

ASSEMBLE



ASSOCIATION OF EUROPEAN MARINE BIOLOGICAL LABORATORIES EXPANDED

WP4 - NA2 Improving virtual access to marine biological stations data, information and knowledge

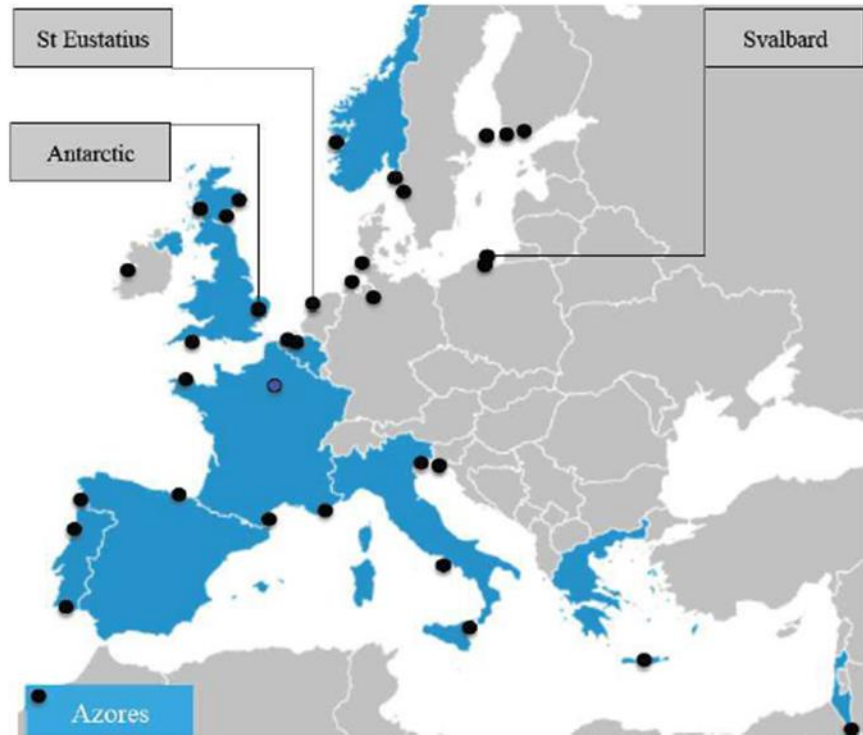
AssemblePlus General Assembly

10-12th October 2018



Objectives

- Improve virtual access to marine biological stations data, information and knowledge by:



- Shaping and Assemble Plus community **data management plan**
- Improving **interoperability** with related e-infrastructures
- Installing an **on-line system** to provide access to marine biological stations **information**
- Improving **virtual access to data resources** from Assemble Plus stations, with a specific focus on **genomic** and **long term biodiversity** data series
- Offering a **virtual analysis platform** as a service for genomic and biodiversity data analysis

Work package number	NA2	Lead beneficiary				VLIZ	
Work package title	Improving virtual access to marine biological stations data, information and knowledge						
Participant number	17	10	20	16	8	1	12
Short name of participant	VLIZ	HCMR	MBA	MPIMM	UGOT	UPMC	SZN
Person/months per participant:	34.5	18	8	9	4.5	4.5	3
Participant number	14	9	4	11	15	18	21
Short name of participant	CCMAR	UPV/EHU	UH	HUJI	AWI	SAMS	NERC-BAS
Person/months per participant:	3	2	1.5	1.5	1.5	1.5	1.5
Participant number	23	2	5	3	6	7	19
Short name of participant	MSS	NIB	IOPAN	NIOZ	UG	NUIG	USTAN

Person/months per participant NA2.3: **Multidisciplinary datasets inventory and archival:** UPMC(x3), NIB, NIOZ, UH, IOPAN, UG, NUIG, UGOT, UVIGO, EHU, HCMR, HUJI, SZN(x2), UiB, CCMAR (x2), AWI, MPIMM, VLIZ, SAMS, USTAN, MBA, NERC, MSS

Objectives NA2.4: **LT biodiversity data series & genomic data:** UPMC, UH, EHU, HCMR, HUJI, SZN, CCMAR, AWI, VLIZ, SAMS, MBA, NERC, MSS

Improve virtual access to marine biological stations data, information and knowledge

- Shaping and developing a virtual analysis platform for genomic and biodiversity data analysis;
- Improving interoperability with related e-infrastructures;
- Installing an on-line system to provide access to marine biological stations information;
- Improving virtual access to data resources from ASSEMBLE Plus stations, with a specific focus on genomic and long term biodiversity data series;
- Offering a virtual analysis platform as a service for genomic and biodiversity data analysis.

Task NA2.1. Design Data Management Plan and interoperability with related e-infrastructures

Task Leader: VLIZ; **participants** HCMR, MBA, MPI.

Deliverables: D NA2.1a,b, c: **Data Management Plan (M6, M24, M48)**

A data management plan (DMP) will be drafted as a reference document for the ASSEMBLE Plus community. This document will outline how the **research data** collected or generated within the **JRAs and as part of the TA** will be handled during the ASSEMBLE Plus project and will remain in force after the end of the project.

The DMP will build on current best practices at the national, EU and international levels, and recommendations and guidance from the ICSU World Data System (ICSU-WDS), Research Data Alliance (RDA) and the International Oceanographic Data Exchange (IODE). The DMP will define the methodology and standards to be applied, whether and how these data will be shared and/or made open, and how they will be curated and preserved.

What is a DMP

- **What is a data management plan?**
 - Document defining what data will be generated, how data will be curated and preserved and whether and how data will be shared or made available.
- **Why do we need it?**
 - H2020 requirements
 - Essential part of planning research activities



What is a DMP

- **Elements of a Data Management Plan**

- Framework and scope
- Data Collection
- Documentation and Metadata
- Storage and Backup
- Selection and Preservation
- Data Access, Sharing and Rights
- Responsibilities and Resources





Workshop: The ASSEMBLE Plus Data Management Plan, 1-2 March 2018, Oostende, Belgium *by Klaas Deneudt*

On 1 and 2 March 2018, the **ASSEMBLE Plus** Data Management Plan Workshop was held at the Flanders Marine Institute (VLIZ), Oostende. The aim of this workshop was to discuss the Data Management Plan (DMP), the first version of which was completed by 30 March.

Participants included partners of Work Package 4 – Networking Activity 2 (NA2), “Improving virtual access to marine biological stations data, information, and knowledge”, and the Work Package leaders of the Joint Research Activities (JRAs).

Experts from related initiatives and partnerships were also invited: LifeWatch, ELIXIR, European Marine Observation and Data Network Biology (EMODNet Biology), the European Marine Biological Resource Centre Working Group on E-Infrastructure (EMBRC WGEI), Digital Curation Centre (DCC), International Oceanographic Data Exchange network (IODE) and Ocean Biogeographic Information System (OBIS).

The meeting commenced with presentations framing the **ASSEMBLE Plus** project and NA2. Specific focus was placed on the design of the DMP and its interoperability with related e-infrastructure (NA2.1). A presentation followed on what Horizon 2020 “Open Research Pilot” data means, and what its Findable, Accessible, Interoperable and Reusable (FAIR) data policy encompasses. The DMP of EMBRC was presented

as a side-by-side comparison to the information required for the **ASSEMBLE Plus** DMP.

A discussion on the most critical points of the DMP was carried out, including: the types of data expected, data file sizes, where data will be stored, open access (or not), and clarifications on the EU template, along with four JRA presentations. The management of TA project data was discussed, as was the format of the spreadsheet that the JRAs need to fill in with information that will allow an **ASSEMBLE Plus** data inventory to be created.

On the second day the discussion points continued, and JRA representatives filled in a template form to provide specific information that is required for the DMP. Overall, the meeting was successful due to the gathering of information for the DMP.

Important future action points will be the management of very large data files, the Open Access and the FAIR policy in relation to TA project data, whether more harmonisation (file types, databases used, naming conventions etc.) is possible, and the storage of **ASSEMBLE Plus** data products in the first Flemish marine data bank (IMIS) and the Marine Data Archive.

Deliverable D4.1



Acronym: ASSEMBLE Plus

Title: Association of European Marine Biological Laboratories Expanded

Grant Agreement: 730984

Deliverable D4.1

ASSEMBLE Plus First Data Management Plan

03/2018

Lead parties for Deliverable: VLIZ

Due date of deliverable: M6

Actual submission date: 06/04/2018

All rights reserved

This document may not be copied, reproduced or modified in whole or in part for any purpose without the written permission from the ASSEMBLE Plus Consortium. In addition to such written permission to copy, reproduce or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright must be clearly referenced.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730984. This output reflects the views only of the author(s), and the European Union cannot be held responsible for any use which may be made of the information contained therein.



Deliverable D4.1
ASSEMBLE Plus Data Management Plan

1. Introduction.....	5
1.1 The ASSEMBLE Plus project.....	5
1.2 The DMP meeting.....	6
1.3 The ASSEMBLE Plus data collection.....	6
1.4 The ASSEMBLE Plus data origins	8
1.5 Other users of ASSEMBLE Plus data	8
2. FAIR Data	10
2.1 Making data findable.....	10
2.1.1 Metadata	10
2.1.2 File naming	11
2.2 Making data openly accessible.....	12
2.2.1 Open Access	12
2.2.2 Storage.....	13
2.2.3 Data access	13
2.3 Making data interoperable.....	14
2.3.1 File formats.....	14
2.3.2 Data formatting and vocabulary standards	14
2.4 Making data re-useable.....	16
2.4.1 Quality control.....	16
2.4.2 Long-term re-use	17
2.4.3 Data-creation standards	18
3. Allocation of Resources	19
3.1 Costs for making data FAIR.....	19
3.2 Responsibilities.....	19
4. Data Security	21
5. Ethical Aspects.....	23
ANNEX 1 Best-practice guidelines	24
ANNEX 2 Tables.....	26

Data Management Plan for the Transnational Access programme

DMP's for TNA's (suggested H2020 template)

- 1. Data Summary**
 - Explanation of the data source and format
- 2. How will you make your data FAIR** (findable, accessible, inter-operable, re-useable)?
 - Discoverability, meta data, keywords, versioning
- 3. What Access** (Open, Partial, Closed) **will your data have?**
 - If some of the data is not OA, explain why and how access could be granted
 - What software is required to access them?
- 4. How will you make your data inter-operable?**
 - Vocabularies used for the meta data
- 5. How will the data be made re-usable?**
 - License, embargo periods, QA processes, how useable is the data by other parties?
- 6. Allocation of costs and Data security**
 - Mostly a formality
- 7. Ethics**
 - Could be taken from their TA application
- 8. Publications**
 - In an OA journal?

Approach

- From Call 2 onwards:
 - A short DMP section is part of the application form they fill in
 - A fuller DMP will be sent to each accepted applicant after they have signed their UAC
- Call 1: retrospective request after the Activity Report (13 sent so far, 3 returned)



TA Data Management Plan

User name:
Project title:
Proposal code:
Date:



1. Data Summary
Questions 1.1. State the purpose of the data collection/generation 1.2. Explain the relation to the objectives of the project 1.3. Specify the types and formats of data generated/collected 1.4. Specify if existing data is being re-used (if any) 1.5. Specify the origin of the data 1.6. State the expected size of the data (if known) 1.7. Outline the data utility: to whom will it be useful
Suggested responses (copy below where applicable): 1.1, 1.2) The answers can be taken from your original ASSEMBLE Plus proposal. Focus on the aspects that specifically answer the questions, but remember that this document must be self-contained 1.3) Where do the data come from? What file formats will they be? 1.4) Yes and briefly why/No. 1.5) The origin of the data are the experiment. The origin of the re-used data is the work of 1.6) For the entire data collection; does not have to be exact ("of the order X Mb") 1.7) These data will be useful to those working in marine conservation/the blue economy/xx research .
Responses:
2. FAIR Data 2.1 FAIR – Making data findable, including provisions for metadata
Questions 2.1.1. Outline the discoverability of data (metadata provision) 2.1.2. Outline the identifiability of data and refer to standard identification mechanisms. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? 2.1.3. Outline the naming conventions used 2.1.4. Outline the approach towards search keywords 2.1.5. Outline the approach for clear versioning 2.1.6. Specify the standards for metadata creation (if any). If there are no standards in your discipline, describe what type of metadata will be created and how
Suggested responses (copy below where applicable): 2.1.1) Data will be made discoverable by describing the generated datasets in the ASSEMBLE Plus/EMBRC dataset catalogue of the Marine Data Archive. 2.1.2) In this catalogue every dataset receives a unique identifier. There is possibility to add a DOI for the completed dataset. 2.1.3) Assigned names for datasets are expected to include topical, spatial and temporal scope of the observations or experiments. 2.1.4) Datasets will be made retrievable by annotation with free and ASFA thesaurus keywords.

Task NA2.2 Creation of the Knowledge Transfer Platform

Task Leader: VLIZ.

An online Knowledge Transfer Platform will be implemented as part of the ASSEMBLE Plus web portal. This online system will become a central tool for provision with relevant stand exchange of information of researchers and research groups, their expertise and scientific output in terms of publications keholder groups. The online system will include two modules:

Knowledge Outputs module will provide the infrastructure to provide users with access to an unrestricted, user friendly and searchable database of Knowledge Outputs which may be relevant to them, as provided for by NA3. Outputs from TA and from research developed at each marine station will be submitted through a customized form which includes

Publications module will be integrated as a publication catalogue and archive in the web portal. This module will serve as an open archive for publications resulting from the ASSEMBLE Plus project. The use of the Knowledge Transfer Platform will be registered and reported using in-built tools ([D NA2.4](#)).

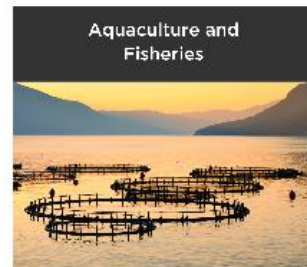
Knowledge Transfer Platform

Explore the knowledge outputs of ASSEMBLE Plus

The knowledge outputs of ASSEMBLE Plus can be accessed via the Knowledge Transfer Platform. These knowledge outputs arise from all ASSEMBLE Plus activities – the research of our [JRA participants](#), research results arising from ASSEMBLE Plus collaborations, publications arising from the [Access](#) programme, or from associated activities, and so on. These knowledge outputs are organised by the sectors listed below, and within each sector they are categorised as:

- **publications:** books, scientific publications, technical documentation, protocols..
- **data:** datasets created by research conducted within ASSEMBLE Plus
- **activities:** training activities and material, learning modules, ASSEMBLE Plus science meeting outputs...

Click on the links below to access these knowledge outputs



- [Aquaculture](#)
- [Fisheries](#)



- [Marine technology](#)
- [Marine energy](#)



- [Scientific instrumentation](#)
- [Scientific diving](#)
- [Cryobanking](#)



- [Genomic observatories](#)
- [Functional genomics](#)



- [Climate change](#)
- [Environmental management](#)

Marine Genomics



Genomics observatories are research facilities that produce genomic-level biodiversity observations that are contextualised, localised geographically, with activities performed in compliance with international data acquisition standards, and for which genomic data acquisition is a long-term activity, monitoring the stream of genetic variations in marine ecosystems. [JRA1](#) will work on the establishment of Genomic observatories across the AssemblePlus partnership.



The Lund declaration. Europe must focus on the grand challenges of our time

(2009). The Lund declaration. Europe must focus on the grand challenges of our time. [S.n.]: Lund. 2 pp.



- Publications
- Books and reviews
- Reports
- Protocols
- Technical documents
- Fact sheets

diversity molecular seaweed bacteria virus
plankton benthos barcoding expression



- Datasets
- Software and model codes
- Maps



- Training sessions
- ASSEMBLE Plus science meetings
- Multimedia

Publication search

Sector	End User	KO Type
Marine Genomics	- All -	Scientific Publication
Year	Author	Keyword
YYYY		
		APPLY → RESET →

101 records found

1 2 3

Almada, V.C.; Almada, F.; Francisco, S.M.; Castilho, R.; Robalo, J.I. (2012). Unexpected high genetic diversity at the extreme northern geographic limit of *Taurulus bubalis* (Euphrasen, 1786). PLoS One 7(8): e44404. <https://hdl.handle.net/10.1371/journal.pone.0044404>



MORE INFO →

DOWNLOAD →

Annunziata, R.; Martinez, P.; Amone, M.I. (2013). Intact cluster and chordate-like expression of ParaHox genes in a sea star. BMC Biology 11. <https://hdl.handle.net/10.1186/1741-7007-11-68>



MORE INFO →

DOWNLOAD →

Annunziata, R.; Perillo, M.; Andrikou, C.; Cole, A.G.; Martinez, P.; Amone, M.I. (2014). Pattern and process during sea urchin gut morphogenesis: the regulatory landscape. Genesis 52(3): 251-268. <https://hdl.handle.net/10.1002/dvg.22738>

MORE INFO →

Unexpected high genetic diversity at the extreme northern geographic limit of *Taurulus bubalis* (Euphrasen, 1786)



Almada, V.C.; Almada, F.; Francisco, S.M.; Castilho, R.; Robalo, J.I. (2012). Unexpected high genetic diversity at the extreme northern geographic limit of *Taurulus bubalis* (Euphrasen, 1786). PLoS One 7(8): e44404. <https://hdl.handle.net/10.1371/journal.pone.0044404>

In: PLoS One. Public Library of Science: San Francisco. ISSN 1932-6203

Available in

[Authors](#)

Vlaams Instituut voor de Zee: Open access 306923 [[download pdf](#)]

Keywords

Biodiversity; Marine Genomics; Marine Sciences; Scientific Community; Scientific Publication; Marine

Project

[Top | Authors](#)

- Association of European marine biological laboratories

Authors

[Top](#)

- Almada, V.C.
- Almada, F.
- Francisco, S.M.
- Castilho, R.
- Robalo, J.I.

EMBRC Open Access Publications Archive

[Article 29.2 of the Model Grant Agreement](#) sets out detailed legal requirements on open access to scientific publications: under Horizon 2020, **each beneficiary must ensure open access to all peer-reviewed scientific publications** relating to its results.

1. **Self-archiving / 'green' OA's** – the author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online repository before, at the same time as, or after publication. Some publishers request that open access be granted only after an embargo period has elapsed.
2. **Open access publishing / 'gold' OA's** - an article is immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers.

Purpose

- A central location for all publications produced under the ASSEMBLE Plus and ASSEMBLE Marine programmes
- Will contain all OA publications and “self-archived”

Expected contributors

- TA users
- JRA scientists
- Potentially from other scientific activities related to A+

Task NA2.3 Set up virtual open access entry point to data resources (1/3)

Task Leader: VLIZ; **participants:** HCMR, MBA, MPI (and all other partners).

Deliverables: D NA2.2a, b, c: Virtual Access hits to ASSEMBLE Plus data resources (M24, M36, M48)

The central entry point that allows users to access common ASSEMBLE Plus data resources (from JRAs, NA and TA) will be created based on existing components.

The datasets module of the Integrated Marine Information System will be plugged into the ASSEMBLE Plus website and will allow users to query and select data resources based on a set of relevant search criteria. In order to create a digital inventory of the data resources available at the ASSEMBLE Plus stations, each of the partners will describe their datasets in a digital catalogue, using an online form. The scope, characteristics, state and accessibility of the data will be described following common standardized formats compliant to ISO19115 metadata standard. In case the data is already accessible through local online databases a web link to the existing interfaces will be included in the dataset description.

Marine biological stations dataset inventory

- In IMIS information system



Institute Acronym	Number of datasets
VLIZ	384
UGENT	201
HCMR	106
AWI	78
NERC	48
IMAR	25
MBA	22
IOPAN	15
NIB	13
UG	13
SBR	11
CCMAR	8
SAMS	7
SZN	7
IAMC	6
OOV	5
CIIMAR	3
HBS	3
OOB	3
SLC	2
UIB	2
Total	963

- After additional enquiry



Institute Acronym	Number of datasets
SZN	18
ARI	15
PiE-UPV	14
IUI	12
SAMS	6
NIOZ-CNSI	5
SLC	5
TZS	4
HCMR	3
MSS	3
SOI	2
ECIMAT	1
Total	88

- 12 stations have responded as requested
- 6 stations have responded not to have the intention to describe additional datasets
- 6 stations have not answered to the request
- 5 marine stations need more time

Dataset search

Sector	End User	KO Type
<input type="text" value="- All -"/>	<input type="text" value="- All -"/>	<input type="text" value="- All -"/>
Keyword		
<input type="text"/>	<input type="button" value="APPLY →"/>	<input type="button" value="RESET →"/>

895 records found

[1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) ... [Last](#)

Vatova A., 1935. The fishery grounds near Alexandria, II. A bottom sample taken at Alexandria, Notes and Memoirs No 9. Fisheries Research Directorate, Egypt

Citation: Tsikopoulou Irini (2016). Digitization of The fishery grounds near Alexandria. II A bottom sample taken at Alexandria. Notes and Memoirs No 9. Fisheries Research Directorate, Egypt

Report on the Danish Oceanographical Expeditions 1908-1910 to the Mediterranean and adjacent seas - Introduction

Citation: Tsompanou M, Faulwetter S, Nikolopoulou S, Mavraki D, Gerovasileiou V, Fanini L, Chatzinikolaou E (2016): Digitization of the Introduction to the Report on the Danish Oceanographical Expeditions 1908-10 to the Mediterranean and adjacent seas by J. Schmidt (1912)v1.3. Hellenic Center for Marine Research. Dataset/Samplingevent.

Steuer Ad., 1935. The fishery grounds near Alexandria. I Preliminary report. Notes and Memoirs No 8, Fisheries Research Directorate, Egypt

Citation: Irini Tsikopoulou and Stamatina Nikolopoulou (2016). Digitization of The fishery grounds near Alexandria. I Preliminary report. Notes and Memoirs No 8, Fisheries Research Directorate, 1935, Egypt

Task NA2.3 Set up virtual open access entry point to data resources (2/3)

Task Leader: VLIZ; **participants:** HCMR, MBA, MPI (and all other partners).

Deliverables: D NA2.2a, b, c: Virtual Access hits to ASSEMBLE Plus data resources (M24, M36, M48)

Data will be deposited into a central ASSEMBLE Plus data repository for archival purposes.

Data infrastructure components that are specific to the genomics observatories data and that were developed during the FP7 MicroB3 project will be revitalized and connected to the entry point to meet the GOs data requirements. This includes links to other publicly available virtual environments, services, standards registries, and data repositories (e.g. Moorea Biocode, the biocode commons, the GSC, workflows developed during MicroB3, ENA-EBI data repository, etc.). Data gathered in JRA1 (Genomics Observatories) will be developed into standardized and user friendly formats and incorporated in a “one stop shop” virtual access entry point to GO data and services developed by MicroB3. In addition to providing access to, the portal will provide information on community data policies and standards, experimental protocols/logs, and the outcomes of benchmarking exercises.



Free search

OSD 2014 Environmental



Archive

klaasdeneudt

Quarantine (2)

testfolder

EMBRC

Quarantine (0)

ASSEMBLEPLUS

JRA1 Data

JRA2 Data

JRA3 Data

Ocean Sampling Day

OSD 2014 Data

OSD 2014 Environmental Data

Raw metagenomic Data

Workable metagenomic data

OSD and MyOSD 2015 Data

OSD and MyOSD 2016 data

Files↑

Files↑	Type	Datatype	Dataprovider	Archive date
OSD1_20140621T1110Z.pdf	pdf	Documents	Marine Biological Association	2018-02-27

Archived by	Deneudt Klaas	Archive date	2018-02-27 13:34:03
Author(s)	Balestreri Cecilia (MBA)	Dataprovider(s)	Marine Biological Association
Start Year	-	End Year	-
Data type	Documents	File size	1.08 MB
Path	EMBRC/Ocean Sampling Day/OSD 2014 Data/OSD 2014 Environmental Data/		
Summary	-		



File Upload Tool

Select files

Upload

Drag and drop files here

Folder drag and drop is only available in Chrome 21+ and Opera 15+



Free search



- Archive
 - klaasdeneudt
 - Quarantine (2)
 - testfolder
 - EMBRC
 - Quarantine (0)
 - ASSEMBLEPLUS
 - JRA1 Data
 - JRA2 Data
 - JRA3 Data
 - Ocean Sampling Day
 - OSD 2014 Data
 - OSD 2014 Environmental
 - Raw metagenomic Data
 - Workable metagenomic
 - OSD and MyOSD 2015 Data
 - OSD and MyOSD 2016 data

EDIT FILE METADATA

File properties
Extra attributes
Description

Filename

File creationdate

Author(s)

Dataprovider

Email Dataprovider

Condition of Use

Datatype

Summary

Changes

Underlined items are required.

Type	Datatype	Dataprovider	Archive date
pdf	Documents	Marine Biological Association	2018-02-27

2018-02-27 13:34:03
 Marine Biological Association
 1.08 MB
 SD 2014 Environmental Data/

Summary

Task NA2.3 Set up virtual open access entry point to data resources (3/3)

Task Leader: VLIZ; **participants:** HCMR, MBA, MPI (and all other partners).

Deliverables: D NA2.2a, b, c: Virtual Access hits to ASSEMBLE Plus data resources (M24, M36, M48)

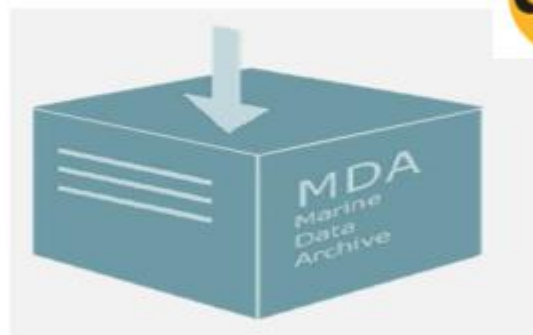
Data that can be brought into the public domain will be prepared for data publication. A technical quality control will verify that all required information is included. Each of the datasets proposed for publication will be made citable and will be labeled with a digital object identifier (DOI). All gathered information will be checked, improved and validated during a common data workshop to which all providers will be invited.

The use of the virtual open access platform will be registered and monitored using built-in tools. Data availability and “hits” will be assessed by the Advisory Board and reported as a yearly deliverable to the EC ([D-NA2.2a, b, c at M24, 36, 48](#)).

Data archiving & publication service



IMIS
Integrated Marine
Information System



Task NA2.4 Improve data access and standardization of genomic and long term marine biodiversity observation

Task leader: VLIZ with contributions from HCMR, MBA, MPI, all genomic & biodiversity observatory stations

Most marine biological stations collect and manage **historical times series of biodiversity data**. However, many valuable, historic datasets still remain inaccessible to the larger scientific community. The stations that manage these biodiversity data series and the **genomic observatory stations** will be invited to dedicated sessions during the **common data workshop** during which they will receive **training in providing access** to these types of data.

This includes documenting, annotating, archiving, quality controlling and dealing with the IPR issues through existing licensing approaches (e.g. use of moratorium periods and accreditation by producing data citations and data papers). By M18 the described data series will be made accessible by the stations to the task team for final validation and incorporation in the VA platform. A contact point of the task team will be ready to assist the stations with any data related technical issues.

Marine biological stations Long Term Biodiversity data

- Enquiry by Sidonie Gras for EMBRC: 14 responses from 7 countries

	T	S	DO	pH	Fluor	Turbidity	NO3	PO4	NO2	SiOH4	NH4	Chl a	Phytoplankton	Zooplankton	Presence	Diversity	Molecular	DN15	DC13	
Institut de la Mer de Villefranche	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Station Biologique de Roscoff	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Observatoire Océanologique de Banyuls	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Royal Belgian Institute of Natural Science	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
VLIZ - Flanders Marine Institute	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
ECIMAT - University of Vigo	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Interuniversity Institute for Marine Sciences (IUI)	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Husö Biological Station	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Tvärminne Zoological Station, University of Helsinki	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Archipelago Research Institute, University of Turku	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Stazione Zoologica Anton Dohrn	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
British Arctic Survey (BAS)	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Marine Biological Association (MBA)	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Marine Scotland Science (MSS)	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green

KEY

Green	19%	Data collected and available for download
Yellow	26%	Data collected and available upon request
Red	11%	Data collected but not available
Grey	44%	Data not collected

Need to address additional AssemblePlus stations



Genomics Observatory data

- Reinstated the OSD sampling and metadata registration tool
- Needs testing before use for 2019 event
- Discussions to be organised with GF BIO and EBI on:
 - Metadata standards
 - Data infrastructure
 - Data brokering

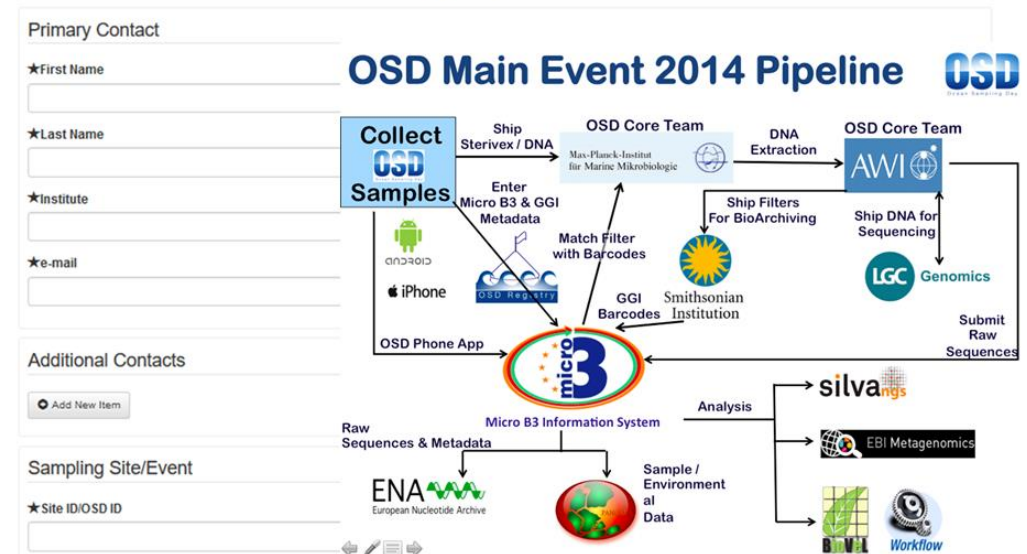
OSD Sample and Metadata Registration

On this page you will be able to register your OSD summer solstice 2016 samples.

Please consider the following before you start filling out the form:

- Please, have your filled-out OSD log-sheet at hand.
- Please provide all information in English.
- Please enter all decimal values using point notation.
- Almost each field has a descriptive text below. There you also find the respective unit of measurement we require (same as in OSD Handbook).
- Geographic Coordinates are in [decimal degrees](#).
- We will ask you about data you already gave during site registration. This allows data cross-checking.
- The form is very similar to the OSD log-sheet you have to fill out and sent with the samples.
- If you have questions about the data you are asked for. Please consult [Ocean Sampling Day Handbook 2.0](#)
- Please, if all breaks loose, you can email your questions to us. We are happy to help :)

OSD Contextual Data Form per Environmental Sample (v1.1)



Task NA2.5 Set up virtual platform for data analysis

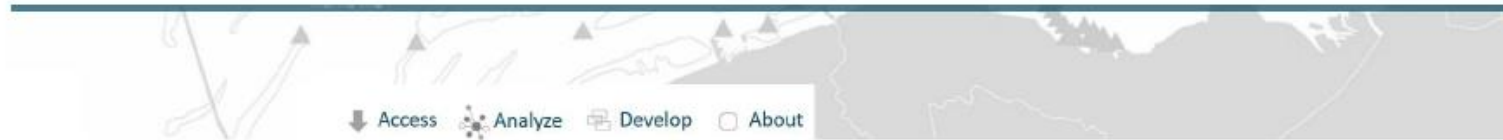
Task leader: HCMR with contributions from VLIZ, MPI, and UGOT.

Deliverables: D NA2.3a, b, c: Virtual Access runs to analysis platform (M24, M36, M48)

The **standardized data from long-term biodiversity and genomics observatories will be connected to a virtual analysis platform**. This platform will be an online portal (based on R web services and Taverna workflow systems) that enables the user to select from available datasets and perform predefined processing and analysis workflows on the selected data. Examples of workflows offered by the platform are existing outlier detection, visualization, statistical analysis and index calculation **workflows for biodiversity observation data and bioinformatics pipelines** (QIIME based demultiplexing, quality filtering, OTU picking, etc.). Depending on the type of workflow, output will be displayed as maps, reports and/or data tables. An authentication system will allow users to log on to the platform and access the pre-installed workflows.

A **workshop (M24)** will be organized to allow the data providers to **test and validate the analysis platform**. During the workshop a **webinar** will be prepared for teaching and stimulating the further use of the platform.

Use of the virtual analysis platform will be monitored using built-in tools. Number of users and “runs” per user will be assessed by the Advisory Board and reported on a yearly deliverable to the EC (DNA2.3a, b, c at M24, 36, 48).



Access

Retrieve and access data resources holding marine biodiversity and ecosystem data. A range of data systems offering data on species names, traits, distribution and genes.

Analyze

Online tools that facilitate data analysis of marine biodiversity and ecosystem data. Analysis is performed on data from known data resources and/or data uploaded by the user.



Runs

Welcome to the BioVeL Portal
For technical support or questions about the BioVeL Project, please visit the contact page

Choose an analysis.

Four analysis options with icons: Taxonomic Refinement, Ecological Niche Modelling, Metagenomics, and Phylogenetics.

RvLab interface showing Admin Toolbar, Workspace File Management, and a list of input files including '18S-ODV-input-corrected-dec15-formatted', 'Dese-Depth-labels.csv', and 'taxadis_job1737.csv'.

Home page content with sections: Exploration, Notes on data structure, Visualisations, Why multivariate analysis?, (Dis)similarity-based methods, BIOENV, Cluster analysis, Cluster analysis & ordination, Non-metric multidimensional scaling, Principal coordinates analysis, SIMPER, Constrained analyses, Canonical correlation analysis, Canonical.

Home

Welcome to the Guide to STAs:

This guide contains descriptions of a range of **multivariate approaches** are right if you are familiar with what data structure

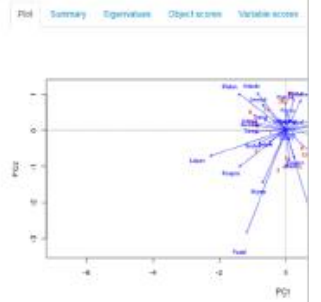
Please visit [our blog](#) to keep up with recent

This resource is described in: **Buttigieg PL, Ramette A (2014) A Guide to Analysis in Microbial Ecology: a comprehensive review of multivariate data analyses 90: 543-550.**
Please cite this paper if you make use of the MASAME applications.
This site is best viewed with Google Chrome [here](#)...

Multivariate Analysis Application for Microbial Ecology (MASAME)

Perform a principal components analysis...

Form for performing a PCA analysis with fields for Data upload, Transformations, PCA parameters, Download results, Description, CV parameters, and Separator.





Thank You!