Marine Microbial Biodiversity, Bioinformatics & Biotechnology

Grant agreement n°287589

Acronym : Micro B3
Start date of project: 01/01/2012, funded for 48 month

Deliverable 4.4

Finalised interoperability structures

Version: 1
Circulated to: Micro B3 Consortium (11.12.2013)
Approved by: Name (Date)

Expected Submission Date: 31.12.2013
Actual submission Date: 24.12.2013

Lead Party for Deliverable: EMBL-EBI

Mail: petra@ebi.ac.uk **Tel.: +44 1223 492565**

Dissemination level:	
Public (PU)	x
Restricted to other programme participants (including the Commission Services) (PP)	
Restricted to a group specified by the consortium (including the Commission Services) (RE)	
Confidential, only for members of the consortium (including the Commission Services) (CO)	

The Micro B3 project is funded from the European Union's Seventh Framework Programme (Joint Call OCEAN.2011-2: Marine microbial diversity – new insights into marine ecosystems functioning and its biotechnological potential) under the grant agreement no 287589. The Micro B3 project is solely responsible for this publication. It does not represent the opinion of the EU. The EU is not responsible for any use that might be made of data appearing herein.

Summary

Understanding the impact of climate change on marine life requires holistic approaches that bring together the biodiversity, biomolecular and environmental domains. Data from such studies will allow analysis of relations between genetic, functional and morphological diversity, and changes of environmental, chemical and physical, parameters. Micro B3 has built strong connections between biodiversity, biomolecular and environmental communities and their data resources to provide the infrastructure needed to reflect these integrated approaches. The present deliverable, D4.4, deals with the details of interoperability between oceanographic data (available via a network of oceanographic data centres managed by the SeaDataNet), marine biodiversity data (available via the EurOBIS) and biomolecular data (available via the European Nucleotide Archive at EMBL-EBI), and drives at an integrated view of these data (to be made available via the Micro B3 Information System). We have agreed upon a minimal set of descriptors (sampling site and platform, event date/time, event latitude, event longitude and sample depth) that must, in order to meet Micro B3 standards, be presented as searchable fields in compliant data resources; we refer to this as the **Micro B3 core**. We have also agreed upon three search service protocols (the OpenSearch, OGC WMS and WFS protocols) that will be used by Micro B3 data resources to support searches against the Micro B3 core descriptors. Finally, we describe how data submission routings for two flagship Micro B3 projects, Ocean Sampling Day (OSD) and Tara Oceans, will be configured to support compliance at the data resources with the Micro B3 core.

Table of contents

1.0 Objectives of the deliverable 4.4

Establishing cross-domain links between molecular researchers building and archiving genome/metagenome collections of marine samples on one side and oceanographic experts providing environmental data associated with the collected sample on the other side will contribute profoundly to research in marine ecosystem biology and modelling and become an important legacy of the Micro B3 project.

MicroB3 interoperability structures reflect the link between molecular microbial research and oceanography and provide integrated view of marine microbial diversity.

Deliverable D5.5 reported on new services implemented at the molecular data archive, the European Nucleotide Archive (ENA) at the EMBL-EBI, to improve discoverability of marine-related read and sequence data. Deliverable D3.1 reported on services of the oceanographic data management infrastructure, SeaDataNet, and on a network of oceanographic data centres including the marine biodiversity data resource, EurOBIS. Deliverables D5.6 and D3.3 drafted connections between the oceanographic, biodiversity and molecular domains conceptually as well as technically. Workflows for both marine data submission and retrieval have been proposed and services discussed for the discovery and delivery of data between components of the Micro B3 infrastructure. These protocols are the Open Geospatial Consortium (OGC) standards – the Web Map Service (WMS), Web Feature Service (WFS) and the OpenSearch protocol.

Deliverable D4.4 describes the **Micro B3 core** – the finalised Micro B3 interoperability structure - which is based on the above mentioned deliverables, on intensive discussions among work packages 3, 4 and 5, and on conclusions of the Micro B3 interoperability workshop held on 23 - 24 October 2013 in the EMBL-EBI, Hinxton, UK (Chapter 3). The Micro B3 core represents a platform-independent Micro B3 interoperability standard. Implementation against this standard will vary across resources and over time, but we include in the Chapters 6, 7 and 8 examples of an expected implementation at ENA, EurOBIS and SeaDataNet.

2.0 Micro B3 core and enhanced interoperability

We have identified a minimal set of core interoperability descriptors, which are central to integration of data from independent resources and allow connection between diverse data across molecular, biodiversity and oceanographic domains.
The table below lists the Micro B3 core descriptors together with their definitions and recommended syntax/typing.

Field	Description	Control vocabulary/format *	Example
SAMPLING_ Site	Refers to the site/station where data/sample collection is performed.	SDN:C17 or OSD Sites Registry**	Poseidon-E1-M3A Time Series Station
SAMPLING_ Platform	Refers to the large infrastructure from which data/sample collection is performed, e.g. a ship or a coastal observatory.	SDN:L06	research vessel
EVENT_ Date/Time	Date and time when the sampling event started and ended, e.g. each CTD cast, net tow, or bucket collection is a distinct event.	Date and time in UTC; Format: yyyy-mm-ddThh:mm:ssZ	2013-06-21T14:05:00Z/ 2013-06-21T14:46:00Z
EVENT_ Longitude	Longitude of the location where the sampling event started and ended, e.g. each CTD cast, net tow, or bucket collection is a distinct event	Format: ###.#### Decimal degrees; East= +, West= - Format: Use WGS 84 for GPS data	035.6666E 035.6702E
EVENT_ Latitude	Latitude of the location where the sampling event started and ended, e.g. each CTD cast, net tow, or bucket collection is a distinct event	Format: ##.#### Decimal degrees; North= +, South= - Format: Use WGS 84 for GPS data	24.6666N 24.6643N
SAMPLE_ Depth	The distance below the surface of the water at which a measurement was made or a sample was collected.	Format : ##.# Positive below the sea surface. SDN:P06::ULAA for m	14.7 m

Table 1: **Core interoperability descriptors.**

* SDN:C17::XXXX is a controlled register for platform codes, observation stations and sites (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=C17 for human interface)

*SDN:L06::XX is a controlled Terms list describing "CATEGORIES" of platforms. (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=L06 for human interface)

* SDN:P02::XXXX is a controlled Terms list describing "WHAT" is measured. (http://www.seadatanet.org/urnurl/SDN:P02::XXXX for XML response) (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=P02 for human interface)

* SDN:P06::XXXX is a controlled Terms list describing "UNITS" of measurements. (http://www.seadatanet.org/urnurl/SDN:P06::XXXX for XML response) (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=P06 for human interface)

** OSD Sites Registry is a controlled register for OSD sampling Sites maintained by the Micro B3 IS. (http://oceansamplingday.blogspot.co.uk)

These six descriptors are part of the mandatory Micro B3 checklist (described in D4.3 and summarised in the Table 2 below), which is the minimal set of reporting descriptors that must accompany all Micro B3/OSD-collected marine microbial samples. Since Micro B3 core

is fully included in the Micro B3 reporting standard, every Micro B3-compliant dataset will be discoverable by services using the interoperability core in their search queries.

Field	Description	Control vocabulary/format *	Example
SAMPLING_ Campaign	Refers to a finite or indefinite activity aiming at collecting data/samples, e.g. a cruise, a time series, a mesocosm experiment.	Free text;	OSD-SS2014
SAMPLING_ Site	Refers to the site/station where data/sample collection is performed.	SDN:C17 or OSD Sites Registry**	Poseidon-E1-M3A
SAMPLING_ Platform	Refers to the large infrastructure from which data/sample collection is performed, e.g. a ship or a coastal observatory.	SDN:L06	research vessel
EVENT_ Date/Time	Date and time when the sampling event started and ended, e.g. each CTD cast, net tow, or bucket collection is a distinct event.	Date and time in UTC; Format: yyyy-mm-ddThh:mm:ssZ	2013-06-21T14:05:00Z/ 2013-06-21T14:46:00Z
EVENT_ Longitude	Longitude of the location where the sampling event started and ended, e.g. each CTD cast, net tow, or bucket collection is a distinct event	Format: ###.#### Decimal degrees; East= +, West= - Format: Use WGS 84 for GPS data	035.6666E 035.6702E
EVENT_ Latitude	Latitude of the location where the sampling event started and ended, e.g. each CTD cast, net tow, or bucket collection is a distinct event	Format: ##.#### Decimal degrees; North= +, South= - Format: Use WGS 84 for GPS data	24.6666N 24.6643N
SAMPLE_ Depth	The distance below the surface of the water at which a measurement was made or a sample was collected.	Format: ##.# Positive below the sea surface. SDN:P06:46:ULAA for m	14.7 m
SAMPLE_ Protocol_Label	Identifies the protocol used to produce the sample, e.g. filtration and preservation.	Term list; See the SAMPLE_Protocol_Short_Label in the OSD Protocols Section for details.	NP0223
SAMPLE_ ID	A unique identifier (barcode) for each sample, e.g. one ID for each filter generated during sampling. The IDs are generated, printed and sent to the sampling groups following their registration to OSD. See the Registration Section for details.		SI NMNH barcode ID
ENVIRONMENT_ Biome	Refers to classes of ecologically similar communities of plants, animals, and other organisms.	Terms list: EnvO (v1.53)	ENVO:00000447 for “marine biome”

ENVIRONMENT_ Feature	Refers for example to geographic features, or natural & artificial habitats that characterise the environment.	Terms list: EnvO (v1.53)	ENVO:00000569 for "marine habitat"
ENVIRONMENT_ Material	Refers to the matter that was taken from the environment by the sampling event.	Terms list: EnvO (v1.53)	ENVO:00002042 for "surface water"
ENVIRONMENT_ Temperature	Temperature of water at the time of taking the sample.	Format: ##.# SDN:P02::TEMP SDN:P06::UPAA for °C	16.2 °C
ENVIRONMENT_ Salinity	Salinity of water at the time of taking the sample.	Format: ##.# SDN:P02::PSAL SDN:P06::UGKG for PSU	39.1 psu

Table 2: **Mandatory Micro B3 reporting descriptors.**

* SDN:C17::XXXX is a controlled register for platform codes, observation stations and sites (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=C17 for human interface)

*SDN:L06::XX is a controlled Terms list describing "CATEGORIES" of platforms. (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=L06 for human interface)

* SDN:P02::XXXX is a controlled Terms list describing "WHAT" is measured. (http://www.seadatanet.org/urnurl/SDN:P02::XXXX for XML response) (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=P02 for human interface)

* SDN:P06::XXXX is a controlled Terms list describing "UNITS" of measurements. (http://www.seadatanet.org/urnurl/SDN:P06::XXXX for XML response) (http://seadatanet.maris2.nl/v_bodc_vocab_v2/search.asp?lib=P06 for human interface)

** OSD Sites Registry is a controlled register for OSD sampling Sites maintained by the Micro B3 IS. (http://oceansamplingday.blogspot.co.uk)

The Micro B3 core, Table 1, is the necessary minimum that each interoperability resource, SeaDataNet, EurOBIS, ENA and the Micro B3 Information System (Micro B3 IS), will support to allow data discovery and connection between resources.

Each infrastructure, however, has a broad range of well-established services tailored to the community they serve. Enhanced interoperability address these resource-specific system components that could be consumed by Micro B3 partners and thus extend interoperability across Micro B3.
For EurOBIS and ENA, for example, enhanced interoperability will include queries using identifiers of the principal taxonomic index that each resource supports. Alternatively, requests for datasets can further be specified with particular environmental parameters, such as temperature or salinity. Each partner will evaluate costs of support for, and benefits to Micro B3, for these enhanced interoperability features.

3.0 Micro B3 interoperability workshop

Micro B3 WP3, WP4 and WP5 focus their efforts on establishing interoperability structures that will allow (1) the sharing of information among oceanographic, biodiversity and molecular domains and (2) the provision of cross-domain data into the Micro B3 Information system (Micro B3 IS).

The ENA at EMBL-EBI organised, for WP3/4/5, on 23 - 24 October 2013 in the EMBL-EBI, Hinxton, UK, the Micro B3 Interoperability Workshop. Individual workpackages were represented as follows:
WP1 – Frank Oliver Glöckner (Jacobs Uni)
WP3 – Peter Thijsse (MARIS), Dick Schaap (MARIS), Stefanie Dekeyzer (VLIZ), Catherine Borremans (IFREMER),
WP4 – Guy Cochrane (EMBL-EBI), Petra ten Hoopen (EMBL-EBI), Nicole Silvester (EMBL-EBI), Rajesh Radhakrishnan (EMBL-EBI)
WP5 – Renzo Kottmann (MPIMM), Ivaylo Kostadinov (MPIMM)

Agenda for the Micro B3 interoperability workshop:

23 Oct (EBI, South building, V2-62 Sawston meeting room)

12.00 - 1.00	working lunch
1.00 - 1.30	welcome and brief summary of the 'Micro B3 interoperability core' (Guy, Petra)
1.30 - 3.30	'core interoperability' - establishing data connection across resources for the interoperability core descriptors, decision on technologies and time line of implementation (discussion all, chaired by WP4)
3.30 - 4.00	coffee break
4.00 - 5.00	'core interoperability' (continue, discussion all, chaired by WP3)
5.00 - 6.30	data acquisition - finalising support for sampling Sites in recording/digitising/submitting acquired marine data and metadata to facilitate compliance (discussion all, chaired by WP5)
7.30	dinner at Hinxton Red Lion

24 Oct (EBI, South building, V2-62 Sawston meeting room)

9.00 - 10.30	'enhanced interoperability' - establishing enhanced data connection across resources based on resource-specific services (15 min each)
9.00 - 9.15	SeaDataNet (Dick/Peter)
9.15 - 9.30	ENA (Nicole)
9.30 - 9.45	Micro B3 IS (Renzo)
9.45 - 10.00	PANGAEA (Renzo)
10.00 - 10.15	VLIZ (Stefanie)
10.15 - 10.30	IFREMER (Catherine)
10.30 - 11.00	coffee break

11.00 - 12.15 next steps (discussion all, chaired by WP4)
12.15-12.30 summary and wrap up (Petra, Guy)
12.30 working lunch

Action points from the Micro B3 interoperability workshop

1. EurOBIS, SeaDataNet and ENA to provide by Dec 2013 a concise technical descriptions of their OpenSearch query and result services and their WMS and/or WFS offerings or future offerings, including timelines where not yet delivered and relationship of services to data model. These Micro B3 interoperability solutions are described in the Chapters 6, 7 and 8.
2. Micro B3 IS to provide by Dec 2013 specification for OSD data submission brokering, version 1.0 to be available in Mar 2014.
3. Cross-domain submitter re-direction
   a/ ENA to provide for NODCs descriptive and linking text to guide molecular data submitters from NODC to ENA.
   b/ SDN to push the re-direction to the SDN page and NODCs.
   c/ SDN to provide for ENA link(s) to the SDN page guiding oceanographic data submitters from ENA to NODC.
4. Pilot OSD 2012 data and metadata submission to the ENA
   a/ ENA to create confidential pilot OSD sample records, assesses the samples for the Micro B3 compliance and returns missing info to OSD coordinators.
   b/ OSD coordinators to contact sampling Sites to obtain missing info and to inform about OSD data public release.
   c/ MPIMM/GFBIO (Ivaylo Kostadinov) to retrieve data from the Argonne National Laboratory ftp site, map the pilot OSD samples to experimental metadata and submits to the ENA.
   d/ ENA to update the pilot OSD sample metadata with missing info and releases all pilot OSD 2012 data
5. Pangaea to send CDI summary data to SDN (for Micro B3 core discovery service).
6. SDN to advise on conversion between unit systems for salinity.
7. ENA to explore mapping WoRMS vs. NCBI (with advice from EurOBIS).
8. ENA to explore searches on the attribute <sample size> based on the values in the attribute field <sampling protocol>.
9. Micro B3 IS to explore possibility to build faceted search interface for marine community.

4.0 OpenSearch and OGC WMS-WFS protocols

Three services were proposed for discovery and delivery of data between the Micro B3 infrastructures. These suggested protocols are the Open Geospatial Consortium (OGC) standards – the Web Map Service (WMS) and Web Feature Service (WFS) and the OpenSearch protocol.

OpenSearch protocol

OpenSearch is a collection of formats that help search engine and search client communicate, enabling them to perform search requests and syndicate results. OpenSearch was created by the Amazon company and is now used by hundreds of search engines and clients. A description of a search engine, a location for retrieval of search results and a formal representation of results (RSS or Atom) allow publication on a web interface.

OGC WMS protocol

The OGC Web Map Service standard is an interface for requesting geo-registered map images from a geospatial database. The WMS request specifies the area of interest and the WMS response dynamically produces spatially referenced map images. The standard defines operations to return a description of the maps and a query about features displayed on the map. However, it does not retrieve feature data or data values.

OGC WFS protocol

The OGC Web Feature Service standard offers fine-grained access to geographic information at the feature and feature property level allowing search clients to retrieve only specific information they seek rather than files containing much more than requested. The standard defines operations for feature discovery, retrieval of features or values for feature properties and access to, or edits of, features.

5.0 Summary of Micro B3 interoperability structures

The broad scope and diversity of data supported in Micro B3 data resources limits the set of core interoperability descriptors to a small number of universally used attributes that can, relatively easily, be mapped between archives (see Table 1). Potentially any attribute can be made searchable and can be returned in a query response. However, the context in which each fine-grained piece of information appears, as well as performance of services delivering the response, have to be taken into account.

Micro B3 WP3, WP4 and WP5 committed themselves to the following:

1. SeaDataNet, ENA, EurOBIS and Micro B3 IS will work towards **support of Micro B3 core by mid 2014**.

2. **Focus** of the Micro B3 interoperability will be **on data discovery**. This eliminates considerations of data access authentication, which applies to SeaDataNet data outside of the discovery domain.

3. **A query** between portals will be facilitated by the **OpenSearch format**.

4. **A response** from portals will be **given in the OpenSearch format or OGC WMS and/or WFS format**. Both protocols are briefly summarised in the Chapter 4.

5. **Implementation** of support for attributes in remote queries will be given **in the following order of priorities**:
    1. Geo-location + time stamp
    2. Interoperability core (as defined in the Chapter 2, Table 1)
    3. Micro B3 mandatory reporting checklist (as defined in the Chapter 2, Table 2)

6. Remote queries of the geo-location and time stamp will include only the event start points.

7. The **Micro B3 IS** will support remote **queries on sample identifiers** issued by the PANGAEA and ENA.

8. **A response from the ENA portal will be given in two steps.** 1. Step – return ENA sample ID based on geo-location and time stamp query, 2. Step – return molecular data for the given ENA sample ID. Further details on the Micro B3 interoperability solution at the ENA are available in the Chapter 6.

9. **A response from the EurOBIS will include the scientific organism name** and **taxon count.** Further details on the Micro B3 interoperability solution at the EurOBIS are available in the Chapter 7.

10. **A response from the SeaDataNet portal will include CDI metadata** for the given ge-location and time stamp with appropriate grouping by oceanographic discipline. Further details on the Micro B3 interoperability solution at the SeaDataNet are available in the Chapter 8.

11. **Data submission to the Micro B3 infrastructures** is shown in the Figure 1 for two enterprises, the Ocean Sampling Day (OSD) campaign and the Tara Oceans expedition. Material for molecular analysis is dispatched to a Sequencing Centre and molecular data submitted to the ENA. For the OSD, all sample metadata as well as environmental and biodiversity data will be collected at a single entry point interface developed at the Micro B3 IS, who will then package relevant information to PANGAEA and ENA. For the TARA Oceans, all sample metadata as well as environmental and biodiversity data are archived at the central Sample Registry as described in the D3.2 and only relevant sample metadata submitted to the ENA. For both campaigns the PANGAEA provides oceanographic and morphology-based biodiversity information to SeaDataNet and EurOBIS, respectively.

12. **Data retrieval for the OSD campaign** is described in the Figure 2. Although the primary access and an integrated view of OSD samples will be provided by the Micro B3 IS a user can access specific pieces of information from the relevant archive directly. Information will be retrieved from National Oceanographic Data Centres (NODC) including Pangaea via the SeaDataNet management system for oceanographic data and via the EurOBIS for morphology-based biodiversity data. The ENA will provide read and sequence data. Using the Micro B3 interoperability structures archives will be able to connect to each other and refer to specific information on the partner's side.

13. **Data management infrastructure for the Tara Oceans expedition** is schematically shown in the Figure 4 of the deliverable 3.2.

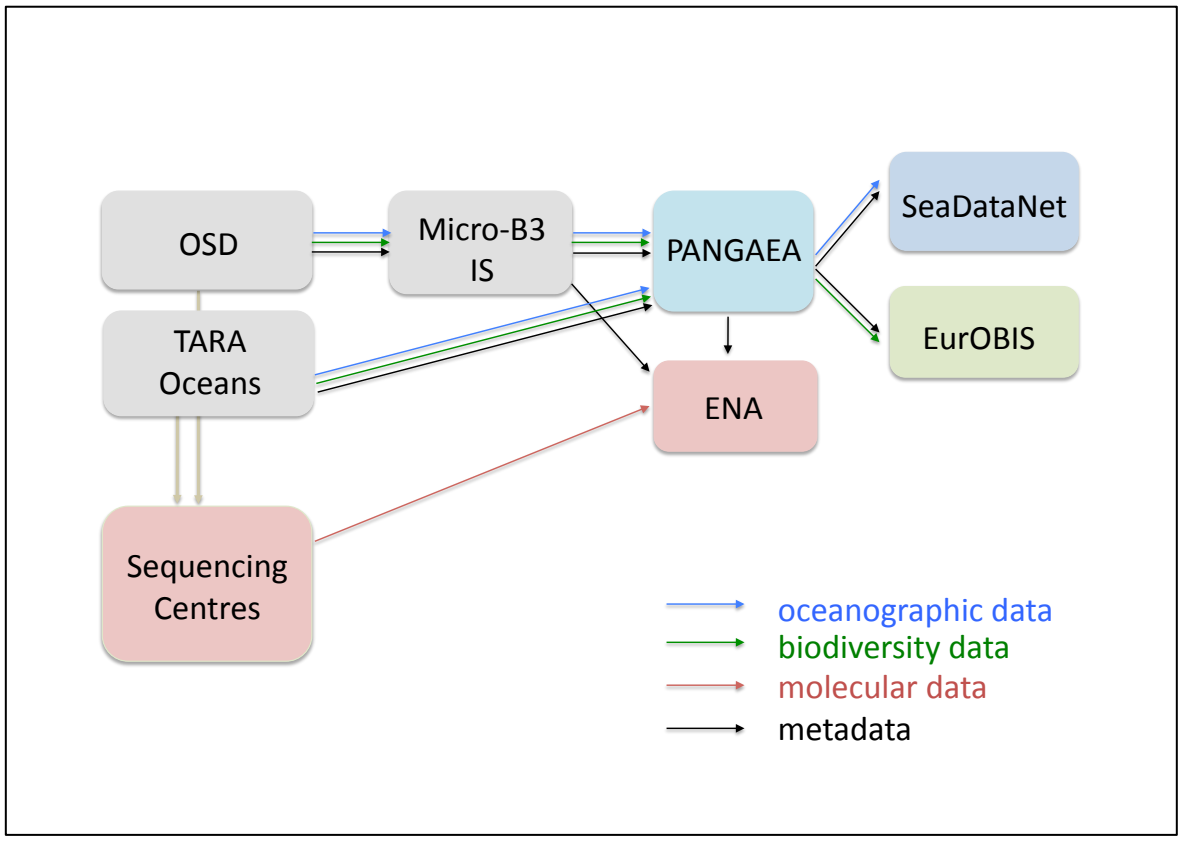

Figure 1: Data submission to the Micro B3 for the OSD and Tara Oceans campaign.

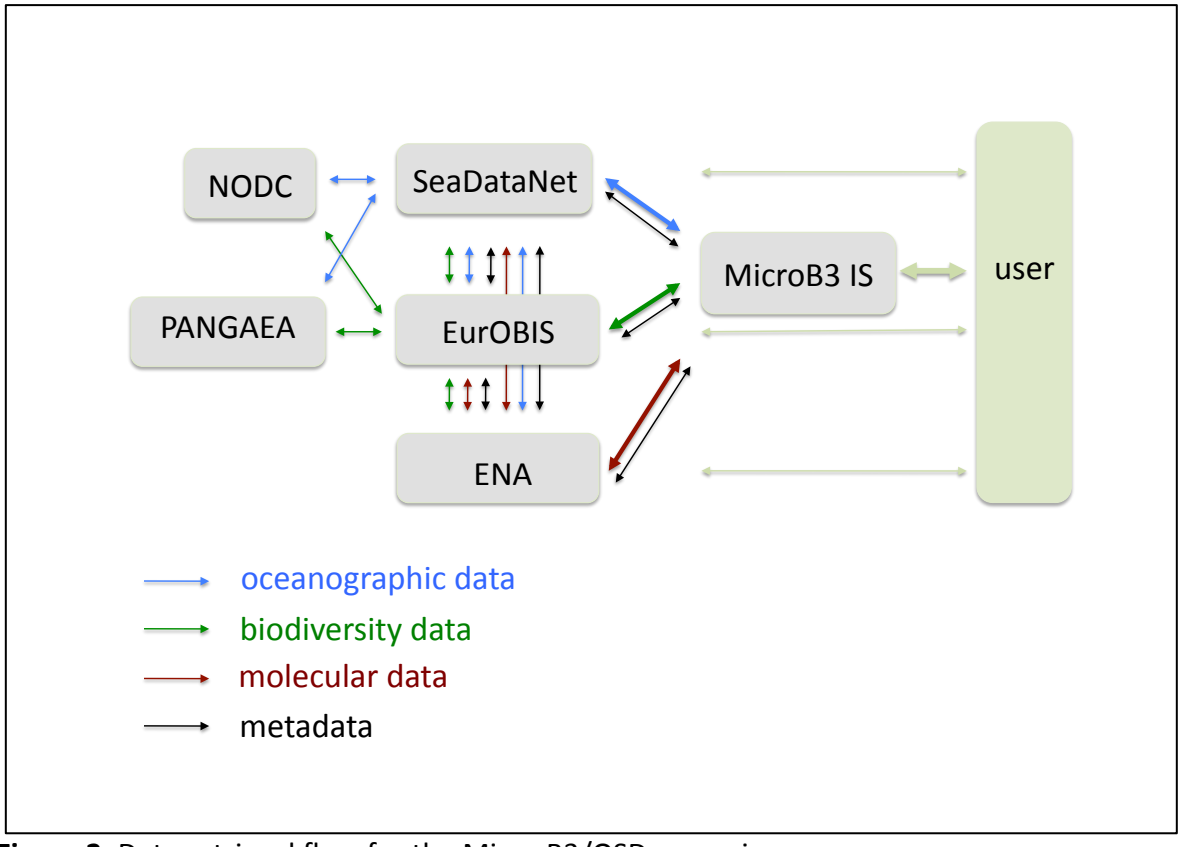

Figure 2: Data retrieval flow for the Micro B3/OSD campaign.

6.0 Example of Micro B3 core and enhanced interoperability implementation at ENA

6.1 Current functionality of ENA Advanced Search

ENA has an advanced search that supports searching by individual fields and returning of a choice of fields based on user-selected results. Currently, ENA supports sample search by location and sampling date fields. Other MicroB3 interoperability core descriptors (depth, sampling site and sampling platform) as well as sampling campaign and additional mandatory MicroB3 checklist attributes, will be added once data become available.

Querying ENA

Full information on building queries for ENA Advanced Search can be found at http://www.ebi.ac.uk/ena/about/browser#data_warehouse. Included here is specific information most likely to be of use in ENA-MicroB3 interoperability.

Search by location

There are currently 6 functions to use for searching based on location, Table 3.

Function	Description	Parameters	Example
geo_box1	All locations within a box defined by the lower left (SW) and upper right (NE) points.	south-west latitude, south-west longitude, north-east latitude, north-east longitude	geo_box1(-20, 10, 20, 50)
geo_box2	All locations within a box defined by a centre point and a radius in km.	latitude, longitude, radius (km)	geo_box2(35, 100, 300)
geo_circ	All locations within a circle defined by a centre point and a radius in km.	latitude, longitude, radius (km)	geo_circ(35, 100, 300)
geo_lat	All locations within a latitude range given by a latitude and a radius in km.	latitude, radius (km)	geo_lat(0, 100)
geo_north	All locations north of a given latitude (inclusive).	Latitude	geo_north(80)
geo_south	All locations south of a given latitude (inclusive).	Latitude	geo_south(-80)

Table 3: ENA search by location functions.

Search by collection date

At present, the collection_date field (which maps to the MicroB3 event date/time) only supports single dates (not ranges) and does not include time. There is a plan in January 2014 to include time either within the collection date field or, as a second field, to allow higher granularity of date/time data within ENA advanced search.

This field can be searched using the format *collection_date <operator> <date>* where *operator* is one of =, !=, <, <=, >, >= and *date* is in the format DD-MM-YYYY or DD-MON-YYYY, e.g., collection_date > 01-01-2012

Example query:

http://www.ebi.ac.uk/ena/data/warehouse/search?query="geo_circ(50,-2,200)"&result=sample&fields=accession,location,collection_date,tax_id,scientific_name&display=report

Returning search results

There are several formats available for results to be returned from the ENA advanced search. Of these, the one likely to be of most interest to MicroB3 is the tab separated report format.

Tabulated report

A user can specify which attributes, and in which order, they want returned from a search using the *fields* parameter. To use this, the *display* parameter **must** be set to report, as can be seen in the example above. All columns are separated by a tab character and if the column represents a list of items, each item is separated by a semi-colon.

E.g.:

```
accession    location     collection_date     tax_id scientific_name
ERS000068    51.0 N 3.1 W 01-Jan-1969  574521 Escherichia coli O127:H6 str.
E2348/69
ERS000075    51.46 N 2.6 W01-Jan-1998  522373 Stenotrophomonas maltophilia K279a
```

6.2 Proposed additional functionality for querying ENA Advanced Search

OpenSearch based queries

ENA will support OpenSearch-formatted queries using the parameter extension, to mimic the native ENA advanced search query URLs. The format of this (still to be finalized for the OpenSearch description document) is shown below.

```
<Url
  xmlns:parameters="http://a9.com/-
/spec/opensearch/extensions/parameters/1.0/"
  type="text/html"
  template="http://www.ebi.ac.uk/ena/warehouse/search">
    <parameters:Parameter name="query" value="{searchTerms}"/>
    <parameters:Parameter name="result" value="{resultId}"/>
    <parameters:Parameter name="fields" value="{returnFields}"
minimum="0"/>
    <parameters:Parameter name="display" value="{displayFormat}"
minimum="0"/>
    <parameters:Parameter name="download" value="{downloadFormat}"
minimum="0"/>
    <parameters:Parameter name="offset" value="{startIndex}" minimum="0"/>
    <parameters:Parameter name="length" value="{numResults}" minimum="0"/>
    <parameters:Parameter name="limit" value="{maxNumResult}" minimum="0"/>
 </Url>
```

An example query using this is:

```
<Query role="example" searchTerms="geo_circ(50,-2,200)"
        result="sample" />
```

Additional geospatial function(s)

As the plan is to keep the OpenSearch URLs matching the native ENA advanced search URLs in the first instance, we will not be supporting the geospatial extension for OpenSearch queries. The BBOX function in the geospatial extension uses opposing corners to ENA's geo_box1 function, therefore we may supply a similar function to the assist easy mapping from the OpenSearch method to the ENA method. OpenSearch's geospatial extension also provides support for searching a point. We will add a geo_point(lat,lon) function to support this.

Additional output format options

ENA will add two new output formats to support the return of results in RSS format using OpenSearch response elements and GML3, for WFS servers. The names of these display options will be confirmed later.

6.3 Proposed additional results formats

RSS XML with OpenSearch response elements

ENA has agreed to add an XML-based results format to support OpenSearch response elements. We have chosen to implement RSS format rather than Atom as this maps more easily to our current html results. An example of a query response in this format is given below.

```
<?xml version="1.0" encoding="UTF-8"?>
<rss version="2.0"
      xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/"
      xmlns:atom="http://www.w3.org/2005/Atom">
   <channel>
     <title>ENA Advanced Search</title>

<link>http://www.ebi.ac.uk/ena/data/warehouse/search?query="geo_circ(50,-
2,200)&result=sample"</link>
     <description></description>
     <opensearch:totalResults>335</opensearch:totalResults>
     <opensearch:startIndex>1</opensearch:startIndex>
     <opensearch:itemsPerPage>10</opensearch:itemsPerPage>
     <atom:link rel="search" type="application/opensearchdescription+xml"
href="http://www.ebi.ac.uk/ena/data/warehouse/opensearch/opensearchdescript
ion.xml"/>
     <opensearch:Query role="request" searchTerms="geo_circ(50,-2,200)"
        result="sample" />
     <item>
       <title>ERS000068</title>
       <link>http://www.ebi.ac.uk/ena/data/view/ERS000068</link>
       <description>
```

```
         This sample has been submitted by European Bioinformatics
Institute on
         2010-02-26; Escherichia coli O127:H6 str. E2348/69
       </description>
     </item>
     <item>
       <title>ERS000075</title>
       <link>http://www.ebi.ac.uk/ena/data/view/ERS000075</link>
       <description>
         This sample has been submitted by European Bioinformatics
Institute on
         2010-02-26; Stenotrophomonas maltophilia K279a
       </description>
     </item>
     ...
   </channel>
 </rss>
```

GML3

ENA has also agreed to add an XML-based result format to support WFS servers. ENA will write an application schema to assist parsing of GML3 XML results.

Application schema

Below is a working draft of the application schema and is likely to change once development commences. Please note that the core interoperability attributes will be included in the SampleType feature.

```
<xs:schema targetNamespace=""
    xmlns=""
    xmlns:gml="http://www.opengis.net/gml"
    xmlns:xs="http://www.w3.org/2001/XMLSchema"
    elementFormDefault="qualified"
    attributeFormDefault="unqualified">
    <xs:import namespace="http://www.opengis.net/gml"
        schemaLocation="..\GML3.0.1\base\feature.xsd"/>
    <xs:element name="SampleSet"
        type="gml:AbstractFeatureCollectionType"
        substitutionGroup="gml:_FeatureCollection">
    </xs:element>
    <xs:element name="Sample"
        type=" SampleType"
        substitutionGroup="gml:_Feature"/>
    <xs:complexType name="SampleType">
        <xs:complexContent>
            <xs:extension base="gml:AbstractFeatureType">
                <xs:sequence>
                    <xs:element name="description" type="xs:string"/>
                    <xs:element name="tax_id" type="xs:integer"/>
                    <xs:element name="scientific_name" type="xs:string"/>
                     <xs:element name=="location" ref="gml:position"/>
                </xs:sequence>
            </xs:extension>
        </xs:complexContent>
    </xs:complexType>
</xs:schema>
```

Example GML3 result

Below is an example of what the GML3 result might look like, using the draft application schema above.

```
<SampleSet xsi:schemaLocation="">
    <gml:boundedBy>
        ...
    </gml:boundedBy>
    <gml:featureMember>
        <Sample gml:id="ERS000068">
            <description>100</description>
            <tax_id>574521</tax_id>
            <scientific_name>Escherichia coli O127:H6 str.
E2348/69</scientific_name>
            <location>
                <gml:Point>
                    <gml:pos>-3.1 51.0</gml:pos>
                </gml:Point>
            </location>
        </Sample>
    </gml:featureMember>
    <gml:featureMember>
        <Sample gml:id="ERS000075">
            <description>100</description>
            <tax_id>522373</tax_id>
            <scientific_name>Stenotrophomonas maltophilia
K279a</scientific_name>
            <location>
                <gml:Point>
                    <gml:pos>-2.6 51.46</gml:pos>
                </gml:Point>
            </location>
        </Sample>
    </gml:featureMember>
    ...
</SampleSet>
```

Implementation of proposed development

ENA will endeavor to add the new functionality by mid 2014. New functionalities described in this chapter may be subject to alterations when ENA has real implementation experience.

7.0 Example of Micro B3 core and enhanced interoperability implementation at EurOBIS

7.1 Current functionality of EurOBIS: Advanced Search

Currently, the distribution data in EurOBIS can be searched by scientific name, sampling date, location and data provider.

Search by scientific name

In the free text field you can type the name of the taxon you want to map. Genus and subgenus names should be included for species. If the field is left blank, you will receive distribution data for all species in EurOBIS. You can also choose the name of the taxon via a lookup which connects you to the World Register of Marine Species (WoRMS). You can choose to either or not include synonym records and/or child records.

Search by sampling date

Sampling date includes fields for start date (year, month, day) and end date (year, month, day). Also the season in which the sampling dates should occur can be specified. At present the sampling date field does not include time. If the fields are left default, you will receive distribution data for all sampling dates in EurOBIS.

Search by location

You can define the geographic location of the sampling points via the selection you made on the map, or you can fill in the coordinates manually (in degrees and minutes).

Search by data provider

EurOBIS currently holds 516 datasets, provided by 159 institutes.

Returning search results

There are several formats available for results to be returned from the EurOBIS search: (1) Tab delimited text file, (2) Table in HTML format, (3) Excel sheet (XLS), (4) XML in DiGIR format (OBIS core), and (5) KMZ file (for import into Google Earth).

7.2 Current functionality of EurOBIS: OGC WMS and WFS

Below is an example of a WFS query on the EurOBIS database, based on a specific YearCollected (2000), MonthCollected (6), DayCollected (27), bounding box (min lon = -7, min lat = 55, max lon = -6, max lat = 56) and AphiaID (141433).

```
http://geo.vliz.be/geoserver/wfs?request=GetFeature&version=1.0.0&srsN
ame=EPSG:4326&typeName=Eurobis:eurobis_points&propertyName=ObservedInd
ividualCount,SampleSize,AphiaID_accepted,YearCollected,MonthCollected,
```

```
DayCollected&FILTER=<Filter xmlns:ogc="http://www.opengis.net/ogc"
xmlns:gml="http://www.opengis.net/gml"><ogc:And><ogc:BBOX><ogc:Propert
yName>the_geom</ogc:PropertyName><gml:Box
srsName='EPSG:4326'><gml:coordinates>-7,55 -
6,56</gml:coordinates></gml:Box></ogc:BBOX><ogc:PropertyIsEqualTo><ogc
:PropertyName>AphiaID</ogc:PropertyName><ogc:Literal>141433</ogc:Liter
al></ogc:PropertyIsEqualTo><ogc:PropertyIsEqualTo><ogc:PropertyName>Ye
arCollected</ogc:PropertyName><ogc:Literal>2000</ogc:Literal></ogc:Pro
pertyIsEqualTo><ogc:PropertyIsEqualTo><ogc:PropertyName>MonthCollected
</ogc:PropertyName><ogc:Literal>6</ogc:Literal></ogc:PropertyIsEqualTo
><ogc:PropertyIsEqualTo><ogc:PropertyName>DayCollected</ogc:PropertyNa
me><ogc:Literal>27</ogc:Literal></ogc:PropertyIsEqualTo></ogc:And></Fi
lter>&outputformat=json
```

"**typeName**" refers to the layer of the VLIZ geoserver that is queried, in this case the layer Eurobis:eurobis_points.

"**propertyName**" refers to the attributes that can be included in the response. Possible attributes are: "DateLastModified", "CatalogNumber", "ScientificName", "YearCollected", "MonthCollected", "DayCollected", "Locality", "Longitude", "Latitude", "CoordinatePrecision", "MinimumDepth", "MaximumDepth", "Sex", "ObservedIndividualCount", "SampleSize", "Citation", "InstitutionCode", "AphiaID", "AphiaID_accepted", sp, imis_dasid, qc, the_geom.

"**outputformat**" refers to the format of the response. Several options are possible: GML2, GML3, Shapefile, JSON, JSONP and CSV (http://docs.geoserver.org/latest/en/user/services/wfs/outputformats.html).

"**Filter**": possible attributes that can be filtered are "DateLastModified", "CatalogNumber", "ScientificName", "YearCollected", "MonthCollected", "DayCollected", "Locality", "Longitude", "Latitude", "CoordinatePrecision", "MinimumDepth", "MaximumDepth", "Sex", "ObservedIndividualCount", "SampleSize", "Citation", "InstitutionCode", "AphiaID", "AphiaID_accepted", sp, imis_dasid, qc, the_geom (http://docs.geoserver.org/latest/en/user/styling/sld-reference/filters.html).

In the example above, observation data was requested for a certain bounding box (min lon = -7, min lat = 55, max lon = -6, max lat = 56). It is also possible to request data for a specific observation point and define a buffer around that point (e.g. a buffer of 1000m around an observation point with coordinates lon = -7 and lat = 55):

```
http://geo.vliz.be/geoserver/wfs?request=GetFeature&version=1.0.0&srsN
ame=EPSG:4326&typeName=Eurobis:eurobis_points&propertyName=ObservedInd
ividualCount,SampleSize,AphiaID_accepted,YearCollected,MonthCollected,
DayCollected&FILTER=<Filter xmlns:ogc="http://www.opengis.net/ogc"
xmlns:gml="http://www.opengis.net/gml"><ogc:And><ogc:DWithin><ogc:Prop
ertyName>the_geom</ogc:PropertyName><gml:Point
srsName='EPSG:4326'><gml:coordinates>-7,55</gml:coordinates>
</gml:Point><ogc:Distance
units="meter">1000</ogc:Distance></ogc:DWithin><ogc:PropertyIsEqualTo>
<ogc:PropertyName>AphiaID</ogc:PropertyName><ogc:Literal>141433</ogc:L
iteral></ogc:PropertyIsEqualTo><ogc:PropertyIsEqualTo><ogc:PropertyNam
e>YearCollected</ogc:PropertyName><ogc:Literal>2000</ogc:Literal></ogc
:PropertyIsEqualTo><ogc:PropertyIsEqualTo><ogc:PropertyName>MonthColle
```

cted</ogc:PropertyName><ogc:Literal>6</ogc:Literal></ogc:PropertyIsEqu
alTo><ogc:PropertyIsEqualTo><ogc:PropertyName>DayCollected</ogc:Proper
tyName><ogc:Literal>27</ogc:Literal></ogc:PropertyIsEqualTo></ogc:And>
</Filter>&outputformat=json

Below is the JSON response of the first WFS request example (with bounding box):

{"type":"FeatureCollection","features":[{"type":"Feature","id":"eurobi
s_points.fid-7ed931ad_14309024b6b_-
4be4","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1071006}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4be3","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1071007}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4be2","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1071008}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4be1","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1071009}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4be0","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1071010}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4bdf","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1084639}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4bde","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1084640}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-

```
4bdd","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1084641}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4bdc","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1084642}},{"type":"Feature","id":"eurobis_points.fid-
7ed931ad_14309024b6b_-
4bdb","geometry":{"type":"Point","coordinates":[-
6.903852,55.112316]},"geometry_name":"the_geom","properties":

{"YearCollected":2000,"MonthCollected":6,"DayCollected":27,"ObservedIn
dividualCount":null,"SampleSize":null,"AphiaID_accepted":141433,"imis_
dasid":640,"fid":1084643}}],"crs":{"type":"EPSG","properties":{"code":
"4326"}},"bbox":[-6.903852,55.112316,-6.903852,55.112316]}
```


In this example, the requested returned attributes were: ObservedIndividualCount, SampleSize, AphiaID_accepted, YearCollected, MonthCollected and DayCollected. Where the ObservedIndividualCount = null, this refers to a presence record.

7.3 Proposed additional functionalities for querying EurOBIS

Additional fields can be added to the attributes table if needed by MicroB3. For example, a field that will include the sample identifiers issued in the framework of MicroB3.

If the current OGC functionalities of EurOBIS are not sufficient to support the MicroB3 IS, OpenSearch functionalities can be implemented for EurOBIS.

8.0 Example of Micro B3 core and enhanced interoperability implementation at SeaDataNet

8.1 Current functionality of SeaDataNet: Advanced Search

The SeaDataNet Data Discovery and Access service currently has 2 human user interfaces for discovering relevant marine and oceanographic data sets. These include a Quick Search with faceted search and an Extended Search where users can make a combination of a large number of search criteria to build and refine their queries. These interfaces also include a shopping basket mechanism for requesting access to the data sets, which are managed in distributed data centres (NODCs and incl PANGAEA), and for downloading.

The Extended Search allows for searching by the following criteria:

Free search:

Search on any text field by entering a free text string.

Disciplines - Parameter groups:

Search on disciplines as included in the SeaDataNet vocabulary P08 and the related parameter groups as in P03. Multiple entries are possible.

Discovery parameters:

Search on parameters as included in the SeaDataNet vocabulary P02. Multiple entries are possible. This menu adapts to possible choices made in the Disciplines - Parameter groups menu.

Lat - Lon box:

Search via giving coordinates in WGS84 and in digital degrees. Upper-left and Lower-right corners of the Lat-Lon box.

Cruise/Station name:

Free text search on name of station or research cruise on which the observation was taken.

Projectname:

Free text search on name of project in which the observation was taken.

Datasetname:

Free text search on name of data set.

Sea regions:

Search on sea regions as included in the SeaDataNet vocabulary C16 which has an hierarchical structure. Multiple entries are possible.

Instrument type:

Search on instrument types as included in the SeaDataNet vocabulary L05. Multiple entries are possible.

Instrument depth (m) from .. to .. :

Search on the position of instruments in the depth vertical in meters. It is possible to give upper and/or lower values.

Platform type:

Search on platform types as included in the SeaDataNet vocabulary L06. Multiple entries are possible.

Waterdepth (m) from .. to ..:

Search on the local waterdepth at the site of the observations. It is possible to give upper and/or lower values.

Measuring area type:

Search by measurement area type which can be a point, curve (trajectory) or surface (polygon).

Originator:

Search on originator of the data set as included in the SeaDataNet EDMO directory (European Directory of Marine Organisations).

CDI partner:

Search on the data centre that is connected to the SeaDataNet infrastructure and that gives access to data sets. As included in the SeaDataNet EDMO directory (European Directory of Marine Organisations).

Country:

Search on the country as included in the SeaDataNet vocabulary C32 as derived from ISO.

Temporal resolution:

Search on the sampling interval as included in the SeaDataNet vocabulary L03.

Date (yyyymmdd) from .. to ..:

Search on the start and end date of the observation. It is possible to give upper and/or lower values. Also it is possible to give partial criteria such as only yyyy.

Duration ..to .. until:

Search on the duration of a timeseries as the time between start and end date – time of the observation. It is possible to give upper and/or lower values.

8.2 Current functionality of SeaDataNet: OGC WMS and WFS

OGC WMS and WFS facilitate SeaDataNet to provide CDI map layers with locations and metadata for observation data sets as included and available in the SeaDataNet CDI service. This interoperability solution is already operational and used by SeaDataNet to share CDI metadata with e.g. the SIMORC service (SIMORC = System of Industry Metocean data for the Offshore and Research Communities (http://www.simorc.org)). This is illustrated in the images below. Go to Data Access in the SIMORC portal and activate the extra layer of SeaDataNet Metocean Data. Also activate the i box and use the identification tool to retrieve the CDI metadata. A direct exchange takes place with the WMS - WFS services as provided by the SeaDataNet CDI portal service. By WMS protocol SIMORC users can activate maps to see the locations of the SeaDataNet data sets. Users can click on sites on the map to retrieve by WFS protocol metadata of that site. Moreover it includes URLs back to the SeaDataNet CDI details.

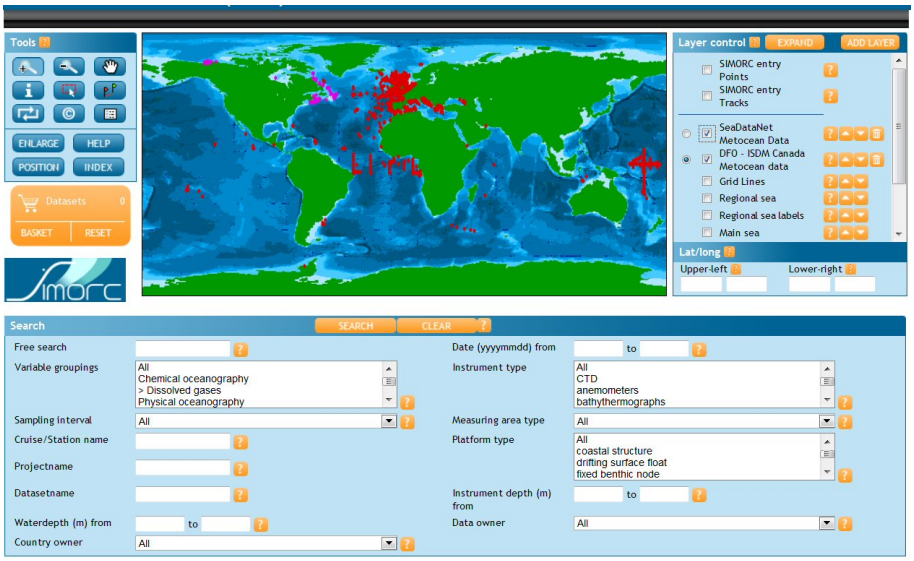

Figure 3: SIMORC User Interface with extra map layer activated for SeaDataNet metocean sites.

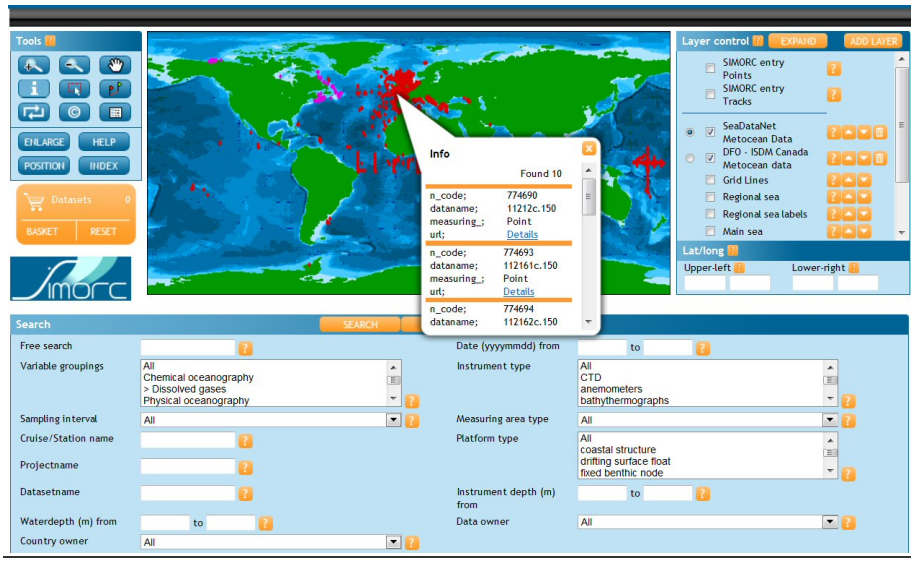

Figure 4: SIMORC User Interface with extra map layer interrogated by WFS to display metadata of SeaDataNet sites.

The SeaDataNet metadata as displayed above include also dedicated URLs to additional metadata and data access as provided by the SeaDataNet CDI service. Clicking on the URL will give a link to SeaDataNet for the selected site and data set.

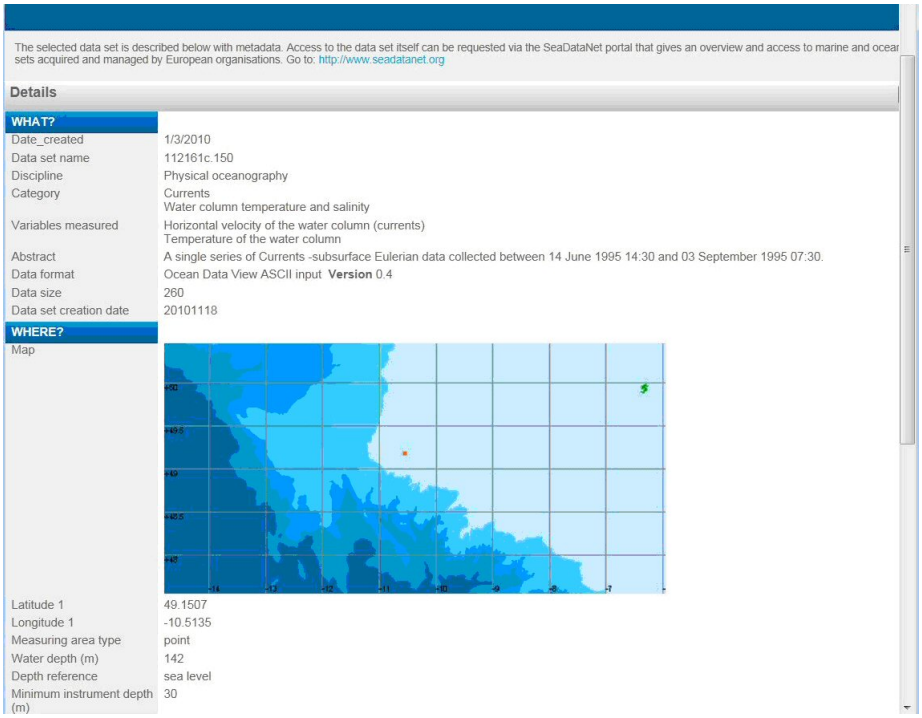

Figure 5: Extra metadata as retrieved from SeaDataNet for selected site and data set.

Another example is the exchange of WMS - WFS to the EMODNet portals, such as for Chemistry and Hydrography. In the case of EMODNet Hydrography the CDIs of bathymetric surveys are shared into the hydrography viewer service and it is possible to make a Lat - Lon search box in the hydrography viewer, which returns the CDI interface and results for that CDIs. This is illustrated in the following figures.

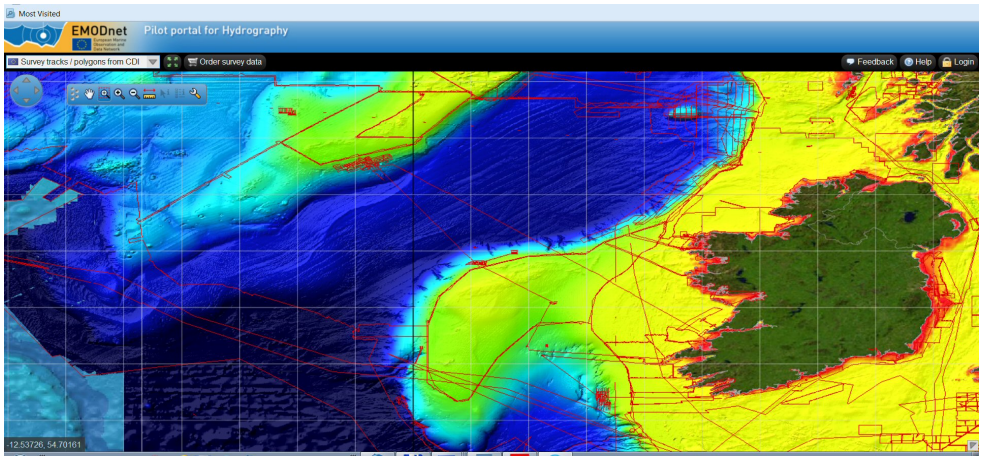

Figure 6: EMODNet Hydrography Viewer User Interface with extra map layer to display metadata of SeaDataNet bathymetric surveys.

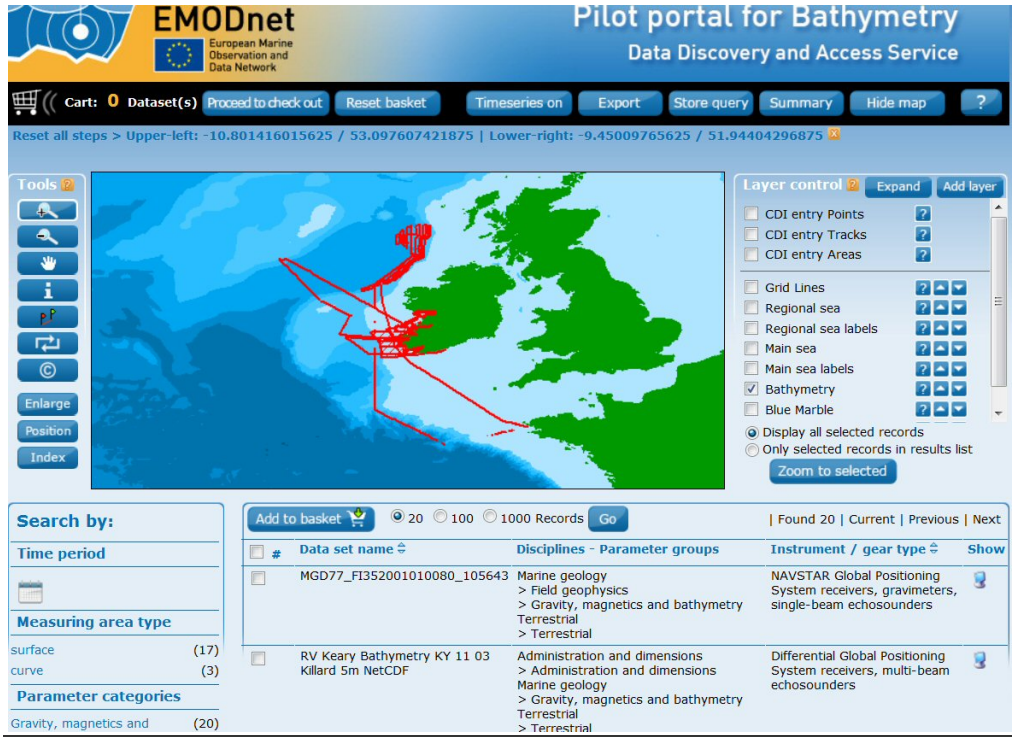

Figure 7: Search on Lat Lon box in EMODNet Hydrography Viewer User Interface returns SeaDataNet CDI interface with CDI results displayed in EMODNet Hydrography look & feel.

The WMS and WFS can be made for various collections of CDIs, following a query profile.

An example of the URLs is given below as in use for SIMORC:

http://geoservice.maris2.nl/wms/seadatanet/cdi_v2/simorc

Getcapabilities

http://geoservice.maris2.nl/wms/seadatanet/cdi_v2/simorc?service=WMS&request=GetCapabilities

an example URL for WMS:

http://geoservice.maris2.nl/wms/seadatanet/cdi_v2/nodc?styles=&format=image/png&transparent=true&request=getmap&version=1.1.1&srs=EPSG:4326&layers=points&width=580&height=290&bbox=-180,-90,180,90

an example URL for WFS:

http://geoservice.maris2.nl/wfs/seadatanet/cdi_v2/simorc?service=WFS&version=1.0.0&request=getfeature&outputformat=gml3&typename=points&maxfeatures=10&bbox=-0.2793103448275872,57.4448275862069,1.5206896551724127,59.244827586206895

Below is also the XSLT which we use to transform the data to HTML:

```
<?xml version="1.0" encoding="utf-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
        xmlns:ms="http://mapserver.gis.umn.edu/mapserver"
        xmlns:gml="http://www.opengis.net/gml"
        xmlns:ogc="http://www.opengis.net/ogc"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
<xsl:param name="pagelen" select="25"/>
<xsl:template match="/">

        <xsl:if test='count(//gml:featureMember/* | //gml:featureMembers/*) > 0'>

                <table width='99%' border='0' cellspacing='0' cellpadding='0'>
                <tr>
                  <xsl:choose>
                        <xsl:when test='count(//gml:featureMember/* |
//gml:featureMembers/*) = 0'>
                          <td class='buttons' align='left'>
                                No results found.
                          </td>
                        </xsl:when>
                        <xsl:otherwise>
```

```
                        <td class='buttons' align='right'>
                                <div id='browse_buttons'>
                                 Found <xsl:value-of
select='count(//gml:featureMember/* | //gml:featureMembers/*)'/>
                                </div>
                        </td>
                       </xsl:otherwise>
                 </xsl:choose>
                </tr>
                </table>

<!-- Resultaten Mapserver: gml:featureMember, Geoserver: gml:featureMembers  -->

                <xsl:for-each select="//gml:featureMember/* |
//gml:featureMembers/*">
                <table class='infowindow-browse-table' border='0' cellpadding='0'
cellspacing='0'>
                 <tr class='browse-subject'>
                        <td colspan='2'>&#32;</td>
                 </tr>

<!-- Laat alles zien behalve de GML (sla elementen over die een element gml:* bevatten) -->

                <xsl:for-each select="*[not(gml:*)]">
                        <xsl:element name="tr">

                         <td class='browse_data'>
                                <xsl:value-of select="local-name(.)"/>;
                         </td>

                         <td class='browse_data'>


                                <xsl:choose>

<!-- maak url clickable -->
                                 <xsl:when test='starts-with(., "http://")'>
                                        <xsl:element name='a'>
                                         <xsl:attribute name='href'>
                                                <xsl:value-of select="."/>
                                         </xsl:attribute>
                                         <xsl:attribute
name="target">_blank</xsl:attribute>
                                         Details
                                        </xsl:element>
                                 </xsl:when>
                                 <xsl:otherwise>
                                        <xsl:value-of select="."/>
```

```
                                    </xsl:otherwise>
                                </xsl:choose>
                            </td>
                        </xsl:element>
                </xsl:for-each>

                    </table>
        </xsl:for-each>

                </xsl:if>

</xsl:template>

</xsl:stylesheet>
```

Follow-up action:

The CDI User Interface at present has an option to store the users search profile of selection criteria and configuration settings of active map layers as a bookmark and to share this bookmark with other users, facilitating repeated searching in a dynamic way. It is planned to extend the bookmark also with a dynamic generation of the WMS service so that users can make their own WMS - WFS layers for specific searches. This will be implemented by MARIS during 2014.

8.3 Proposed additional functionalities for querying SeaDataNet

OGC CS-W:

OGC CS-W provides the option to serving out INSPIRE compliant CDI files for wide interoperability with other metadata portals. This concerns metadata while the access to data will be realised by including dynamic URLs towards the SeaDataNet User Interface.

In an earlier version of the CDI User Interface this was already implemented by using GI-CAT software service that facilitated the conversion of SeaDataNet CDI XMLs in the old 19115 DTD format into INSPIRE ISO 19139 XML files that were shared by GI-CAT CS-W as part of the EuroGEOSS Brokerage service towards the GEOSS portal. In that case SeaDataNet provided by SOAP web service a set of aggregated CDI XML files. The aggregation was done per CDI data centre and per parameters discipline (P08) resulting in ca 500 collections, while the XML contained URLs towards the related subsets of the 1.3 million CDI granules.

Recently the XML output of SeaDataNet has been upgraded to the new SeaDataNet ISO 19139 Schema and INSPIRE compliance and the EuroGEOSS Brokerage Service exchange has been amended. This works now fully operational and dynamic.

http://seadatanet.maris2.nl/gi-cat-seadatanet/sdn-cdi-aggr-seadatanet_v3.xml

The same exchange has been established and is operational for the interoperability of SeaDataNet with the IODE Ocean Data Portal (ODP).

Follow-up action:

MARIS will expand the existing interoperability services for the CDI service with a CS-W, using GeoNetWork, to provide a wider interoperability to external portals. This will work again with CDI collections, facilitating a 2-step approach for external discovery services to identify the offer of SeaDataNet and to bring users to the SeaDataNet portal for more detailed discovery and the actual data shopping. The GeoNetWork CS-W is being configured for the SeaDataNet ISO19139 XML format and this will be finalised end January 2014.

OpenSearch:

The OpenSearch request interface is simple, consisting of a description of a HTTP GET request with a series of optional key-value parameters that can be used to constrain the search:

- Free search
- Geospatial (area or point + radius)
- Temporal (from to)

The OpenSearch protocol is rather simple: No use of vocabularies, no special search fields, which has a disadvantage when querying a very large collection with specific data like SeaDataNet. It is necessary to split the total metadata in many OpenSearch access points (virtual aggregations).

Format HTTP get:

<Url type="text/html"

template="http://example.com/?q={searchTerms?}&pw={startPage?}"/>

Example:

http://www.google.com/?q=question

http://www.google.com/?q={searchTerms}

SeaDataNet already has experience with OpenSearch: an application prototype for OpenSearch has been set up as a result of work in the EU FP7 Genesi-DEC project. An OpenSearch communication layer has been developed on top of CDI database. This prototype is running only for the Geo-Seas subset as yet.

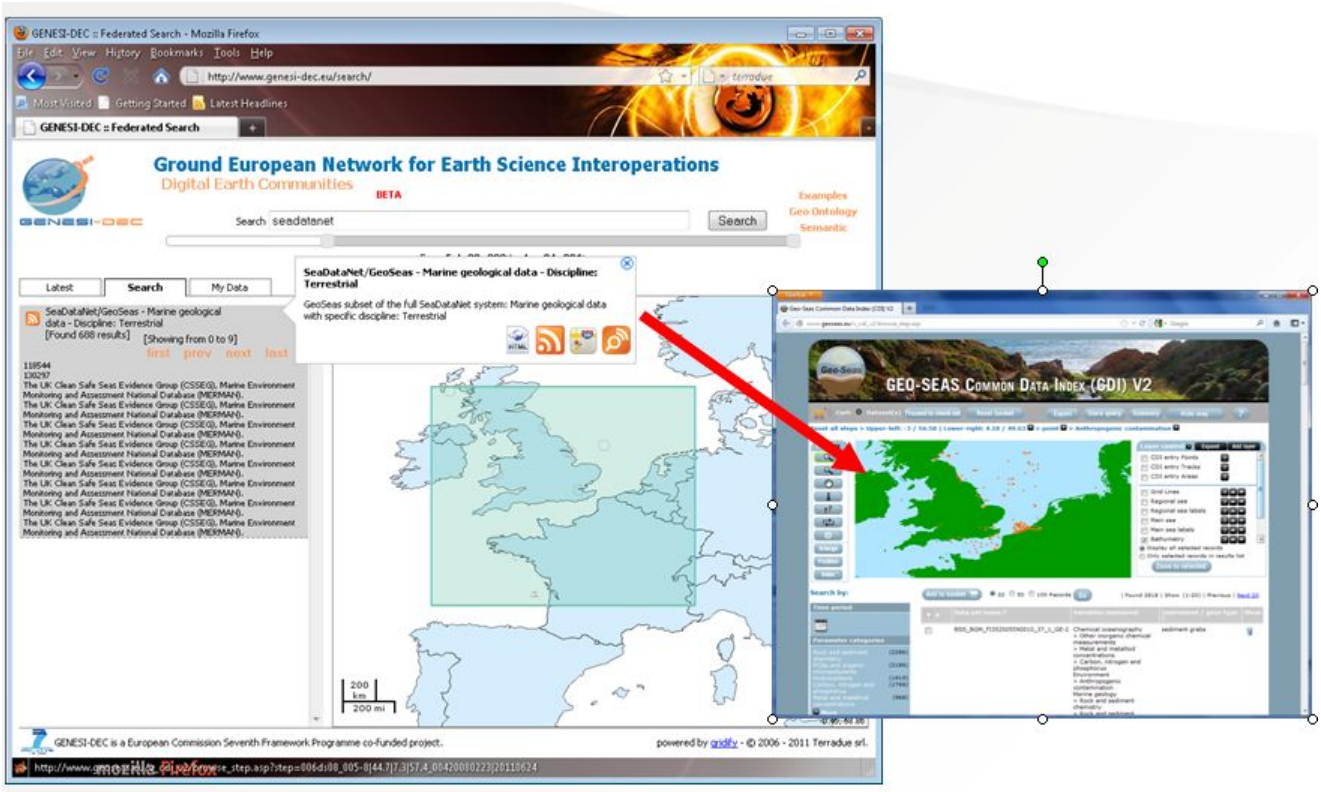

Figure 8: Search result of OpenSearch in Genesi-DEC portal with link to Geo-Seas CDI service for further ordering of related data sets.

OpenSearch entry points have been created in RDF and registered to the Genesi-DEC portal (www.genesi-dec.eu). Each RDF file describes an access point of an aggregation of the marine geological data (first filter) and then per discipline (P08).

To provide interoperability towards other portals an identical approach can be made at the level of collections or granules. The downside of going for granules is that the virtual series will contain a high number of metadata records (> 100.000). And the OpenSearch protocol is limited in selective search fields. To avoid too many results the best solution is to go for collections by applying an aggregation per parameter or instrument or data provider when enabling OpenSearch entries. Following this method the SDN CDI directory could again be split up in a couple of hundred entry points, which would make the exchange more feasible in practice.

Follow-up action:

It is planned to implement an OpenSearch interoperability service on top of the CDI service at collection level for Micro B3. This will be implemented by MARIS in the first part of 2014.