



GRANT N°: 871153
PROJECT ACRONYME : JERICO-S3
PROJECT NAME : Joint European Research Infrastructure for Coastal Observatories - Science, services, sustainability
COORDINATOR : Laurent DELAUNEY - Ifremer, France - jerico-s3@ifremer.fr

JERICO-S3 MILESTONE	
Joint European Research Infrastructure network for Coastal Observatory Science, Services, Sustainability	
MS#, WP# and full title	JERICO-S3 MS.35 - Gap analysis of Best Practices draft
5 Key words	data management, gap analysis,
Lead beneficiary	
Lead Author	Miguel Charcos Llorens
Co-authors	Pauline Simpson, Peter Thijssse, Alain Lefebvre, Leonidas Perivoliotis, Juan G. Fernández, Joaquín Tintoré
Contributors	
Submission date	April 28th 2021

→ **Please specify the type of milestone:**

- Report after a workshop or a meeting (TEMPLATE A)
- Report after a specific action (TEMPLATE B) (test, diagnostic, implementation,...)
- Document (TEMPLATE B) (guidelines,...)
- Other (TEMPLATE B) (to specify)

Diffusion list			
<u>Consortium beneficiaries</u>	Third parties	Associated Partners	other

PROPRIETARY RIGHTS STATEMENT

THIS DOCUMENT CONTAINS INFORMATION, WHICH IS PROPRIETARY TO THE **JERICO-S3** CONSORTIUM. NEITHER THIS DOCUMENT NOR THE INFORMATION CONTAINED HEREIN SHALL BE USED, DUPLICATED OR COMMUNICATED EXCEPT WITH THE PRIOR WRITTEN CONSENT OF THE **JERICO-S3** COORDINATOR.

*According to the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) and the 78-17 modified law of 6 January 1978, you have a right of access, rectification, erasure of your personal data and a right of restriction to the data processing. You can exercise your rights before the Ifremer data protection officer by mail at the following address: IFREMER – Délégué à la protection des données- Centre Bretagne – ZI de la Pointe du Diable – CS 10070 – 29280 Plouzané - FRANCE or by email: dpo@ifremer.fr // jerico@ifremer.fr
 Ifremer shall not hold your personal data for longer than necessary with regard to the purpose of the data processing and shall destroy it thereafter.*



TABLE OF CONTENT

Objectives	3
Implementation process	3
Main report	3
Conclusion	3
Synthesis of main conclusion	3
Next steps (work plan)	3
Annexes and references	4

1. Objectives

The main objective of this Milestone is to provide the steps on how to deal with the Gap analysis of best practices draft (D6.6).

2. Implementation process

- Identify all the best practices related to data flow within the Jerico Research Infrastructure (Jerico-FP7, Jerico-Next and Jerico S3), pointing out the date of the document, the link, and if it is already included in the OBPS system. BestPractices_Deliverables
- Share the documents identified with experts of Ocean Best Practices System and WP5 in order to verify the relevance of the documents taken into consideration.
- Share the information with all partners under WP6 and WP5.
- Read the best practices that were identified and identify various aspects related to the observing platform, the observed variables and the steps in the data management process that are described.
- Create a summary table with all the information as the foundation to describe the gap analysis

3. Main report

[T6.4.1 Gap analysis Draft](#)

(c.f. Annex 2. The document is also available in the JERICO-S3 archives)

4. Conclusion

4.1. Synthesis of main conclusion

- The processes of the HFR are the most complete of the four mature platforms that we are considering in this analysis.
- The best practices provided by JERICO about data management of data from gliders, ferryboxes and moorings are not sufficient to describe the management of the entire data cycle.
- Recommendations on how to evolve new best practices for glider and ferryboxes:
 - Follow the guidance of the Eurogoos white paper to be taken into account in order to comply with the European vision.
 - The glider best practice should align with the current efforts of the Ocean Glider community in data management as well as data and metadata formats and standards.
 - The handbook from fixO3 could be used as a starting point for the mooring best practices.
 - The development of best practices for gliders, ferryboxes and moorings should take into account the current practices of the PSS and IRS.

4.2. Next steps (work plan)

- The best practices from PSS's and IRS's for each of these observation systems will be gathered and shared with Task 6.2.
- Use inputs from regions to guide partners on how to include their best practices on the OBPS system.

4.3. Annexes and references

4.3.1: Annex 1: [BestPractices_Collection_WP6-T6.4.1](#)

This excel file is too large to be printed as an annex, but is available as a file in the JERICO-S3 archives.

4.3.1: Annex 2: [T6.4.1 Gap analysis Draft](#)

DATA MANAGEMENT BEST PRACTICES ANALYSIS

WP6 of JERICO-S3 is intended to support data management activities of the regional sites in the JERICO-S3 project (in WP3 and 4) and possible other research infrastructures. For this purpose, it is critical to provide best practices (BP) for data management that define the whole data flow from data acquisition to dissemination. We can differentiate various levels of maturity of platforms and parameters. On one hand, T6.2 addresses the platforms of higher data management maturity, High-Frequency Radars (HFR), Gliders and Ferryboxes with physical and biogeochemical (BGC) parameters. On the other hand, T6.3 focuses on less mature communities around biological imagery, biological sensors and carbonate systems. For several of these platforms and parameters, previous JERICO projects have contributed to best practices defining the operations and data flows, and JERICO-S3 continues on this path with an analysis of the existing and working via experts on drafting new BP's.

The goal of this analysis is to identify the existing JERICO best practices, to analyze their status and to recognize possible gaps. This document presents the methodology and results of the gap analysis. [Section 1](#) describes the ideal scenario of the best practices related to data management based on the WP6 outputs on data management. For this purpose, we outline the target descriptions that we expect from a best practice providing a complete set of the data management of mature platforms. [Section 2](#) provides an overview of the best practices covering various types of platforms. This overview focuses on JERICO and JERICO-NEXT deliverables. Although this document aims to cover mature platforms, we have also included some of the results around biological sampling in order to support the gap analysis by the expert panel in T6.3 for less mature platforms. [Section 3](#) provides a gap analysis for the mature platforms with physical and BGC parameters. It summarizes gaps that exist between the current situation summarized in [section 2](#) and the ideal state described in [section 1](#). [Section 3](#) also reports recommendations for harmonizing the existing best practices in order to cover the whole data flow for each mature platform.

1. Data Management: Future State

In order to allow harmonization of the best practices from a single platform perspective, covering the whole data flow from acquisition to dissemination, it is important that the best practices describe the process for the entire data cycle that is defined in the data management plan of JERICO-S3 (D6.1). This means that ideally for each platform we need a guideline that describes each step of the process.

As described in the data management plan of JERICO-S3 ([D6.1](#)), these JERICO best practices should be adopted by partners as data management guidelines to guarantee access to FAIR (Findable Accessible Interoperable Reusable) data through the major European data providers,

which is only possible via FAIRness in the whole dataflow. We used the data management document and the deliverables of previous JERICO projects to first identify the important steps of the data flow. The major processes involved in the steps of the data cycle are as follows:

- Data acquisition
- Data transfer
- Data pre-processing
- Data processing
- Data post-processing
- Data archiving
- Data dissemination
- Data publication and visualization
- Quality control
- Data flow to aggregators

In addition, we also identified references to aspects of the data management such as,

- Data policies
- Data standards
- Data formats
- Metadata descriptions.

The list of available platforms in the networks along the European coastal line, and the expected variables observed from them, is described in the data management plan of JERICO-S3 ([D6.1](#)) for physical, biogeochemical and biological data.

2. Analysis of the Current State of Best Practices

As a start of this task we have created an overview of the existing JERICO best practices related to data management. [Appendix A](#) summarizes the list of best practices related to data management that were identified among the deliverables of JERICO and JERICO-NEXT. It includes the links to the documents in the JERICO-RI website as well as the IOC Ocean Best Practices System (OBPS) repository. We present in this section the way the management of data from mature platforms is described by best practices. The identification numbers in [Appendix A](#) are used between “[]” brackets in the analysis to refer to specific documents. The study of these documents was performed considering the following criteria:

- The steps of the data management process that were referred to in the document.
- The platforms to which these best practices apply when available. Some of them were focused on specific sensors and did not apply to a specific platform.
- The sensor types or variables to which the best practice apply.

Concerning mature platforms, we identified 15 documents related to data management best practices. [Table 1](#) addresses the steps of the data cycle for the platforms HFR, gliders,

ferryboxes and moorings. [Table 2](#) explores the descriptions of the documents related to data policies, standards, formats and metadata description.

Table 1: Summary of the JERICO best practices references to the data management steps. Identification numbers of the documents (see Appendix A) are used to indicate that the document describes a specific step of the data cycle for a given platform.				
	Glider	Ferrybox	HFR	Mooring
Data flow definition	4 (overview)	1 (RT), 4 (overview), 7	4 (overview), 13 (all flow)	4 (overview)
Data Acquisition			13	
Data Transfer	4 (overview)	1, 4 (overview), 7	4 (overview), 13	4 (overview)
Data Pre-processing		1	13	
Data Processing			3 (preliminary), 6 (algorithm improvements), 11, 13	
Data Post-processing	10 (DM corrections),		3 (preliminary), (time aggregation strategy), 12, 13	
Quality Control	4 (overview), 9 (BGC QC), 10 (DM QC)	1, 4 (overview), 7, 9 (BGC QC)	3 (preliminary), 4 (overview), 6 (algorithm improvements), 11 (totals and radials), 12, 13	4 (overview), 9 (BGC QC)
Data archiving	5	1, 5	5, 13	5
Data Dissemination	4 (overview), 5	4 (overview), 5, 7	3 (preliminary), 4 (overview), 5, 11 (standard definitions), 13	4 (overview), 5
Data Publication and Visualization	4 (overview)	1, 4 (overview)	4 (overview), 13	4 (overview)
Data flow to aggregators	5	5, 7	5, 13	5

Although one document [[AppendixA: 13](#)] talks about cabled observatories, it focuses on the operational side of the platform and it was not included in this section. All of the best practices that we identified in this section were present in OBPS. The overview of the existing best practices is as follows:

- Two documents apply to all platforms and describe general practices on data management practices [[AppendixA: 4](#)] and data access policies [[AppendixA: 5](#)]
- One document [[AppendixA: 9](#)] applies to BGC QC of gliders, ferryboxes and moorings.
- Eight documents apply specifically to individual platforms
 - Five related to HFR [[AppendixA: 3, 6, 11, 12, 13](#)] which are focused on sea surface velocities
 - One related to gliders [[AppendixA: 10](#)] which is focused on water temperature and salinity
 - Two related to ferrybox [[AppendixA: 1, 7](#)] which is focused on water temperature and salinity as well as BGC variables as Chl, oxygen and turbidity

Table 2: Summary of the JERICO best practices references to data and metadata descriptions. Identification numbers of the documents (see [Appendix A](#)) are used to indicate that the document describes the data or metadata properties for a given platform.

	glider	ferrybox	HFR	Mooring
Data policies	5	5	5, 13	5
Data standards	4 (overview)	1, 4 (overview)	3, 4 (overview), 11 (for processing), 13	4 (overview)
Data formats	4 (overview)	1, 4 (overview)	3, 4 (overview), 11 (for processing), 13	4 (overview)
Metadata description	5, 10 (Corrected products)	5	3, 5, 11 (for processing), 13	5

We did not include in this section the results related to biological data. However, these results are summarized in [Appendix B](#) for future reference for the analysis of T6.3. Among the documents that we analyzed concerning less mature platforms, we recognized four documents [[Appendix A: 2, 8, 14, 15](#)] related to biological data. Two documents [[Appendix A: 8, 14](#)] are specifically related to phytoplankton and describe the standards and the formats of the data.

3. Gap analysis and recommendations

The processes of the **HFR** are the most complete of the four mature platforms that we are considering in this analysis. We identified a considerable amount of literature about data management practices of HFR which are the result of a progressive work leading to the creation of the report on Best Practice for new HFR network systems [[AppendixA: Part 1 of 13](#)]. This set of documents covers the whole data flow from acquisition to dissemination. The data, metadata and QC models are well described in document [[AppendixA: 11](#)] as considered by the European HFR Node, SeaDataNet and CMEMS. The representatives from the different regions should be contacted in order to promote the use of these guidelines. Adopting the process of HFR data management will help to learn from the different partners in order to update the existing procedures.

The best practices provided by JERICO about data management of data from **gliders**, **ferryboxes** and **moorings** are not sufficient to describe the management of the entire data cycle. Documents [[AppendixA: 4](#)] and [[AppendixA: 5](#)] outline some general recommendations on data management and policy but it is necessary to create specific descriptions of the process for these three platforms. These descriptions are critical for the alignment of the processes of all the partners. Some of the processes for ferrybox data are sketched in document [[AppendixA: 7](#)] and they can be considered as a starting point to expand the current information into a complete procedure of the data management of data from these platforms. In this process, it is recommended to follow the guidance of the [Eurogoos white paper](#) be taken into account in order to comply with the European vision.

Concerning **gliders**, the only contribution [[AppendixA: 10](#)] that was found among JERICO best practices relates to advanced delayed mode products. It is also recommended to elaborate a complete procedure for the management of glider data. This best practice should align with the [current efforts](#) of the Ocean Glider community in [data management](#) as well as data and metadata formats and standards.

We did not identify best practices specific to the data management of **mooring platforms**. Although moorings are not part of the platforms of T6.2 activities to support data management, it is clear that there is a lack of the description of this process. The [handbook from fixO3](#) could be used as a starting point for this purpose.

The development of best practices for gliders, ferryboxes and moorings should take into account the current practices of the PSS and IRS. It should compile the existing procedures into a common document that helps to standardize the processes. In this process, these regions must be contacted to collect the information as an input for the data management best practices under T6.2 with the support of T6.4.1 through a data management committee with representatives from the PSS and IRS's involved in defining and applying best practices.

APPENDIX A: Overview of JERICO Best Practices

Table A1: List of JERICO deliverable related to data management best practices. Column ID provides an identification number which is used as a reference of the best practice in this document. Column D# contains the deliverable name and version.						
ID	Title	Publication date	Project	D#	Deliverable link	OBPS repository link
1	Report on current status of Ferrybox.	2014-02	JERICO-FP7	D3.1 Version 1.7	https://www.jerico-ri.eu/download/filebase/jerico_fp7/deliverables/D3.1%20Ferrybox%20best%20practices%20v1.7.pdf	http://hdl.handle.net/11329/317
2	Marine biological data: quality control and management practices.	2017-02	JERICO-NEXT	D5.4 Version 4.0	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable-5.4.pdf	http://dx.doi.org/10.25607/OBP-217
3	Recommendation Report 1 for HFR data implementation in European marine data infrastructures	2017-03	JERICO-NEXT	D5.13 Version 1.0	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable-5.13_V1.pdf	http://dx.doi.org/10.25607/OBP-943
4	Report on data management best practice and Generic Data and Metadata models.	2017-07	JERICO-NEXT	D5.9 Version 2.1	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable_5.9_v2.1.pdf	http://hdl.handle.net/11329/354
5	Recommendation on a free and open data access policy	2017-09	JERICO-NEXT	D5.1 Version 1.3	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT_Deliverable_5.1_v1.3.pdf	http://dx.doi.org/10.25607/OBP-950
6	Report on first methodological improvements on retrieval algorithms and HF radar network design.	2017-09	JERICO-NEXT	D3.3 Version 1.0	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable_3.3_v1.0.pdf	http://dx.doi.org/10.25607/OBP-947

7	Specifications for a European FerryBox data management system	2017-09	JERICO-NEXT	D5.3 Version 1.1	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT_Deliverable_5.3_v1.1.pdf	http://dx.doi.org/10.25607/OBP-216
8	Novel methods for automated in situ observations of phytoplankton diversity and productivity: synthesis of exploration, intercomparisons and improvements.	2017-10	JERICO-NEXT	D3.2 Version 5	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT_Deliverable_3.2_130819_V5.pdf	http://dx.doi.org/10.25607/OBP-219
9	Best practices for quality control of sensor based biochemical data.	2017-11	JERICO-NEXT	D5.11 Version 1.3	http://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT_Deliverable-5.11_v1.3.pdf	http://dx.doi.org/10.25607/OBP-16
10	Guidelines for the delayed mode scientific correction of glider data	2018-09	JERICO-NEXT	D5.15 Version 4.1	http://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable_5.15_Final.pdf	http://dx.doi.org/10.25607/OBP-430
11	Recommendation Report 2 on improved common procedures for HFR QC analysis. JERICO-NEXT WP5-Data Management,	2018-10	JERICO-NEXT	D5.14 Version 1.0	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT-Deliverable_5.14_V1.pdf	http://dx.doi.org/10.25607/OBP-944
12	Report on final assessment of methodological improvements and testing on infrastructures.	2019-07	JERICO-NEXT	D3.4 Version 2.0	http://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT_Deliverable_3.4_180719_final.pdf	http://dx.doi.org/10.25607/OBP-948
13	Report on Best Practice for new network systems: Part 1: HF-radar ; Part 2: cabled coastal observatories.	2019-07	JERICO-NEXT	D2.4 Version 1.0	http://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO_NEXT_Deliverable_2.4_final.pdf	http://dx.doi.org/10.25607/OBP-1005
14	Report on the technical and analytical improvements of	2019-08	JERICO-NEXT	D3.2 Version 5.0	http://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO-NEXT_Deliverable_3.2_180719_final.pdf	http://dx.doi.org/10.25607/OBP-945

	innovative techniques and recommendations on their use.				able_3.2_130819_V5.pdf	
15	Document describing the biological data.	2019-08	JERICO-NEXT	D5.5 Version 3.0	https://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO_NEXT_Deliverable_5.5_final.pdf	http://hdl.handle.net/11329/1548
16	Report on Best Practice in the utilization of sensors used for measuring nutrients, biology related optical properties, variables of the marine carbonate system, and for coastal profiling.	2019-09	JERICO-NEXT	D2.5 Version 1.0	http://www.jerico-ri.eu/download/jerico-next-deliverables/JERICO_NEXT%20_Deliverable%20_2.5_Version%201.0.pdf	http://dx.doi.org/10.25607/OBP-898

APPENDIX B: Summary of best practices related to biological data

Table B1: Summary of the JERICO best practices references to the data management steps for biological data. Identification numbers of the documents (see [Appendix A](#)) are used to indicate that the document describes a specific step of the data cycle for a given platform.

	Biological data in general (includes phytoplankton) [Appendix A:2, 15, 16]	Phytoplankton [Appendix A:8, 14]
Sample source	Nets or grabs, water samples	Ferrybox, mooring, bottle sample, CTD rosette frame mounted sensors
Data flow definition	RT & DT	partially
Data Acquisition		
Data Transfer	2	
Data Pre-processing		8
Data Processing		8, 14
Data Post-processing		8
Quality Control	2	8
Data archiving	2, 15	8
Data Dissemination	2, 15	
Data Publication and Visualization	2, 15	8
Data flow to aggregators	2	
Data policies		
Data standards		
Data formats		8
Metadata description	15, 16	8