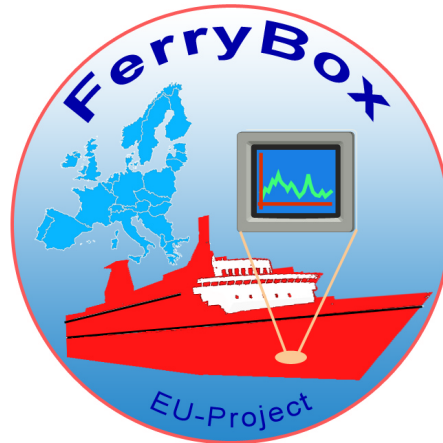


FerryBox

From On-line Oceanographic Observations to Environmental Information



The FerryBox Project Data and Information Management Plan

Incorporating general mechanisms and tools for project internal data and information distribution and dissemination

Contract number : EVK2-2002-00144

Deliverable number : D-3.1/A & D-3.3/A

Revision : 1.2

Coordinator:

Professor Dr. Franciscus Colijn

GKSS Research Centre
Institute for Coastal Research
Max-Planck-Strasse
D-21502 Geesthacht

<http://www.ferrybox.org>



Preface

This document provides a condensed management plan for data, documents and information tackled, elaborated and exchanged in the project.

It reflects the topics agreed among the partners for production, exchange and treating of data and documents.

This document incorporates and accomplishes the generic aspects on delivery and mechanisms for delivery and distribution of project data and information (refer to DOW, deliverable no. 3.3). Due to the thematic and logical overlapping of these topics they are merged and assembled in this single document.

This document does not contain specific matters of management of FerryBox data which are documented in a separate guideline and reference document (refer to deliverable no. D-3.1/B and D-3.3/B).

This project data management plan can be refined if necessary and as appropriate in course of the project.

Author: This document was compiled and edited by the leader of work package 3, K.D. Pfeiffer (HYDROMOD) following mutual agreement of the partners on proposed topics.

Contributors: Several partners contributed suggestions to this project data management plan.

Implementation

Revision 1.0 of this document was implemented and in December 2003. It is replaced by the updated revision 1.2 which was implemented and became valid in July 2004 after an review of an intermediate version.

If necessary updated revisions of the FerryBox Project Data and Information Management Plan can be implemented. They become valid after review by the project partners in charge of project data and information management issues. The review period will be specified by the project's data manager and review comments will be subsequently accomplished. Revised versions of this document entirely replace previous ones.



Document Reference Sheet

This document has been elaborated and issued by the European FerryBox Consortium.

P 1		GKSS	GKSS Research Centre Institute for Coastal Research	Coordinator
P 2		NERC.SOC	Southampton Oceanography Centre	
P 3		NIOZ	Royal Netherlands Institute of Sea Research	
P 4		FIMR	Finnish Institute of Marine Research	
P 5		HCMR (formerly NCMR)	Hellenic Centre for Marine Research (formerly National Centre for Marine Research)	
P 6		NERC.POL	Proudman Oceanographic Laboratory	
P 7		NIVA	Norwegian Institute of Water Research	
P 8		HYDROMOD	HYDROMOD Scientific Consulting	
P 9		CTG (formerly CIL)	Chelsea Technology Group (formerly Chelsea Instruments Ltd.)	
P 10		IEO	Spanish Institute of Oceanography	
P 11		EMI	Estonian Marine Institute (in co-operation with the Estonian Maritime Academy)	

This document is sole property of the European FerryBox Project Consortium.

It must be treated in compliance with its confidentiality classification.

Any unauthorised distribution and/or copying without written permission by the author(s) and/or the FerryBox Consortium in terms of the *FerryBox Consortium Agreement* and the relevant project contracts is strictly prohibited and shall be treated as a criminal act and as a violation of copyright and whatsoever applicable laws.

The responsibility of the content of this document is fully at the author(s).



The European FerryBox Project is co-funded by the European Commission under the Fifth Framework Programme of the European Commission 1998-2002 – Energy, Environment and Sustainable Development (EESD) Programme under contract no. EVK2-2002-00144.





Document Control Table

Project acronym:	FerryBox	Contract no.:	EVK2-2002-00144		
Deliverable No.:	D-3.1/A & D-3.3/A	Revision:	1.2		
WP number and title:	FerryBox WP-3	Management of FerryBox Data			
Work Package Manager:	Klaus D. Pfeiffer – HYDROMOD Scientific Consulting				
Work Package Team:	FerryBox Work Package 3 Team				
Document title:	The FerryBox Project Data and Information Management Plan				
Document owner:	European FerryBox Project Consortium				
Document category:	Project management document				
Document classification:	R – Restricted				
Status:	Implemented				
Purpose of release:	Released for implementation				
Contents of deliverable:	Project Data Management Plan for the European FerryBox Project Incorporating general project data and information related aspects and topics of mechanisms and tools for project internal data and information distribution and dissemination.				
Pages (total):	39	Figures:	0	Tables:	0
Remarks:	Specific topics of management of FerryBox data are not included in this report but provided in a separate document. Revision 1.0 implemented and in force since December 2003. Updated revision 1.1 reviewed for implementation by the FerryBox WP-3 Team. Revision 1.2 implemented July 2004.				
Editor:	Klaus D. Pfeiffer	Leader FerryBox WP-3	HYDROMOD		
Main Author:	Klaus D. Pfeiffer	Leader FerryBox WP-3	HYDROMOD		
Contributors:	FerryBox Work Package 3 Team				
FerryBox main contact:	FerryBox project coordinator:	Contact for this report:			
GKSS Research Centre Institute for Coastal Research Max-Planck-Strasse D-21502 Geesthacht, Germany	Professor Dr. Franciscus Colijn Tel.: +49 4152 87 – 42 00 Fax.: +49 4152 87 – 42 99 E-mail: Franciscus.Colijn@gkss.de	Klaus D. Pfeiffer Tel.: +49 4103 9 12 23 – 0 Fax: +49 4103 9 12 23 – 23 E-mail: pfeiffer@hydromod.de			
Project website:	http://www.ferrybox.org				

Revision Control

Revision	Date (MM-YYYY)	Released by	Distribution	Purpose of release and remarks
0.x	2002 / 2003	Project Data Manager	European Commission and project consortium	Preliminary / working revisions
1.0	10-2003	Project Data Manager	European Commission and project consortium	First year reporting Implemented December 2003
1.1	05-2004	Project Data Manager	project consortium	For project internal periodic review
1.2	07-2004	Project Data Manager	European Commission and project consortium	Periodic review and update Implemented July 2004



Table of Contents

The FerryBox Project Data and Information Management Plan.....	1
1 Introduction.....	2
2 Assignment and Responsibilities of the DIM	4
3 Implementation of the FerryBox PDIM	5
4 Project Data and Information Management Issues.....	6
4.1 Data and information management activities	6
4.2 Data reporting and summary documentation	7
4.3 Data and information types	7
4.3.1 Documents	7
4.3.2 Other data formats and file types	10
4.3.3 File compression and file archives	11
4.4 Data and information classification and availability	12
4.4.1 Public data and information	12
4.4.2 Data and information available to the consortium and the EU	13
4.4.3 Confidential data and information restricted for project related research	14
4.4.4 Confidential data with limited selective access	15
4.4.5 Project internal data and information of commercial character	15
4.4.6 Data and information available in addition to an EU RTD Project Cluster	16
4.5 Related responsibilities of the consortium members	17
4.5.1 Responsibility of the data and information provider / publisher	17
4.5.2 Responsibility of the data and information recipient / user	17
4.6 Applicable security measures	18
4.6.1 Computer / IT security measures	18
4.6.2 Additional security measures to distribute sensitive information	19
5 Data and Information Availability and Demands.....	20
6 Mechanisms and Support Tools for Project Internal Data and Information Distribution and Dissemination.....	21
6.1 The FerryBox project website	21
6.1.1 The public section	21
6.1.2 The restricted website section	22
6.2 FerryBox e-mail forwarding lists	23
7 General Aspects for Data Documentation	25
7.1 General project-internal agreements	25
7.2 Documentation and meta data for FerryBox data	26
7.3 Example for comprehensive data documentation	27
7.3.1 Mandatory minimum items	27
7.3.2 Data acquisition, measurements and sample collection	28
7.3.3 Data processing and sample analysis	28
7.3.4 Quality control and quality assurance information	29
7.3.5 Data Formats	29
7.3.6 Field data	30
7.4 Site-related data documentation	34
8 References	35



1 Introduction

In this project related Data and Information Management Plan the general rules, guidelines and procedures for data and information production, dissemination and archiving will be laid down. This document will be successively refined according to both needs and demands by the project partners as well as with respect to requirements to fulfil the project work. These are to be mutually agreed upon among the project teams concerned prior to binding implementation.

This document does not include specific topics on management of FerryBox data which will be elaborated and documented in a separate guideline documentation and report.

Within the FerryBox project large amounts of data and information are and will be acquired, processed and integrated. Measured data result from both FerryBox operations during the project. Other project-relevant data and information is already available or produced in parallel to the project within various national and/or affiliated projects conducted by the project partners. This data and information has to be consistently documented and merged. This will ensure its effective utilisation within the project itself as well as beyond by the project partners and/or authorised third parties. When finally archived, this will ensure long-term availability and usability after the project has been completed.

The prime goals of project related data and information management (PDIM) are coordination, harmonisation and standardisation of related activities. Project related data and information management is considered as being a project related service to ease intra-project data and information exchange as well as document and information production, dissemination and long term usability later on. Further PDIM aims on provision of a certain reporting and document standard which includes also a corporate identity and visibility of the project deliverables as well as elements of document control in terms of quality assurance and document control procedures.

Generally data and information management activities ought to be in compliance with the applicable guidelines in data and information management and thereto related rules, international standards and regulations. In addition they should be not too complicated and sophistic to be mutually agreeable and consistently implemented in a multi-disciplinary and multi-national research project.

State of the art data and information management does not necessarily require demarcation between (numeric) data (i.e. measurements) but also includes comprehensive related data documentation (in particular corresponding meta data) and a variety of other information. In the area of electronic media and possibilities of object inter-linking the discrimination between pure numeric (measured) data, results from numerical models, reports and other hyper data (e.g. images, photos, audio or video files) becomes more and more ambiguous. In the FerryBox project we consider that a wide range of data and information and especially therefrom derived key results are accordingly incorporated, assembled and compiled into final deliverables and especially in the final reports in appropriate ways. Despite these possibilities important data shall be available and accessible both internally and externally on lower levels (e.g. as plain ASCII data or assembled in standard respectively very commonly used applications and formats).



Efficient information and data exchange among the different scientific teams comprise an essential requirement. The dissemination of project-related public information and scientific results achieved during the project's operative phase via the Internet and the World Wide Web (WWW) is a standard in EU research and development projects. This will be applied in the FerryBox project as well and whenever possible and appropriate. Therefore the FerryBox website was set up which comprises a focal point for data and information exchange and dissemination for the project. This website is implemented by HYDROMOD and is maintained by the project coordination team.

The portal page is implemented at a provider and accessible through the project registered URLs

- www.ferrybox.org,
- www.ferrybox.net,
- www.ferrybox.com, and
- www.ferrybox.de.



2 Assignment and Responsibilities of the DIM

Within the FerryBox project data and information management is assigned as a task clustered under the project management activities.

The FerryBox PDIM is responsible for the following:

- (a) To coordinate, monitor and, where necessary, steer and take required management action regarding data and information production, exchange and flow within the project in particular with respect to in time production, availability and delivery of data and information
 - within project / task teams,
 - across work packages / tasks,
 - to the project coordinator,
 - to the funding organisation and other external addressees via the project coordinator respectively on his authorisation and release.

- (b) Further the PDIM is in charge to ensure that data and information produced, made available or delivered within the project
 - complies with requirements of the project teams and addressees (in terms of quality, formats, standards, documentation etc.),
 - are uniquely referenced and possible to identify,
 - are accordingly classified and sufficiently documented for third party use (if allowed),
 - complies with international acceptable / adopted standards with respect to quality and documentation (including corresponding meta data), and
 - are archived, accessible and usable in the long term (if applicable).

- (c) Finally the PDIM supports the project partners and teams in order
 - to harmonise and optimise data and information exchange,
 - to provide efficient distribution and dissemination measures / methods,
 - to develop and provide general guidelines for data and information processing, production, distribution and dissemination in mutual agreement with the project consortium, and
 - to give assistance in technically related matters should this be required.



3 Implementation of the FerryBox PDIM

As mentioned above the FerryBox project data and information management is assigned as a task clustered under the project management activities and this comprises a steering and controlling activity plus related supportive activities.

For this project the following implementation guidelines shall apply:

- The individual work packages are responsible for that the data and information needed to conduct the work are acquired and made available in such a form that they can be used by all teams and persons concerned. In case of problems or difficulties the work package or task leader in charge shall take appropriate action of finding a solution. Only if a problem cannot be achieved on the working level he shall report such problems and difficulties to the data manager who then seeks for appropriate solutions and takes (if required together with the project management) appropriate actions.
- The same shall apply for such data and information which are required to be acquired, produced or distributed across work packages or which are necessary as input for another work package. Such matters shall primarily be agreed upon and solved among the work package teams concerned under coordination of their task leaders.
- Any problems and delays in data and information production, provision, availability and / or distribution shall be reported to the data manager via the task leaders.
- The PDIM supports the project team in technical matters with respect to data and information production and distribution and proposes guidelines and general rules to be adopted by the project teams and laid then to be incorporated into the project's Data and Information Management Plan as binding.
- The PDIM is entitled to take appropriate management action to solve such problems either identified but not to solved on the working level or identified by the PDIM itself within related monitoring and control activities.
- The PDIM is further entitled to return data and information produced within the project and demand corrective actions if such do not comply with respect to adoption of mutually agreed project internal guidelines, international standards (if adopted) and / or if deficiencies in quality and / or documentation are identified.

The implementation of PDIM measures shall:

- Give as much as possible flexibility, freedom and assignment of responsibility on the working level with (if everything goes well) minimum interference from the management's side.
- Provide a minimum set of commonly developed and/or mutually agreed guidelines and rules and thus optimises the co-operation regarding data and information sharing and joint use.
- Allow controlling and corrective actions when required on the project coordination and management levels in case of problems.
- Guarantee that data and information produced and distributed comply with project internal demands and requirements as well as with generally or internationally adopted criteria.



4 Project Data and Information Management Issues

4.1 Data and information management activities

In this section the general requirements for Data and Information Management (PDIM) as generally demanded by the European Commission for research and development projects conducted under the Fifth Frame Programme (FP5) are brought into connection with the project specific requirements and objectives. As no general guideline for data and information management is available for the Fifth Framework Programme elements of the MAST Code for Data Management issued within the Fourth Framework Programme to RTD projects conducted under the Marine Science and Technology (MAST) Programme can facilitate as a general guideline. This code reflects state of the art and applicable international standards in DIM for marine data. As most FerryBox data and information are of similar character certain issues from this code can be adapted to the project.

- PDIM activities concern the following project activities:
- Collection / acquisition, processing and evaluation of field data,
- in case of numerical models – production of model results,
- followed by pre- and post-processing of such data and results,
- including transparent documentation and quality control,
- thereafter dissemination and in time delivery for project internal use,
- and finally their utilisation for exploitation measures, and
- final archiving for long term use.

It is not the task of the PDIM to conduct these actions but to coordinate these in conjunction with other coordinating measures and to bring them consistently in line with the project's operative tasks.

The integration of data and information into the final deliverables and in particular into final reports is conducted by the work package / task teams concerned and supervised by the project management. The PDIM will link up with these teams and thus ensure that processing, merging, integration and documentation of data and information complies with applicable standards and agreed / implemented policies.



4.2 Data reporting and summary documentation

In order to give the European Commission, affiliated or clustered projects and all interested institutions – such as data centres, intergovernmental organisations, research undertakings and commercial units – the possibility to monitor data and information gathering and production activities of FP5 projects minimum conditions for reporting are applied.

4.3 Data and information types

The FerryBox project will collect / produce a wide range of data and information which due to the multi-disciplinary and integrative character of the project are expected to be quite heterogeneous. Due to the heterogeneity of application programmes and software used by the different teams, we have decided to apply some minimum standards to ease and ensure project internal compatibility and availability of data and information and to make data and information accessible and usable for all project teams which are given in the subsequent paragraphs.

4.3.1 Documents

4.3.1.1 Documentation language

All project documents shall be generally elaborated in UK English.

4.3.1.2 MS-Word Documents

Documents which are in the stage of elaboration and thus are to be accessed and worked on by various project teams shall be generally compiled by the text processing software Microsoft Word in Microsoft Office™ (respectively MS Word97 / MS Word 8) compatible format (file extension doc). In case this poses problems for the one or other project team due to software or version incompatibility such temporary MS-Word documents can be distributed in Rich Text Format (file extension rtf). This can be exported from any MS Word program version and is also readable by other word processors and older software versions. Project partners using newer versions of MS Word shall set there store setting accordingly and thereby ensure that they export a compatible file.

Finally compiled documents shall be generally compiled, circulated and in particular distributed to third parties in the Adobe™ Portable Document Format (file extension pdf) and not as MS-Word files. This ensures document originality, integrity and security as well as general readability and accessibility.



The project management may give certain minimum guidelines and hints for production, compilation and layout of documents in order to ensure a common appearance and corporate identity of project documents (in particular for final reports and reports delivered to the EU or distributed outside the project consortium).

The project's first year documentation shall facilitate as a proposal and example including.

4.3.1.3 Final documents in Portable Document Format

For document integrity reasons we do not recommend to distribute MS Word files respectively files in the form as produced by the text processing software. Such documents shall be generally produced in Portable Document Format (file extension pdf). This ensures document originality, integrity and security as well as general readability and accessibility by the public and freely available reader software (the so called Acrobat Reader which can be downloaded from any suitable Adobe™ website – e.g. <http://www.adobe.com>).

In particular all documents which are distributed outside the project consortium shall be generally compiled as secured PDF files meaning that they can be displayed and printed but not altered, extracted as a whole or in parts and/or modified. This requires that an appropriate security password plus the respective security options are to be set during PDF file production. The PDIM will take care of these issues.

If not otherwise agreed, PDF files shall be compiled Acrobat version 4 compatible.

4.3.1.4 Hyper text documents

Documents which are intended to be published on the project website shall be compiled in HTML format readable by state of the art browsers (in particular Netscape and MS Internet Explorer with version 5 or higher). Please avoid special, application or version dependent features and adhere to generally accepted HTML standard (typically HTML 4.0 or HTML 4.01 transitional – refer for example to the Web Design Group (<http://www.htmlhelp.com>) for further information and validation). All internal links shall be set as relative ones in order to ensure easy integration and proper browsing.

4.3.1.5 Photos and images

Photos and images shall be generally in a content complying raster (bitmap) format. We recommend for photos to use generally the JPEG format (file extension jpg) with full 32 bit colouring and good to high quality image compression.

For raster graphs and images with a few colours only and for black and white ones we generally recommend the Graphic Interchange (GIF) format (file extension gif).



Due to file size and / or compression incompatibilities (if image compression is applied) we do not recommend the use of the TIFF format except for photos which require very high quality (for instance to use on large posters and for print material). We also recommend not to use Windows or raw bitmap formats.

Photos and images shall be produced and compiled in an adequate resolution adapted to the intended use. This implies proper scaling regarding image size and/or pixel respectively raster resolution complying with the document, application and intention of use.

As a general guideline for photo and image provision the following is recommended:

- 96 dpi to 150 dpi adapted accordingly in colour depth, compression ratio and quality not exceeding approximately 500 to 600 pixels in length and/or height for images intended for publication on the project website (i.e. not more than preferably some 10 up to a maximum of approximately 100 kilobytes in size).
- 150 to 200 dpi (seldom and for very high quality documents only 300 dpi and more) resolution with adapted colour depth for raster images to be compiled in printable documents. These should be scaled in the size as they will appear in the document prior to integration (i.e. do not drag, drop and “mouse-scale” a poster sized photo of several megabytes file size on a small part of an A4 page).
- 200 to 300 dpi (seldom and for very high quality documents only 600 dpi and more) resolution with full 32 bit colouring for photos to be compiled in high quality print material. These should be properly scaled in the size as they will appear in the document prior to integration.

4.3.1.6 Remote sensing imagery

Exception of the above applies in general for images which need to keep the information contents in original resolution, grey-scaling or colouring. Typically this is the case for remote sensing imagery by optical, infrared, radar or other sensors. This is of special evidence when remotely sensed images are provided for evaluation and further processing.

No further details are specified in this conjunction as the teams concerned with remote sensing images handle this accordingly and in line with their long-lasting experiences in this field.

4.3.1.7 Vector graphics

For compatibility reasons we strongly recommend to distribute project internally vector graphics in the Windows Meta Format (file extension wmf) only. This format can be exported by almost all commonly used graphic and drawing applications including most CAD and GIS software.



4.3.1.8 Numerical data and corresponding meta data

Numerical data in FerryBox will most likely result from measurements or mined observational data and possibly include some key model results. As so far foreseeable most of these data as well as the key model results will be multi-parameter time series or multi-parameter vertical profiles.

We recommend to circulate and distribute such numerical data generally in plain ASCII format (file extension dat) or – as many of the teams are used to this, as MS-Excel spreadsheets – in a simple tabular format (e.g. comma (csv) or tab separated).

These data shall be accompanied by appropriate meta data which has to facilitate use by other project teams and authorised third parties both in the short and long term.

We will propose a simple and common data format both for the numerical data themselves as well as for the meta data including the definition of their contents in an as far as possible harmonised manner and in a way that best compatibility to international applied standards is ensured. Details will be agreed upon and compiled in an attachment to this PDIMP.

4.3.2 Other data formats and file types

For certain purposes the use of other data types might be necessary and appropriate. Those foreseeable within the project are reflected in the paragraphs below. Where necessary, project teams using such data should agree on compatibility and exchange matters internally. However, use of such formats or file types should not exclude any other of the project teams from access and readability of the data, respectively this should not require special investments in readers or retrieval software. This means that the finally elaborated documents, images or data shall be distributed in one of the applicable formats given above.

4.3.2.1 MS-Powerpoint presentation

MS-Powerpoint files or applications are expected to be most likely used for presentations. Should these be integrated in or distributed as separate documents we recommend conversion to PDF files.

4.3.2.2 MS-Excel spreadsheets

MS-Excel files or applications are expected to be commonly used for data provision in the project.



4.3.3 File compression and file archives

The shareware software WinZip or the freely available zip programme is installed at most if not all project partners.

In order to reduce file size and traffic amount the following should be distributed compressed:

- larger ASCII files,
- MS-Excel spreadsheets,
- images in uncompressed bitmap formats (bmp or tif),
- larger documents produced with older MS-Word or MS-Powerpoint versions (which do not have internal compression) as well as in Rich text Format (rtf), and
- larger hypertext documents.

Project teams who do not have WinZip installed may use the UNIX or DOS compression routines ZIP or GZIP which can be handled by WinZip as well.

We do not recommend to use other archiving and compressing software due to possible compatibility problems.



4.4 Data and information classification and availability

The FerryBox project will collect and produce a wide range of data and information which due to the multi-disciplinary and integrative character of the project is expected to be:

- Data and information which are to be made public – to be labelled and classified as **Unclassified** or **Public**.
- Data and information which are to be kept internal to the project, if the project participants so require – to be labelled and classified as **Restricted to the FerryBox Consortium**.
- Data and information which are made available **for project related research to the entire project consortium only** – to be labelled and classified as **Restricted** or **Confidential – for FerryBox research work only**.
- Data and information which are made available **for project related research only to a limited part or even single persons of the project consortium only** – to be labelled and classified as **Confidential – limited access for selective FerryBox research work**.
- Classified **commercial information** which might be produced within the project – to be labelled and classified as **Restricted or Confidential – Commercial in confidence**.

As a guideline for classification of data it is also referred to applicable EU policies and in particular to the rules and regulations laid down in the project contract, applicable attachments and referenced documents.

4.4.1 Public data and information

Public data and information are available to and accessible for the general public without restrictions. This data must be labelled with **Unclassified** or **Public** and applicable ownership statements.

No usage and distribution restrictions apply, except those related to citation, ownership, use and copyright restriction as demanded by the publisher himself.



Typical public data and information:

- Project deliverables which are explicitly mentioned as Public in the project's Technical Annex (work description).
- Data and information which are disclosed to the general public on decision of the project consortium (if jointly owned) or by an individual project partner (if owned solely individually).
- Any other data and information from arbitrary sources which were already in or disclosed to the general public before being incorporated in project related data and/or information assemblies and/or documents.

4.4.2 Data and information available to the consortium and the EU

These types of data and information is owned by and is to be made available by and to the project consortium.

This data is generally **in joint ownership of the project consortium** unless otherwise explicitly mentioned in project related agreements.

Distribution and availability restrictions apply according to applicable rules and regulations for FP5 RTD Projects. Other applicable EU regulations may pertain as well. Also compliance with decisions of the project consortium or the owner (if not the project consortium itself) may apply.

Respective data must be labelled **Restricted** or **Confidential** and applicable ownership statements and availability restrictions must be clearly mentioned. The latter must be in compliance with applicable EU and European Commission regulations. EU Member States having co-ownership and may use such data accordingly.

It must be noted that for certain data and information the EU requires banking in a data centre of the consortium's choice. Such data and information may become available for wider use or even public after a dedicated period of time (commonly after 2 to 5 years). For further details concerning data custody, archiving and final banking of FerryBox data it is referred to the separate guideline and reference document on management of FerryBox data (deliverable number D-3.1/B & D-3.3/B).

In addition arbitrary data providers, other EU project consortia, national institutions and agencies or this particular consortium member who make such data available to the FerryBox project consortium may decide upon other restrictions to be applied individually to such data sets or information as long as such rules do not infringe aforementioned EU regulations.



Typical restricted data and information are:

- All data and information acquired / produced during the project's operative phase which are directly produced or related to "on project work".
- The project deliverables labelled as restricted or confidential in the project's Technical Annex (work description).
- Any other data and information which are made available by one or more consortium members for use of the consortium under this classification and above restrictions.
- Data and information which are not allowed to be disclosed to the general public on decision of the project consortium (if jointly owned) or by an individual project partner (if solely individually owned),
- Any other data and information from arbitrary sources for which above restrictions are applicable and if these restrictions do not infringe the restrictions of the owner or publisher.

4.4.3 Confidential data and information restricted for project related research

As far as foreseeable certain historic or actual data, either owned by individual consortium members or by any other third party when used by or distributed to all or selected other consortium members for project related research work must be classified as

Confidential – restricted for project internal research work only.

Such kind of data and information are commonly:

- Historic data made available, purchased from or solely owned by a third party according to the availability and usability restrictions applied by the data owner or provider.
- Data and information made available by a project partner to the consortium or to selected other consortium members which were already in the possession of or solely owned by the providing party before commencement of the project.
- Any other data and information which were produced solely outside the project by a project partner and made available under these classifications and such restrictions.

Above all to apply this classification, such data and information **must not have been produced or acquired within or in direct relation with the EU funded project work**. In addition no other applicable availability regulations of the EU or any other funding organisation must be infringed.



4.4.4 Confidential data with limited selective access

Data and information might be made available to one or more consortium members or even to selected persons for specific research work but restricted for limited accessibility and use.

Such data and information shall be specially protected by applicable agreements with or guarantees to their provider / owner.

Such data and information must be labelled with

Confidential – limited access for selective FerryBox research work.

If not explicitly excluded it should also be possible to cite and reference such data as well as to list related meta data in the project's data and information dictionary.

However, no restrictions shall apply in incorporating research results and derivatives elaborated within the project using such data for project deliverables: These are be jointly owned by the project consortium as laid down in the previous section.

4.4.5 Project internal data and information of commercial character

This kind of data and information is of commercial character. It could be produced within the project itself or may include other information of commercial character.

Typically such comprise:

Commercial agreements regarding future exploitation and/or commercialisation of research results (both overall or bilateral ones).

- Licenses, patents and related information.
- Data and information related to or arising from commercial spin-off activities.
- Any other commercial data and information if accordingly classified and not infringing applicable EU regulations or other consortium internal agreements.

Such data and information must be handled with utmost care in order to ensure that commercial rights and interests, as well as investments of the providing party, are secured.

If supplied to the entire project consortium the information is owned by it jointly. If provided bilateral or to a limited number of consortium members it should be jointly owned by these parties only.



4.4.6 Data and information available in addition to an EU RTD Project Cluster

In case the project would be clustered with one or more other thematically affiliated EU projects certain data and information have likely to become available for and shared with such a cluster.

In such cases and if not already in the public domain, such data shall be **accordingly labelled independently from their classification**. Restrictions in availability and use must be clearly brought to the attention of the cluster respectively these parties of the project cluster.

The parties of the cluster to which such data and information are made available are then fully responsible (and if necessary must confirm this to the consortium respectively the owner) that such data and information is used in full compliance with any restrictions applied.



4.5 Related responsibilities of the consortium members

4.5.1 Responsibility of the data and information provider / publisher

Each provider / publisher of data – respectively each consortium member distributing data and information within the consortium or to any other third party – is fully and solely responsible for

- the contents of distributed / published data and information,
- the appropriate and compliant classification of data,
- the compliance of the classification and other applied restrictions with any applicable or related rights of himself and any other third party,
- providing complete information to the receiving parties about the applied classification and restrictions,
- any damage which might occur due to non-compliant classification or insufficient or incomplete application of restrictions.

In addition the provider / publisher shall hold free the receiving party from any damage or claims which occur in conjunction with published / provided data in case of improper classification, application of restriction or insufficient information of the receiving party thereon.

Remark : The above indemnification is of particular importance if data and information are made available for publication on the project website. This applies especially in conjunction with third party copyrights for documents and / or photos.

4.5.2 Responsibility of the data and information recipient / user

Each recipient and/or user of data and information within the consortium, or any other third party, is fully and solely responsible for

- the appropriate and compliant treatment and storage of data and information according to their classification and applied third party rights and restrictions,
- any damage which occurs by non-compliant or improper treatment, handling, disclosure or distribution of received data and information.

In addition the recipient and/or user shall indemnify and hold free the providing / publishing party (parties) from any damage or claims which occur if he improperly treats or distributes such data and information.



4.6 Applicable security measures

The FerryBox Consortium has in principle agreed upon the below described security measures with respect to computer safety and security as well as to maintain data and information confidentiality and restrictions.

4.6.1 Computer / IT security measures

Each project partner shall take utmost care and apply in house respective measures to avoid damage on computer systems, software and data of other parties when distributing data and information by electronic media in particular by electronic mail and by files provided for downloading from the project website.

This implicitly requires:

- Installation of an appropriate, state of the art and actual software which scans files intended for distributions for computer viruses and other damaging or hazardous components. Such software shall detect with high probability and reliability such damaging components which besides state of the art programs requires frequent actualisation of data files incorporating related virus characteristics or profiles (i.e. a state of the art virus scanner is not sufficient if improperly configured or related profile data are not up to date).
- Avoiding distribution of executables if not explicitly necessary as well as of components which are capable to perform actions on the recipients computer systems.
- Avoiding incorporation of macros into documents and spreadsheets (if not possible to avoid at least implementation / switch-on of macro virus protection measures).

Application of the above can substantially reduce the risk of virus distribution and infection of recipients' computer systems but never be a full guarantee and protection. It is therefore of utmost importance to inform and notify immediately all project partners concerned in case of suspicions or detection of a computer virus or other damaging components in electronic distributions.

For the same reasons each project partner shall take appropriate measures to protect and secure his own computer systems including remote office and/or home PCs as well as Laptops.



4.6.2 Additional security measures to distribute sensitive information

Each consortium member providing data and information can apply additional security measures should if considered as necessary or applicable or to comply with protection and security measures applied by third party data providers. This may also apply to data and information being, entirely or including in part, sole property of the data provider and which are provided to the project consortium for use and conduction of the research work.

Applicable additional security measures can be

- availability and distribution restrictions,
- protection by separately and secure provided opening passwords (e.g. for PDF and MS-Office files),
- protection by separately and securely provided archive passwords,
- coding or keying of data and information as per accordingly agreed procedures between provider and recipient(s).



5 Data and Information Availability and Demands

The project partners will communicate their data and information demands and requirements as well as such data and information they will acquire, process and produce to the consortium.

Data and information deliveries shall be envisaged by the provider including minimum information on

- the provider,
- the owner,
- classification,
- restrictions in use and/or distribution (if applicable),
- description and content in brief,
- time of delivery / availability within the project's schedule.

Data and information demands shall be addressed including minimum information on

- the demanding consortium member,
- requirements, content and expectations in brief,
- time required within the project's schedule.

Both providers of and users demanding for data and information shall also report the above to the project's data and information management unit.

If contradictions between provision and demands are identified both the work package teams and/or leaders concerned will harmonise these together with the PDIM in order to best optimise and bring in line the schedules of the data provider and the data user as well as maintaining project work flows and time plans.

Upon receipt and accomplishment of the above the PDIM will compile a data availability and distribution schedule which will be incorporated into subsequent revisions of the project's Data and Information Management Plan.



6 Mechanisms and Support Tools for Project Internal Data and Information Distribution and Dissemination

Within the project coordination and DIM work several activities give support to the project consortium in distribution and dissemination of data and information.

6.1 The FerryBox project website

As a coordination activity a project website (e.g. <http://www.ferrybox.org>) has been set up and is maintained for the duration of the project and a considerable extension period beyond.

It shall be noted once more that for material being published (respectively being delivered for uploading) on the project's website (whether on the public or on the restricted area) the aforementioned responsibilities of the publisher apply.

6.1.1 The public section

Within the website's public section unclassified information and results are made available to the general public as well as to the interested scientific and user communities. As long as the project is operative this site will be periodically updated and supplemented as per progress with new results and other information.

When the project has closed out the site will be maintained further on for a limited amount of time in order to support project related exploitation activities. However, then the site will become more and more static and less frequently actualised.

It is the responsibility of each project partner to support dissemination and exploitation activities by submitting and keeping up to date relevant information.

The public section of the project's site is successively filed and supplemented with information as becoming available during the project.

Despite certain periods of more intensive updating as required by project activities the routine actualisation and updating schedule is once per month in average.

The coordinator and the project partners are encouraged to provide suggestions and wishes for upgrading of the public website section and routinely / periodically deliver information in line with the guidelines and procedures as laid down in this data management plan.



Partners shall also suggest in which manner they wish such information to be incorporated and published on the website. The project's web team will adhere to ideas and suggestions as applicable and in line and limits with both technical feasibility and therefore assigned project specific resources.

In particular all data and information delivered with intention for publication on the public section of the project website section must be free of Third Party copyrights or any other restrictions pertaining or infringing with publication on the Internet.

Each partner delivering information or data has the sole responsibility that these regulations are kept and obeyed. The project's web team shall be hold free from the responsibility of possible infringements in case of missing, incomplete or wrong notification or classification by the data and information providers concerned.

6.1.2 The restricted website section

If required and upon availability of applicable material the FerryBox website also incorporates a part with restricted (password protected) access for the consortium members. This part of the project website – the so called Partner's area – is intended to support and ease retrieval, downloading and distribution of data and information limited in access to the consortium members.

The partner's area will contain for example relevant information, templates, reports, data, information and reports delivered to the project management or compiled and made available for the consortium by the project coordinator. This part of the website will also be periodically updated and supplemented with new results, documents and other restricted information.

However, please bare in mind that this security / protection level (applied here as agreed upon among the project consortium) for the restricted area of this site is rather low.

The restricted section of the project's site is successively filed and supplemented with information as becoming available during the project.

Despite certain periods of more intensive updating as required by ongoing project activities the routine actualisation and updating schedule is once per month in average.

The coordinator and the project partners are encouraged to provide suggestions and wishes for upgrading of the restricted website section and routinely / periodically deliver material, data and information in line with the guidelines and procedures as laid down in this data management plan.

Partners should also suggest in which manner they wish such information to be incorporated and made available for distribution in the restricted website section. The project's web team will adhere to ideas and suggestions as applicable and in line and limits with both technical feasibility and therefore assigned project specific resources.



In particular all data and information delivered with the intention of publication on the restricted section of the project website section must be available for the entire project consortium for project specific and internal work and use. No other limitations or restrictions in use shall apply therefore.

Materials which were disseminated to project partners by other means – in particular through e-mail distribution as described in the section below as well as through data and report collections on CDs are not intended for publication on this section of the website. This applies in particular for large material collections and reports.

6.2 FerryBox e-mail forwarding lists

To support distribution of relevant project information to all addressees concerned email alias address are provided by the data management which facilitate distribution and communication within the FerryBox project.

Different forwarding lists are at present distributing e-mails automatically to

- all partner contacts concerned with the project,
- the project scientists at the partners,
- the project participant's administrative contact persons,
- the project's steering committee,
- each team member associated with work packages 2 to 6.

The PDIM compiles and maintains these aliases and an address list. This document will be distributed whenever actualised. These aliases expand to a list of individual addressees (as long as these have an individual email address otherwise once per given collective / institutional address) which are named by the consortium members as recipients concerned.

Furthermore, the email alias ferrybox@hydromod.de can be used to communicate with the data management team. It includes all addressees involved in the FerryBox Data and Information Management on side of HYDROMOD – namely as per present stage **Klaus D. Pfeiffer** (pfeiffer@hydromod.de) and **Dirk Goldmann** (goldmann@hydromod.de).

It is important that each project partner reports changes of recipient's addresses as well as addressees to be added to or deleted from these lists directly the PDIM team or to the project coordinator. Only by this it can be ensured that relevant project related information reaches all the addressees properly and in time).

In conjunction with large attachments it should be kept in mind that some recipients may have applied email size limitations on side of their mailhosts.



Further each sender using this alias bears full and sole responsibility on the contents as well as the compliance of attachments with commonly agreed formats and/or file types. On HYDROMOD's side the address is expanded only and messages are forwarded to the listed addressees exactly as received. To a limited extent HYDROMOD can monitor correct receipt by automatically generated notifications in case an email is undeliverable.

HYDROMOD also does not take any responsibility for whatever damage might occur if via this alias computer viruses or other damaging elements are transmitted.

HYDROMOD especially takes neither responsibility for the contents of such emails nor on the readability on side of the individual recipients. Furthermore, should traffic generated over this alias block our mailing or routing system or hamper HYDROMOD's Internet access due to non compliance with the above HYDROMOD reserves the right to close down this service upon short term notice.

These aliases shall not be disclosed to the general public for security reasons and to avoid e-mail spamming.



7 General Aspects for Data Documentation

Data documentation including compilation and provision of corresponding and comprehensive meta data is a general problem in data and information management. It is also the reason for low data usability and availability in particular if such data and information shall be used outside the originator's institution.

In general and as recommendation on completeness and comprehensiveness of meta data and data documentation the following may provide some guidance:

“Document and supplement your data and information in such a way as you like to have them documented if receiving them from others”.

The topics in the following sections shall provide some hints and a general guideline how state of the art data documentation and meta data shall be done.

The FerryBox project applies a jointly agreed minimum subset of meta data and documentation which has to be provided together with the numerical data.

A separate guideline and reference document is compiled which described in detail compilation, assembly and documentation of FerryBox data (refer to deliverable no. D-3.3/A & D-3.3/B).

7.1 General project-internal agreements

Commonly there appear a lot of problems in usability of data provided by third parties due to a few simple matters which can easily be avoided. These are in particular:

- **Insufficient, inhomogeneous or incomplete geographical references**

This is avoidable by using generally geographical coordinates in compliance with the WGS 84 geoid which is GPS / DGPS compliant as mutually agreed by the consortium.

Should this be impossible in rare cases at least sufficient references (e.g. map edges) shall be provided geographical coordinates. Simultaneously used relative or other coordinates shall be accordingly documented.

The consortium has further agreed to concert all geographical coordinates in decimal degrees which facilitates easy retrieval and incorporation in other applications (especially in GIS and other mapping software).



- **Insufficient, mixed or inconsistent time references**

The project consortium has already agreed to use generally UTC time referencing.

If this is not possible in all cases at least a reference (i.e. the time shift relative to UTC whereas “+” indicates later and “-“ earlier) must be provided and clearly denoted together with the data.

The commonly agreed format for date and time stamps is YYYYMMDD for dates and HHMMSS for time values in 24 hour notation.

If applicable and to ease evaluation of numerical data series associated with date and time stamps we propose to provide an additional time stamp to be provided as a “large integer” value counting the seconds from the beginning of a reference or start time (e.g. seconds from start of a measurement registration).

7.2 Documentation and meta data for FerryBox data

These themes are documented in detail in a separate public guideline document for the management of FerryBox data. For further details one is referred to the deliverable no. D-3.3/B & D-3.3/B.

The subsequent sections provide general suggestions and examples on meta data contents and comprehensiveness.



7.3 Example for comprehensive data documentation

In the following subsections and paragraphs the contents of state of the art data documentation and meta data contents is given as derived from internationally elaborated and widely adopted standards and guidelines (among others EU-MAST Guidelines, ICES – International Council for Exploration of the Sea, NOAA – National Oceanographic and Atmospheric Administration (USA), WOCE – World Ocean Circulation Experiment – Data Management Guidelines, EEA – European Environmental Agency, BODC – British Oceanographic Data Centre, IFREMER/SISMER – the French Oceanographic Data Centre, ESA – European Space Agency).

They should facilitate as examples and something like a “selection list” to adjoin and accompany numerical data with compliant meta data contents facilitating efficient use of data throughout the project and beyond.

7.3.1 Mandatory minimum items

Mandatory minimum information which ought to be provided as documentation (meta data) together with measured oceanographic data are

- data owner,
- data provider,
- data classification,
- restrictions for use, distribution and dissemination,
- type of data,
- brief description of contents,
- contact for further information (institute, person, address, fax, phone),
- general references,
- geographical references / positions,
- depth / height reference (zero level),
- date and time references (DTG – start and end date and time of measurements),
- parameter names,
- parameter abbreviations,
- parameter units,
- format and electronic reading / computer access information,
- data availability, processing, quality control and archiving information.



7.3.2 Data acquisition, measurements and sample collection

Data acquisition and sampling shall be supported by comprehensive documentation. It is recommended to link such to external documents best provided through world wide web links (URLs):

- Description of the equipment used: Name, gear code and details of its deployment as appropriate.
- If possible, references for the method acquisition method and instrumentation deployed.
- Description of the measurement platform and the techniques used for positioning of the platform and equipment.
- Details of calibration of the equipment used, documentation of details on sampling frequency.
- List of persons involved (e.g. party chief, chief scientist, contacts).
- Description of environmental conditions as appropriate.
- Any limitations associated with using the chosen sampling strategy should be documented.

7.3.3 Data processing and sample analysis

Methods applied for data processing and sampling shall be comprehensively documented.

If possible to maintain in the longer term it is recommended to link such to external documents through world wide web links (URLs). Otherwise such documents should be provided together with the data.

Such documentation shall typically include:

- Documentation how the sample was preserved and whether it was processed in-situ, onboard or in the laboratory; name of the laboratory where samples were processed.
- Description of different stages in sample processing (including instrumentation used) from treatment of sample from when it is collected through to final processing of sample to the point where data is generated (details of the methods/instrumentation used should be given with full references).
- Description of any instrumental error corrections made during the course of processing the sample.
- Description of the expected precision, accuracy or reproducibility of the used methodology and the limits of detection.
- Details if the methodologies employed in sample processing have been validated.



- Comments on any limitations associated with used sample processing techniques.
- Use of references to published literature where ever possible in order to shorten and compact the information.
- Description of the different stages in processing and analysing the data including reduction algorithms and statistical analyses, etc. Details of the methods used should be given with full references.
- Description of algorithms and computer programmes used.
- Details of the degree in precision, accuracy and reproducibility of further processed and/or analysed data.
- Limitations incurred in using these sample processing techniques
- Use of references as appropriate - give full bibliographic reference.

7.3.4 Quality control and quality assurance information

Accordingly methods applied for data quality assurance shall be comprehensively documented (as above by URLs or documents distributed together with the data).

Quality control and assurance may apply at each level of the data generation steps. Therefore for each step meta data shall include comprehensive information on the quality of the data and the quality assurance methods applied.

This comprises typically:

A comprehensive cruise description (e.g. in a Cruise Report).

Methodology documentation which have been applied to standardise and/or calibrate the data.

Descriptions of data validation methods used in response to results of calibration and inter-calibration measures as well as, if available, in comparison with standard methods.

Details on detection of limitations, gaps or errors in particular in position and time.

7.3.5 Data Formats

When compiling data files, it is essential that the contents of a data file can be clearly understood by a subsequent user and that the essential information such as station numbering, references to position, time and depth are always clearly identifiable. All data formats must be well described to ensure that there are no misunderstandings and that the data will be usable by the receiver.



Cross reference:

The project consortium has already mutually agreed to develop a simple but comprehensive and easy to handle set of meta data as well as a simple common format for exchange and dissemination of FerryBox data. This public document shall facilitate intra-project data exchange as well as data dissemination to other users. Details are documented in a separate guideline document (refer to deliverable D-3.1/B & D-3.3/B for further details).

7.3.6 Field data

7.3.6.1 Country of the platform

Designates the country of registration of the data acquisition platform (e.g. ship, ferry, buoy, station). It is advisable to use the IOC country and platform code tables to indicate the country and platform.

This is not necessarily if the country of the platform's registration is the same as the country of the organisation or project responsible for collecting the data.

If applicable, it is generally helpful to supply the ITU call sign of a ship so that any data sent in real-time may be easily identified. Where it is not possible to use a code table, supply the name of the platform used in the data collection in plain text.

7.3.6.2 Cruise identifier (if applicable)

The originator's cruise identifier. This may be recorded as part of the data or in a Cruise Summary Report. The cruise identifier is a convenient and useful way to group data collections. Many data centres use this as part of the identification and archiving process.

In general FerryBox data is likely to require specific designations of ferry routes and possibly also blocks of transits or legs or single cruises of voluntary ships of opportunity. The guideline document on management of FerryBox data (deliverable no. D-3.3/A and D-3.3/B) describes in detail how this is implemented and suggested by the project partners.



7.3.6.3 Data units and values – some general remarks

If ever possible, data shall be reported in standard SI oceanographic units (refer to the UNESCO/IOC documentation on standard units of oceanographic parameters for details).

Only if these are not available or reasonable to apply other units could be used. However, in such cases they shall be comprehensively described in the file header respectively the meta data.

Data may be reported as real numbers with or without the decimal, or as integers with the number of decimals indicated. Scientific notation maybe used but if so great care must be taken to ensure that there is no ambiguity in interpretation. Reliable precision should be preserved in the data set.

Data values may be in the same file as the header information or may be stored in separate files accordingly referenced to the compatible header file. In this latter case, a clear and unique link must be provided between the header and data values (see also comments regarding station numbering below).

Some commercially available computer applications use a comma (ASCII 44) as a delimiter for separating fields. Some countries use commas to indicate the position of the decimal. If data are exchanged in comma delimited files, care must be taken to ensure there happens no confusion with the use of commas as decimal points. Generally, the use of a period to indicate the decimal is preferred since there can be no confusion.

7.3.6.4 Date / date stamp

This conveys information about the year, month, and day on which the data were collected. A great many variations are possible and are in use. Generally, numeric schemes are preferable to avoid confusion caused by language differences. A date in the form YYYYMMDD is preferred where YYYY is the four digit year, MM the two digit month and DD the two digit day. Since oceanographic data collections extend back more than 100 years it is extremely important to use a four digit year. If a format other than that described above is used, then it should be fully described in the file header.

Another common practice is to use either day number from the start of the year, or Julian day. Julian day is not the same as day number since the former is the number of days measured from noon, 1 January, 4713 BC. A clear description must be given if either of these forms is used.

7.3.6.5 Time / time stamp

Time shall be reported as hours, minutes and seconds or decimal equivalents, in UTC (GMT) and if not, the time zone must be indicated. The 24 hour clock shall be used and all local time stamps must be converted to UTC.



7.3.6.6 Instrumentation

Information about what instrument was used to make the observation should be included with the data. If this is not possible some other means of identifying the method by which the observation was made should be included.

7.3.6.7 Missing Values

The use of missing value indicators is strongly recommended. There are many ways to indicate a value is missing but care must be taken to ensure that they cannot be confused with valid data. Indicators must be clearly described in the file header.

7.3.6.8 Contact person(s)

At least one contact person shall be named in plain text and is useful in providing a contact for questions about the data collection. The full name and address of the contact person is useful to include with the data collection. This normally may be supplied as plain text. If feasible more contact persons may be named (e.g. discriminated in chief scientist, operator and/or evaluator).

7.3.6.9 Position

An indication of precision and/or the equipment used to establish position should be described in the file header (e.g. GPS, DGPS, etc.). Latitude and longitude coordinates must be specified. Geographical descriptions such as place names and bearings from locations should not be used by themselves (although they may also be included). There are a number of ways of reporting latitude and longitude and care must be taken to indicate the precision of the position when known.

Regardless of how latitude and longitude are specified, it must be clear where its origin is, and how a position is measured with respect to this origin. It is common to refer to latitudes north of the equator as positive and those south as negative. Some groups report longitudes from 0 to 360 degrees while others use +/- 180 degrees measured from Greenwich. Although more difficult to parse and transfer into application programs it could be advisable to use N, S, E and W designators since there can be no doubt once the origin is specified. One should not use equivalent designators from other languages as this can be confusing.



7.3.6.10 Station number

The station number is often assigned incrementally for consecutive stations sampled by the originator. Changes in either space or time should cause the station number to be incremented. As well, stations sampled at ocean weather stations, for example, may be labelled specially.

If applicable, this is an important parameter to include. Very often it is used by a data centre in combination with the cruise number as a unique identifier of a station. If data values are split between files, the unique combination of cruise number and station number is very important in allowing data collections to be properly recombined. Use of unique station identifiers is strongly advised.

7.3.6.11 Data quality

Data originators should give some indication of their own evaluation of the data quality. Originators should indicate whether and which calibration measures have been applied to the data.

Originators may use quality flags to indicate the degree of confidence in the associated variables. If quality flags are used then the quality checking procedures should be documented in detail.

If these procedures are not documented then the flags are likely to be of no use to secondary users.

There exist many schemes used to flag data quality. In principal all are acceptable provided the interpretation of the flag is well documented.

FerryBox data shall be quality controlled as for their level in use which implies different stages for data used in real- or quasi-real-time. There is certainly a quality difference for real-time or quasi-real-time data which are for instance applied for hazard prediction or warning purposes, and finally processed and quality controlled data which can be subsequently used for tasks like monitoring of the oceans or identification of trends in the climate. The latter do require a much higher degree and level of quality checking.

Quality flagging can be applied globally to a particular data set as well as individually to each numerical value.

One of many data quality control flag schemes employed is illustrated here. It uses a one character field with the following interpretation:

- 0 = data are not checked**
- 1 = data are checked and appear correct**
- 2 = data are checked and appear inconsistent but are considered as correct**
- 3 = data are checked and appear doubtful**
- 4 = data are checked and appear to be wrong**
- 5 = data are checked and the value has been altered**



Other QC flagging schemes use less flags (e.g. 3 indicating good and doubtful data and a third flag for missing values).

In the case that data values are altered as a result of quality control practices, it is strongly advised that the original values are preserved in history-like data sets.

7.3.6.12 Remotely sensed data, images, videos and movies

Satellite- or air born images and movie files should be delivered in formats which are common among most used operating systems (e.g. TIFF, GIF, BMP, AVI).

Application of image or video data reduction measures should be avoided in case the images or video sequences are provided for subsequent evaluation.

Images and videos shall be generally referenced at least with a geographical position relative to a reference point as well as with a date and time stamp.

A brief description of image content and nature should also accompany each image or video file as well.

7.4 Site-related data documentation

References or complementary material to site related documentation shall be accordingly given – at least references to respective documentation elaborated in the projects.

Further details and requirements in this conjunctions will be specified.



8 References

The following references give certain helpful hints and additional information to data management in general and documentation of data in particular.

Relevant guidelines and procedures for management of oceanographic data including references to guidelines of ICES, WOCE 1/2, EU MAST 2/3, U.S. JGOFS are given in

<http://ioc.unesco.org/oceanteacher/resourcekit/Module1/DataPolicy/>.

The International Council for the Exploration of the Sea (ICES) provides guidelines and standards for oceanographic data preparation, documentation and exchange at.

<http://www.ices.dk/ocean/procedures/>.

More and partially very comprehensive information on management of oceanographic data can be found e.g. on the websites and digital documentation stores of

- [the Intergovernmental Oceanographic Commission \(IOC\)](#),
- [the British Oceanographic Data Centre \(BODC\)](#),
- [the French National Oceanographic Data Centre \(SISMER\)](#),
- [the United States National Oceanic and Atmospheric Administration \(NOAA\)](#).