



# SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE  
FOR OCEAN & MARINE DATA  
MANAGEMENT

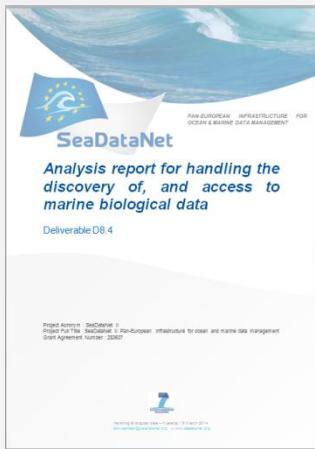
## *Status of tests with handling marine biological data in SeaDataNet - follow-up of SeaDataNet Deliverable 8.4*

TTT 23 September 2014



# Deliverables

- D8.4a Analysis report with required adaptions for marine biological data
- D8.4b Format documentation
- Example files





## **D8.4a Analysis report**

- **Analysis**
  - Types of data
  - Data use and applications
  - Standards and practices in the biological community
  - Requirements
  - Required adaptations of SeaDataNet infrastructure
- **Data transport format**
- **Towards operational data exchange**



**SeaDataNet**

PAN-EUROPEAN INFRASTRUCTURE  
FOR OCEAN & MARINE DATA  
MANAGEMENT

## **D8.4b Format documentation**

- Format:
  - CDI metadata
  - ODV biology variant
    - Mandatory fields (8 + 10)
    - Optional fields: P01 & P06 terms
    - Semantic header
- Version: BioODV\_1.0



**SeaDataNet**

PAN-EUROPEAN INFRASTRUCTURE  
FOR OCEAN & MARINE DATA  
MANAGEMENT

## *Example files*

- Types
  - Grab/core benthos community data with density and biomass values.
  - Zooplankton community with samples from different depths
  - Demersal fish population data with densities for different size classes and individual fish measurements (examples of subsamples are included)
  - Pollutant concentrations in biota specimens



- **Example 1:** Grab/core benthos community data (BioODV\_1.0\_MacroB) - This example demonstrates how to cover both density and biomass of infaunal (in this case macrobenthos) community data.

CDI mandatory fields	CDI optional fields	ODVbio mandatory fields	ODVbio optional fields
All mandatory CDI fields	No specific optional CDI fields	Cruise	DensityPerUnitEffort [unit]*
		Station	AFDWBiomassPerUnitEffort [unit]*
		Type	AFDWBiomass [unit]*
		yyy-mm-ddThh:mm:ss.sss	...
		Longitude [degrees_east]	
		Latitude [degrees_north]	
		LOCAL_CDI_ID	
		EDMO_code	
		Bot. Depth [m]	
		MinimumObservationDepth	
		MaximumObservationDepth	
		SampleID*	
		SamplingEffort [unit]*	
		SubSampleID*	
		SubSamplingCoefficient*	
		ScientificName*	
		ScientificNameID*	
		Sex*	
		LifeStage*	
		ObservedIndividualCount*	

(\*) fields indicated require paired QC flag column



- **Example 2:** Vertical Profile zooplankton community data (BioODV\_1.0\_ZooP) - This example demonstrates how to cover zooplankton data from vertical net sampling of different depth zones.

<u>CDI mandatory fields</u>	<u>CDI optional fields</u>	<u>ODVbio mandatory fields</u>	<u>ODVbio optional fields</u>
All mandatory CDI fields	No specific optional CDI fields	Cruise	DensityPerUnitEffort [unit]*
		Station	...
		Type	
		www-mm-ddThh:mm:ss.sss	
		Longitude [degrees_east]	
		Latitude [degrees_north]	
		LOCAL_CDI_ID	
		EDMO_code	
		Bot. Depth [m]	
		MinimumObservationDepth	
		MaximumObservationDepth	
		SampleID*	
		SamplingEffort [unit]*	
		SubSampleID*	
		SubSamplingCoefficient*	
		ScientificName*	
		ScientificNameID*	
		Sex*	
		LifeStage*	
		ObservedIndividualCount*	

(\*) fields indicated require paired QC flag column



- **Example 3:** Fish trawl data (BioODV\_1.0\_DemFish) - This specific example demonstrates how to cover demersal fish population data resulting from towed net fish trawls. The example demonstrates how to deal with subsampling storing information on the applied SubSamplingCoefficient and storing identifiers for both sample and subsample. The example includes both real counts of fish and derived densities per unit effort. Furthermore exemplary cases for including length-frequency data (counts and densities for different size classes), individual fish length measurements, and age information are available.

CDI mandatory fields	CDI optional fields	ODVbio mandatory fields	ODVbio optional fields
All mandatory CDI fields	Tracks (Curves)	Cruise	SubSamplingCoefficient*
		Station	ObservedMinLength [unit]*
		Type	ObservedMaxLength [unit]*
		www-mm-ddThh:mm:ss.sss	ObservedIndividualLength [unit]*
		Longitude [degrees_east]	DensityPerUnitEffort [unit]*
		Latitude [degrees_north]	Age [unit]*
		LOCAL_CDI_ID	SampleDuration [unit]*
		EDMO_code	
		Bot. Depth [m]	
		MinimumObservationDepth	
		MaximumObservationDepth	
		SampleID*	
		SamplingEffort [unit]*	
		SubSampleID*	
		SubSamplingCoefficient*	
		ScientificName*	
		ScientificNameID*	
		Sex*	
		LifeStage*	
		ObservedIndividualCount*	



- **Example 4:** Biota pollutant concentration data (BioODV\_1.0\_BiotaPol) - This example demonstrates how to cover data on pollutant concentrations in biota.

<u>CDI mandatory fields</u>	<u>CDI optional fields</u>	<u>ODVbio mandatory fields</u>	<u>ODVbio optional fields</u>
All mandatory CDI fields	No specific optional CDI fields	Cruise	Lead_per_unit_wet_weight_of_biotा[unit]*
		Station	...
		Type	
		yyy-mm-ddThh:mm:ss.sss	
		Longitude [degrees_east]	
		Latitude [degrees_north]	
		LOCAL_CDI_ID	
		EDMO_code	
		Bot. Depth [m]	
		MinimumObservationDepth	
		MaximumObservationDepth	
		SampleID*	
		SamplingEffort [unit]*	
		SubSampleID*	
		SubSamplingCoefficient*	
		ScientificName*	
		ScientificNameID*	
		Sex*	
		LifeStage*	
		ObservedIndividualCount*	

(\*) fields indicated require paired QC flag column



## **Roadmap for implementation**

SDN training: Biology workshop (21/05/2014, Ostend)

- 1) Format testing for well-defined biology data types
- 2) Download manager (testing, listing issues, adapting software)
- 3) Ocean Data Viewer (detailing requirements, listing developments, adapting software)
- 4) Machine-to-machine interface: creating data flow for public biology data



## 1) Format testing

### P01 terms for ODV bio

ConceptID	Label		
MINWDIST	Minimum depth below surface of the water body		
MAXWDIST	conceptid	preflabel	
SAMPID01	SDBIOL01	Abundance of unspecified biological entity per unit volume of the water body	
AREABEDS	SDBIOL02		
SSAMID01	SDBIOL03	conceptid	preflabel
SSAMPC01	SDBIOL04	OBSINDLX	Length of unspecified biological entity
SCNAME01	SDBIOL05	OBSMAXLX	
SNANID01	SDBIOL06	OBSMINLX	
ENTSEX01	SDBIOL07	SOVO0001	SAMPPROT
LSTAGE01	SDBIOL08	SOVO0004	ZDRZZ01
OCOUNT01	SDBIOL09	PBBIOTUK	Sample duration
	SDBIOL10	LENTRACK	Length of sampling track
	SDBIOL11		Start Latitude
			End Latitude
			Start Latitude
			End Latitude



# 1) Format testing

IFREMER:

REPHY dataset

Phytoplankton monitoring

479 CDI records in test portal

-> 479 ODV bio files

222.392 occurrences

SEADATANET COMMON DATA INDEX (CDI) V3

Tools

Enlarge Help Position Index

Datasets 0 Basket Reset

SeaDataNet

Add to basket

Timeseries on Summary Zoom to selected Export result Store query

| Refine query | New query | Found 479 | Show (1-20) | Previous | Next 20

Layer control

CDI entry Points CDI entry Tracks CDI entry Areas

Grid Lines Regional sea Regional sea labels Main sea Main sea labels

Display all selected records  Only selected records in results list

**Listing results**

20 100 1000 records Go

#	Data set name	Country	Start date	Disciplines - Parameter groups	Instrument / gear type	Show
	018-P-018 - Agon (a) - Biology - Surface	France	19870629	Biological oceanography > Biota abundance, biomass and diversity > Phytoplankton	discrete water samplers	
	014-P-023 - Géfosse - Biology - Surface	France	20020416	Biological oceanography > Biota abundance, biomass and diversity > Phytoplankton	discrete water samplers	
	010-P-110 - Estuaire de l'Orne - Biology - Surface	France	20070307	Biological oceanography > Biota abundance, biomass and diversity > Phytoplankton	discrete water samplers	
	012-S-032 - Baie de Seine - Zone 4 - Biology - Surface	France	20050405	Biological oceanography > Biota abundance, biomass and diversity > Phytoplankton	discrete water samplers	
	125-P-038 - Cap Saint Martin - Biology - Surface	France	20070612	Biological oceanography > Biota abundance, biomass and diversity > Phytoplankton	discrete water samplers	
	116-P-002 - BIS - Biguglia sud - Biology - Shallow depth	France	20090630	Biological oceanography > Biota abundance, biomass and diversity > Phytoplankton	discrete water samplers	
	110-P-002 - Jai - Biology -	France	19870331	Biological oceanography	discrete water samplers	

```

//SDN_parameter_mapping
//<subject>SDN:LOCAL:MinimumObservationDepth</subject><object>SDN:P01::MINDPCDI</object><units>SDN:P06::ULAA</units>
//<subject>SDN:LOCAL:MaximumObservationDepth</subject><object>SDN:P01::MAXDPCDI</object><units>SDN:P06::ULAA</units>
//<subject>SDN:LOCAL:SampleID</subject><object>SDN:P01::SAMPID01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:SamplingEffort</subject><object>SDN:P01::SASIZE01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:SubsampleID</subject><object>SDN:P01::SSAMID01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:SubSamplingCoefficient</subject><object>SDN:P01::SSAMPC01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:ScientificName</subject><object>SDN:P01::SCNAME01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:ScientificNameID</subject><object>SDN:P01::SNAMID01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:Sex</subject><object>SDN:P01::ENTSEX01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:LifeStage</subject><object>SDN:P01::LSTAGE01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:ObservedIndividualCount</subject><object>SDN:P01::OCOUNT01</object><units>SDN:P06::UUUU</units>
//<subject>SDN:LOCAL:IndividualCountperLiter</subject><object>SDN:P01::SDBIOL01</object><units>SDN:P06::UCPL</units>
//

```

Cruise	Station	Type	yyyy-mm-ddThh:mm:ss.s	Longitude	Latitude	[i] LOCAL_CD	EDMO_co	Bot. Depl	Minimum	QV:SEAD	Maximum	QV:SEAD	SampleID	QV:SEAD	SamplingE	QV:SEAD	Subsampl	QV:SEAD	SubSampl	QV:SEAD	Scientific
001-P-015	1001022_E*		1992-01-07T14:00:00.000	23334994	51.068649	1001022_E		486 0.0	.5	1.5		1	50021	1		9	50021	1	1	1	Prorocent
									.5	1.5		1	50021	1		9	50021	1	1	1	Dinophysi
									.5	1.5		1	50021	1		9	50021	1	1	1	Chaetocer
									.5	1.5		1	50021	1		9	50021	1	1	1	Thalassio
									.5	1.5		1	50021	1		9	50021	1	1	1	Guinardia
									.5	1.5		1	50021	1		9	50021	1	1	1	Guinardia
									.5	1.5		1	50021	1		9	50021	1	1	1	Chaetocer
									.5	1.5		1	50021	1		9	50021	1	1	1	Asterione
									.5	1.5		1	50021	1		9	50021	1	1	1	Pseudo-ni
									.5	1.5		1	50021	1		9	50021	1	1	1	Dinophysi
001-P-015	1001022_E*		1992-02-06T14:00:00.000	23334994	51.068649	1001022_E		486 0.0	.5	1.5		1	50022	1		9	50022	1	1	1	Dinophysi
									.5	1.5		1	50022	1		9	50022	1	1	1	Thalassio
									.5	1.5		1	50022	1		9	50022	1	1	1	Chaetocer
									.5	1.5		1	50022	1		9	50022	1	1	1	Leptocylir
									.5	1.5		1	50022	1		9	50022	1	1	1	Skeleton
									.5	1.5		1	50022	1		9	50022	1	1	1	Guinardia
									.5	1.5		1	50022	1		9	50022	1	1	1	Asterione
									.5	1.5		1	50022	1		9	50022	1	1	1	Nitzschia

Time precision



# 1) Format testing

- IMARES:
  - Seal counts
    - Aeroplane-Cameras
  - Benthos monitoring
    - Benthos samplers
  - Demersal Young Fish
    - Beam trawls
- 467 CDI files
  - > 52 SHP - files
  - > 415 ODV - files

The screenshot displays the SeaDataNet Common Data Index (CDI) V3 interface. On the left, there's a sidebar with various tools like search, zoom, and a 'Basket' section showing 0 items. The main area features a map of the Dutch Wadden Sea with a red hatched polygon overlaid. To the right of the map is a 'Layer control' panel with checkboxes for different data types such as CDI entry Points, CDI entry Tracks, CDI entry Areas, Grid Lines, Regional sea, Regional sea labels, Main sea, and Main sea labels. Below the map is a 'Listing results' table with the following data:

#	Data set name	Country	Start date	Disciplines - Parameter groups	Instrument / gear type	Show
1	Monitoring Seals in the Dutch Wadden Sea - Harbour seals moult 2013	Netherlands	20130815	Biological oceanography > Birds, mammals and reptiles	cameras	
2	Monitoring Seals in the Dutch Wadden Sea - Harbour seals moult 2012	Netherlands	20120815	Biological oceanography > Birds, mammals and reptiles	cameras	
3	Monitoring Seals in the Dutch Wadden Sea - Harbour seals moult 2011	Netherlands	20110812	Biological oceanography > Birds, mammals and reptiles	cameras	







## 1) *Format testing*

- **IMGW:**
  - National Environment Monitoring
  - Macrophytes, zoobenthos, zooplankton, phytobentos
  - 5 ODV files
    - > No CDI records yet
  - >35.000 occurrences
- **HCMR**
- **Aarhus University DKU**



**SeaDataNet**

PAN-EUROPEAN INFRASTRUCTURE  
FOR OCEAN & MARINE DATA  
MANAGEMENT

## 1) *Format testing*

- Non critical issues -> FAQ
  - Feedback phase
  - Testing phase





## 1) *Format testing*

- Non critical issues -> FAQ
- Critical issues:
  - 1) Local CDI ID <=> eventID (Example: IFREMER IMARES, IMGW,)
    - No info on start and end time
    - No info start and end position
    - No info on sampling protocol
    - Not in CDI either

=> Include in data files (ODV)

OR => Make CDI ID = eventID obligatory



## 1) *Format testing*

- Non critical issues -> FAQ
- Critical issues:
  - 1) Local CDI ID <=> eventID (Example: IFREMER IMARES, IMGW,)
    - No info on start and end time
    - No info start and end position
    - No info on sampling protocol
    - Not in CDI either
      - => Include in data files (ODV)
  - OR => Make CDI ID = eventID obligatory
  - 2) Empty Latitude/longitude (Example: IMARES seal data)
    - => some data products not fit for ODV?



## 2) Download manager: format creation

- ODV bio file generation has been tested
- Allow special forms ISO 8601 DateTime  
(eg. 1979-04, 2012-06-03T03, 2011-04-02T06:04:01/ 2011-04-02T06:08:12)

```
<?xml version="1.0" encoding="UTF-8"?>
- <root>
  - <updated>2014-08-29T07:00:00+02:00</updated>
  - <codes type="odv">
    <code local="SurveyNo" to="SDN:LOCAL:Cruise" from="" />
    <code local="Station" to="SDN:LOCAL:Station" from="" />
    <code local="Date" to="SDN:LOCAL:yyyy-mm-ddThh:mm:ss.sss" from="" />
    <code local="Longitude" to="SDN:LOCAL:Longitude, degrees_east" from="" />
    <code local="Latitude" to="SDN:LOCAL:Latitude, degrees_north" from="" />
    <code local="BotDepth" to="SDN:LOCAL:Bot. Depth, m" from="" />
    <code local="Type" to="SDN:LOCAL:Type" from="" />
    <code local="MinDepth" to="SDN:LOCAL:MinimumObservationDepth, m" from="SDN:P01::MINWDIST, SDN:P06::ULAA" qflag="fMinDepth"/>
    <code local="MaxDepth" to="SDN:LOCAL:MaximumObservationDepth, m" from="SDN:P01::MAXWDIST, SDN:P06::ULAA" qflag="fMaxDepth"/>
    <code local="SampleID" to="SDN:LOCAL:SampleID, #" from="SDN:P01::SAMPID01, SDN:P06::UUUU" qflag="fSampleID"/>
    <code local="Area" to="SDN:LOCAL:SamplingEffort, m^2" from="SDN:P01::AREABEDS, SDN:P06::UMSQ" qflag="fArea"/>
    <code local="SubSampleID" to="SDN:LOCAL:SubsampleID, #" from="SDN:P01::SSAMID01, SDN:P06::UUUU" qflag="fSubSampleID"/>
    <code local="SubCoef" to="SDN:LOCAL:SubSamplingCoefficient, none" from="SDN:P01::SSAMPC01, SDN:P06::UUUU" qflag="fSubCoef"/>
    <code local="LatinName" to="SDN:LOCAL:ScientificName, none" from="SDN:P01::SCNAME01, SDN:P06::UUUU" qflag="fLatinName"/>
    <code local="WormsID" to="SDN:LOCAL:ScientificNameID, #" from="SDN:P01::SNAMID01, SDN:P06::UUUU" qflag="fWormsID"/>
    <code local="Sex" to="SDN:LOCAL:Sex, none" from="SDN:P01::ENTSEX01, SDN:P06::UUUU" qflag="fSex"/>
    <code local="Stage" to="SDN:LOCAL:LifeStage, none" from="SDN:P01::LSTAGE01, SDN:P06::UUUU" qflag="fStage"/>
    <code local="Count" to="SDN:LOCAL:ObservedIndividualCount, #" from="SDN:P01::OCOUNT01, SDN:P06::UUUU" qflag="fCount"/>
    <code local="WetWeight" to="SDN:LOCAL:WetWeight, g" from="SDN:P01::SDBIOL05, SDN:P06::UGRM" qflag="fWetWeight"/>
    <code local="DryWeight" to="SDN:LOCAL:DryWeight, g" from="SDN:P01::SDBIOL08, SDN:P06::UGRM" qflag="fDryWeight"/>
  </codes>
</root>
```

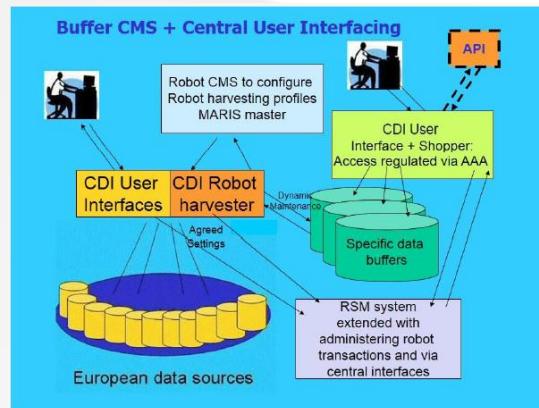


### ***3) Ocean Data Viewer: Format reading***

- Read and display text variables
- Aggregate single localCDI data files
- Allow special forms ISO 8601 DateTime  
(eg. 1979-04, 2012-06-03T03, 2011-04-02T06:04:01/ 2011-04-02T06:08:12)
- Incorporate information from CDI
  
- Develop appropriate filtering and visualisation tools



## 4) Machine-to-machine interface: data flow



- Machine to machine services – API
  - Create\_order: enabling to configure a discovery profile for shopping in an authorised central buffer database;
  - Get\_order\_info: enabling to retrieve info from the RSM about the progress of the order processing;
  - Download\_order: enabling to download the orders, when ready, as zip files with metadata and data sets.
  - Services secured by SDN CAS (AAA) login



## 4) *Machine-to-machine interface: data flow*

- Data flow using data buffer has been discussed within EMODNet biology meeting Horta 09/2014
- Candidates are IFREMER, IMARES, IMGW and possibly Aarhus univ.
- Need to define specific profile (Specify: datatype, region, datacenters, etc) => in the SeaDataNet buffer system a subset is created.
- The SDN robot builds up a (consistently updated) specific subset of data relevant for EMODNet-Biology.
- To create this buffer an agreement of involved data centers is required. => An email will be sent by EMODNet Biology data providers to MARIS to specify involved data
- Deadline of data delivery: december 2014



## ***In conclusion***

- Decide on mentioned critical issues
- Finalise testing phase format
  - Adapt documentation where needed
  - Add to documentation:
    - FAQ
    - Additional examples with list of already used P01 concepts
- Discuss developments on DM and ODV
- Prepare buffer for data flow of biology data in the framework of EMODNet Biology(before december 2014)



**SeaDataNet**

*PAN-EUROPEAN INFRASTRUCTURE  
FOR OCEAN & MARINE DATA  
MANAGEMENT*

[sdn-userdesk@seadatanet.org](mailto:sdn-userdesk@seadatanet.org) – [www.seadatanet.org](http://www.seadatanet.org)