendations_for_rtqc_procedures_v1_2.pdf))

Recommendation for a PAN-European data management system for operational oceanography within EuroGOOS, S Pouliquen and DATAMEQ Working group, DataMEQ-01-2008, V2.1, June 2010 ([http://www.eurogoos.org/documents/eurogoos/downloads/recommendation\\_for\\_a\\_pan\\_eu\\_data\\_sysem\\_from\\_datameq-wg\\_v2.1.pdf](http://www.eurogoos.org/documents/eurogoos/downloads/recommendation_for_a_pan_eu_data_sysem_from_datameq-wg_v2.1.pdf))

SeaDataNet deliverable 5.7 "User manual and instructions for compiling the CDI metadata, coupling table and associated data" <http://www.seadatanet.org/Data-Access/How-to-contribute>

SeaDataNet deliverable 4.6 "User manual and instructions for updating EDMED, EDMERP, EDIOS, EDMO and CSR" <http://www.seadatanet.org/Metadata/How-to-contribute>