PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

SeaDataNet

User manual and instructions for compiling the CDI metadata, coupling table and associated data

Deliverable 5.7

Deliverable number		Short Title	
D5.7		Guidelines for the CDI Data Discovery and Access Service	
Long title			
User manual and instructions for compiling the CDI metadata, coupling table and associated data			
Short description			
The document provides guidelines to data centres on how to maintain the SeaDataNet CDI Data Discovery and Access Service, especially how to submit new entries or update existing ones.			
Author		**Working group**	
D.M.A. Schaap		MARIS, BODC, BSH-DOD, IFREMER	

History

Version	Authors	Date	Comments
1.0	D.M.A. Schaap	6/1/2012	
1.1	M. Fichaut	9/2/2012	Text contribution
1.2	D.M.A. Schaap	29/2/2012	Final document
1.3	F. Gheorghe, D.M.A. Schaap	21/11/2013	Updates for the 1st Innovation Cycle

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

2

Table of contents

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

3

Purpose of the document

This document provides guidelines to the SeaDataNet partners on how to maintain the content and the consistency of data deliveries of the Pan-European CDI Data Discovery and Access Service. It describes how to prepare and submit new CDI metadata and how to update existing ones. It also describes how to maintain the local coupling table that is part of the CDI system and that provides the links from the central CDI entries to the locally stored data.

It does not intend to give instructions for the usage of the common software tools that have been developed within SeaDataNet for editing and generating the CDI XML metadata entries, editing of the coupling table and possible conversion of local data sets to the SeaDataNet common data formats. This is covered by the user's manuals of MIKADO and NEMO software tools which can be retrieved from the SeaDataNet portal in the Software section.

Its purpose is to integrate in one document the available options and methods for the maintenance of the CDI service. These practices are already being used by the "old" partners while new data centres may use it in a complementary way with the already existing information at the SeaDataNet Portal [1, 2, 4], the manuals [14 - 18] and the training workshops material [3].

Note: Guidelines for maintenance of the other metadata service (EDMED, EDMO, EDMERP, EDIOS and CSR) are given in a separate document D4.6

1. Introduction

1.1. Metadata directories overview

SeaDataNet maintains and operates an on-line metadata discovery system (http://www.seadatanet.org/Metadata) with overviews of marine organizations in Europe and their engagement in marine research projects, managing large datasets and data acquisition through research cruise and monitoring programmes and networks for the European waters and the global oceans:

- **EDMO**: European Directory of Marine Organizations (at present > 2100 entries)
- **EDMED**: European Directory of Marine Environmental Data (at present > 3800 entries)
- **EDMERP**: European Directory of Marine Environmental Research Projects (at present > 2500 entries)
- **CSR**: Cruise Summary Reports (at present > 32000 entries)
- **EDIOS:** European Directory of the Ocean Observing Systems (at present > 10000 entries)

Moreover the portal features the CDI Data Discovery and Access service to facilitate searching for marine data sets as managed by the connected data centres and requesting and downloading data sets. (http://www.seadatanet.org/Data-Access)

A schematic overview of the metadata directories and their interconnections is shown in the Figure (1):

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
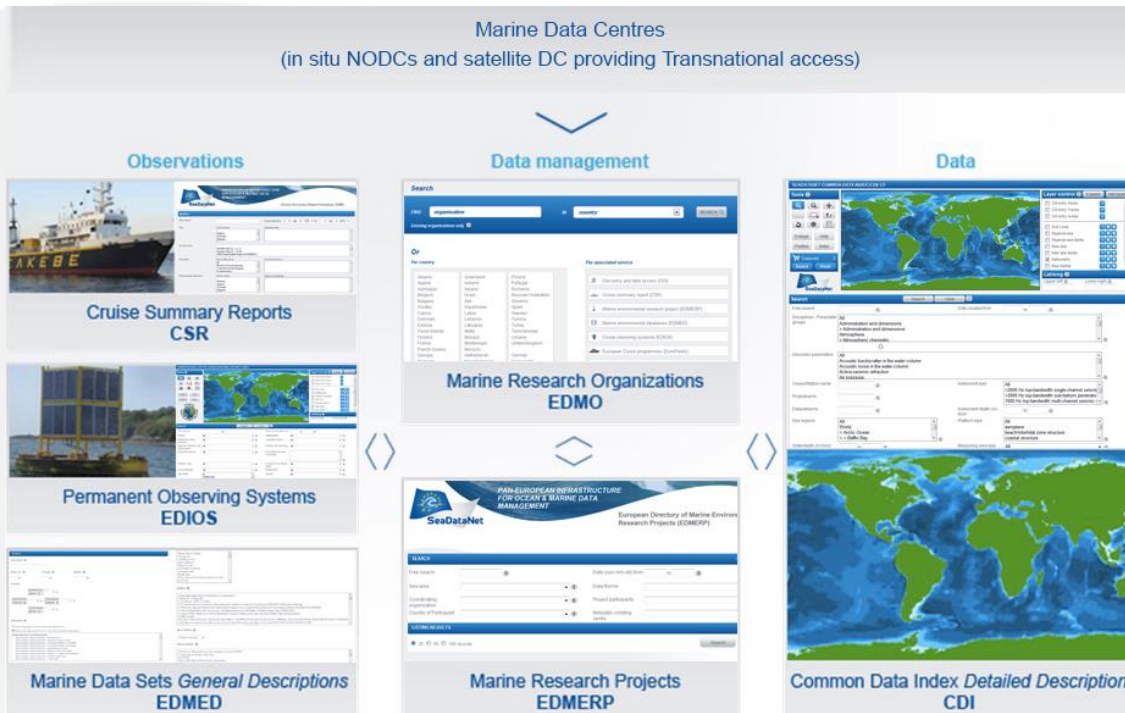sdn-userdesk@seadatanet.org – www.seadatanet.org

4

Figure 1: Metadata Directories Overview

The access to these services can be found at the SeaDataNet portal at:
http://www.seadatanet.org/Metadata

The **Common Data Index** (**CDI**) is the key Data Discovery and Delivery service that provides online unified access to datasets managed by the distributed Data Centres (at present >1000000 entries). It gives users a highly detailed insight in the geographical coverage, and other metadata features of partners data holdings. Users can have access to identified datasets in a harmonized way, using a shopping basket mechanism. They can follow the processing of requests via an online transaction register and can download datasets in the SeaDataNet standard formats. The access to these services can be found at the SeaDataNet portal at:
http://www.seadatanet.org/Data-Access/Common-Data-Index-CDI

In addition, SeaDataNet maintains and operates Common Vocabularies for parameters, platforms types, instruments, etc. which are available as Web services and are used to mark up metadata, data and data products in a consistent and coherent way.
At present the Vocabulary Services comprise over 120000 terms in over 100 lists. More information and access can be found at the SeaDataNet portal at:
http://www.seadatanet.org/Standards-Software/Common-Vocabularies

1.2. Background and future upgrades

At the beginning of the SeaDataNet Project (2006-2007), the metadata directories had different background, set-up and structure. Many of the descriptions were in free text format preventing thus the cross search of information between the different metadata bases. The CDI service provided access to the data holdings in different formats through the partner's interfaces or through e-mail requests to the individual data centres. This SeaDataNet Version 0 was actually a maintenance and continuation of the metadata system as developed earlier in the Sea-Search Project.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

5

During 2008-2010 a major effort took place in order to set up an integrated on-line data service by interconnecting, harmonizing and mutually tuning the metadata directories in terms of formats, syntax and semantics. To this end the ISO 19115 metadata standard was adopted for all directories -and not only for the CDI - , ensuring thus interoperability between the metadata directories on behalf of the marine community. Common Vocabularies, Projects, Organizations and CSR References were used in the metadata descriptions. In addition the metadata became accessible and linked through Web Services to enable constant accessibility of the latest versions. The users could search for data sets in the CDI V1 data base, order and get access by downloading their requested data in a unique format via a shopping basket mechanism and a communication between the SeaDataNet portal and the distributed data systems (SeaDataNet Version 1).

At the end of SeaDataNet I project, also the content of EDMED, EDMERP, EDMO and CDI directories have been completely converted to V1 format. The upgrade of the older CSR V0 entries to the V1 content model as well as the update of EDIOS content in cooperation with EuroGOOS will be continued and completed within the SeaDataNet-II project.

During the last year of the SeaDataNet Project (2010-2011), the infrastructure was improved by adding data product services, OGC compliant viewing services and extending the ability to access various marine and ocean data types **(**CDI Version 2).

It is the goal of SeaDataNet-II (2011-2015) to upgrade the infrastructure further into an operational robust and state-of-the-art Pan-European system providing up-to-date and high quality access to ocean and marine metadata, data and data products. In November 2013 a major upgrade has been launched (CDI V3) which makes use of the ISO 19139 Schema, is compliant to INSPIRE and is supported by the new version (NVS 2.0) of the SeaDataNet Common Vocabularies. Also a further harmonisation and relationship of the CDI metadata format has been achieved. Underway is a further upgrading by making the exchange of CDI XML entries from data centres to the central CDI manager more efficient. The metadata will be created/ updated at the local data centres as usual but the exchange with the central CDI catalogue will become operational. e.g. the metadata will be automatically harvested by the central system via OGC based Catalogue Services for the Web (CS-W).

2. Overview of the CDI Data Discovery and Access Service

2.1. Governance of SeaDataNet directories and CDI in specific

A simplified schema of the overall metadata workflow is shown in Figure (2). In each country contributions for each of the directories originate from a wide range of local organisations which are gathered and validated at the national scale by the SeaDataNet national nodes (usually NODCs). Depending on the local settings, these use MIKADO or the online CMS tools for preparing and forwarding manually their national entries to the managers of the central European directories: BODC manages EDIOS, EDMED and the Common Vocabularies; BSH-DOD manages CSR; MARIS manages EDMERP, EDMO.

Note ► *The cooperation between the local organisations and the national collators is very important because it ensures the effective metadata flow and management from the national systems to the central Pan-European archives.*

The roles of the partners are summarized as follows:

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

6

The national collators are responsible for the content of the directories and they co-ordinate the overall management activities at national level. They are the sole responsible for submitting XML files to the central directories. More specific:

- Provide assistance in metadata preparation
- Collate new records or their updates
- Quality control of metadata descriptions
- Distribute to the central portal in case of XML files
- Contact with central directory managers

The role of the other organizations is to support and complement the coordination activities by:

- Provide new entries
- Complete the information
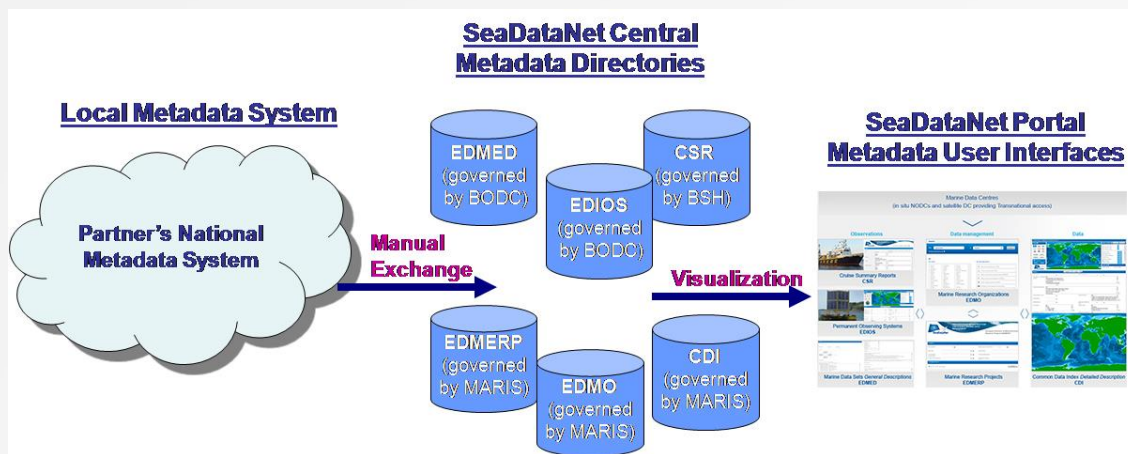- Update existing information

Metadata workflow

Figure 2: General workflow of the metadata directories

For each directory a user interface has been developed which enables the on-line search of the metadata bases and the display of the query on the screen. These metadata viewing services can be reached from the metadata overview page of the SeaDataNet portal: http://www.seadatanet.org/Metadata

However for the Common Data Index (CDI) service the roles are somewhat different. The CDI gives an index and access to data sets that are managed in a distributed way by individual institutes/ data collators. Therefore each institute itself is responsible for connecting and operating its connection of its local data management system to the SeaDataNet infrastructure and for generating and submitting CDI references for its own data to the central European system. But their national NODC will give assistance and support to establish the connection and to prepare and maintain good quality CDI entries and data access. The pan-European CDI system is managed and operated by the SeaDataNet partner MARIS, who also runs a CDI-support desk together with IFREMER to give support.
http://www.seadatanet.org/Data-Access/How-to-contribute

2.2. How does the CDI Data Discovery and Access Service function?

The **Common Data Index** (CDI) service provides the link between discovery and delivery of data sets. The CDI metadata gives users a detailed insight of the availability and geographical

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

7

distribution of marine data, archived at the connected data centres. It provides sufficient information to allow the user to assess the data relevance.

The CDI Service at the SeaDataNet Portal provides registered users access to these distributed data sets, managed by connected data centres. The access consists of the ability to download data sets in common formats. This is arranged via a shopping mechanism.

At present there are 2 CDI User Interfaces which can be reached from the CDI overview page of the SeaDataNet portal: http://www.seadatanet.org/Data-Access
- CDI service – Quick Search
- CDI service – Extended Search

Both discovery interfaces enable users to search for data sets by a set of criteria. The selected data sets are listed by CDI metadata records. Geographical locations are indicated on a map. Clicking on the display icon retrieves the full CDI metadata of the data set. This includes standardised information on the data access restrictions, that apply. The interfaces also feature a shopping mechanism, by which selected data sets can be included in a shopping basket.
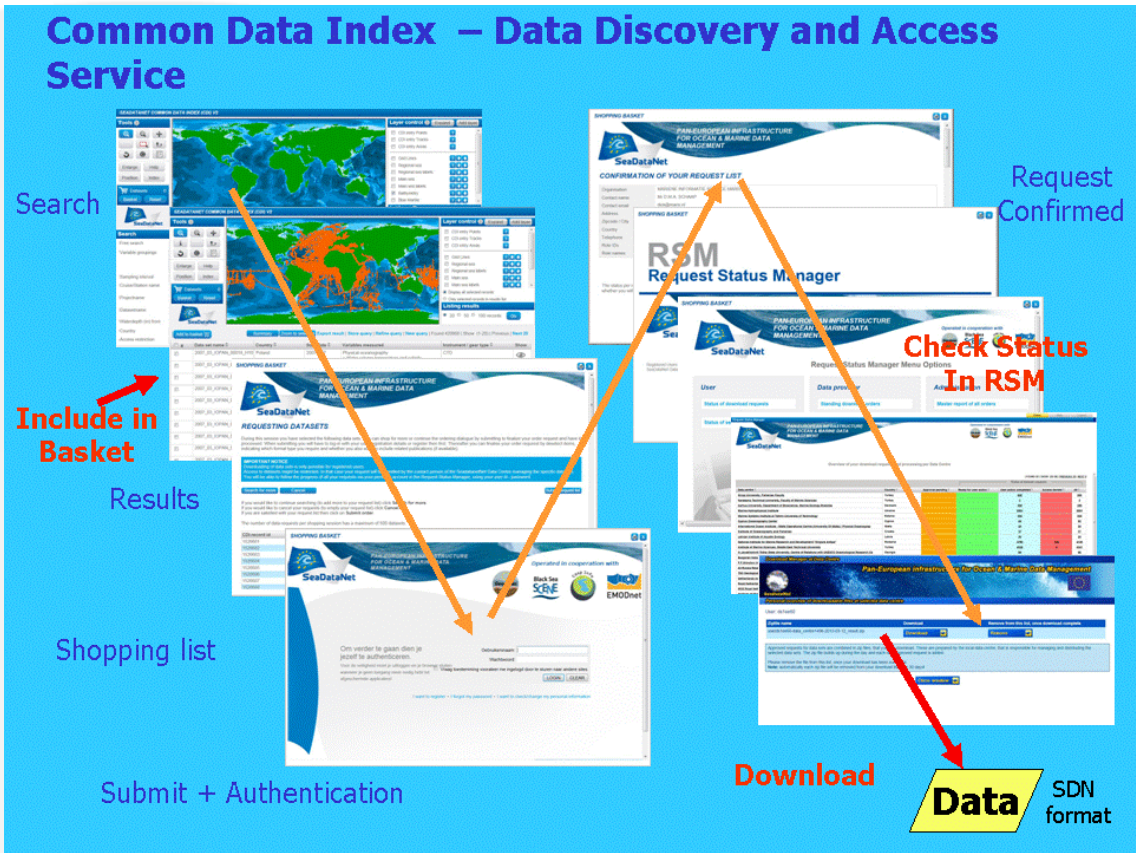

All users can freely query and browse in the CDI directory; however submitting requests for data access via the shopping basket requires that users are registered in the **SeaDataNet Central User Register**, thereby agreeing with the overall **SeaDataNet User Licence**.

All data requests are forwarded automatically from the Portal to the relevant data centres. This process is controlled via the **Request Status Manager (RSM) service** at the portal, that communicates with the data centres via the **Download Manager** (**DM**) java software module, implemented at each of the data centres.

Users receive a confirmation e-mail of their data set requests and a link to the RSM service. By log-in to the RSM service users can check regularly the status of their requests and download data sets from the associated data centres, after access has been granted. Each CDI metadata record includes a data access restriction tag. It indicates under which conditions the data set is accessable to users. Its values can vary from 'unrestricted' to 'no access' with a number of values in between. During registration every user will be qualified by its national NODC / Marine Data Centre with one or more SeaDataNet roles. The RSM service combines for each data set request the given data access restriction with the role(s) of the user as registered in the SeaDataNet central user register. This determines per data set request, whether a user gets direct access automatically, whether it first has to be considered by the data centre, that therefore might contact the user, or that no access is given.

The following image gives an overview of the dialogue for users that use the SeaDataNet portal. Related projects, such as Geo-Seas, Upgrade Black Sea SCENE and the EMODNET pilots have their own portals and CDI user interfaces, but the processing of the data requests is handled in the same way, using the common SeaDataNet CDI database and RSM service.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

8

Common Data Index — Data Discovery and Access Service

The actual delivery of data sets is done between the user and the selected data centre. This is illustrated below for users that use the SeaDataNet portal.
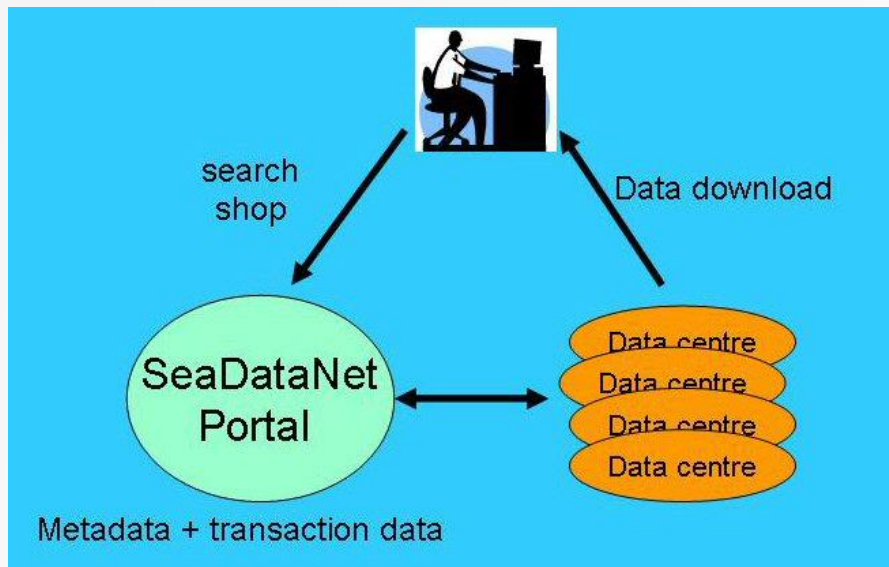
2.3. Technical components of the CDI service

The technical configuration for the CDI Service at the Portal and connected Data Centres is illustrated in the following image.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

9
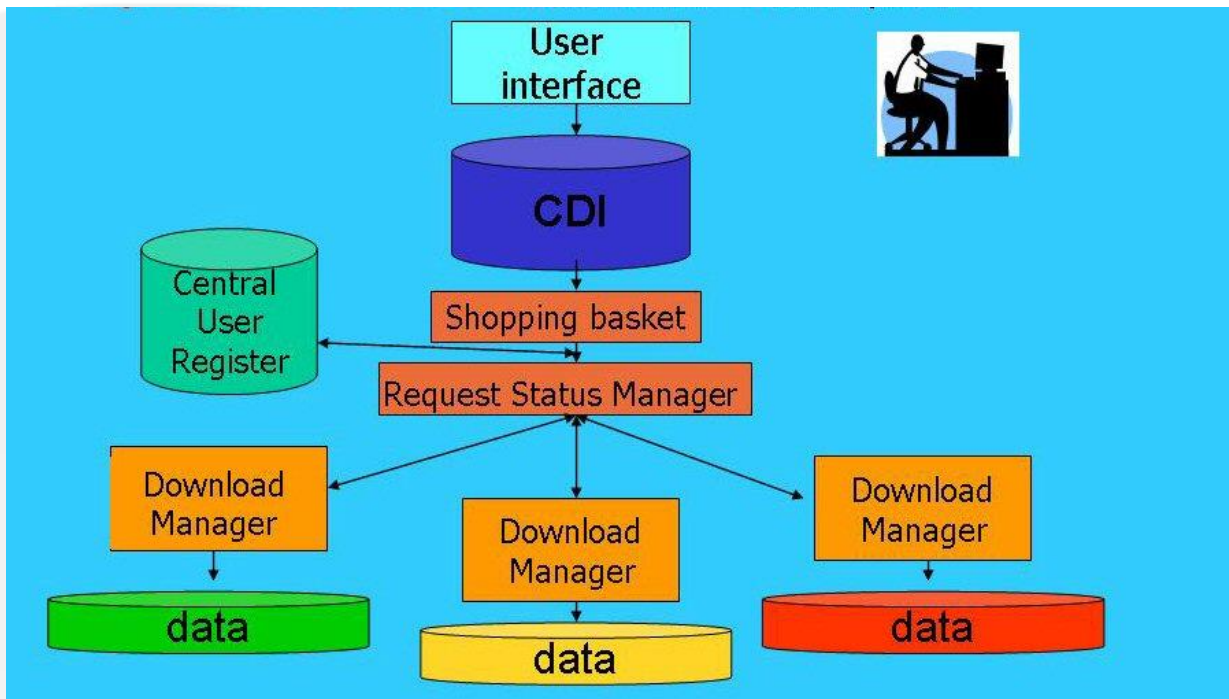
It comprises the following components:

- CDI Metadatabase and Discovery User Interface at the Portal :
  Free searching and browsing by users of metadata of data sets, managed by connected Data Centres. This can have a dedicated look & feel and with a focus on specific data centres and types of data per Portal (there are Portals for SeaDataNet, Geo-Seas, Upgrade Black Sea SCENE and the EMODNET pilots)

- SeaDataNet Central User Register - Web service:
  Containing details of users, their organizations and addresses, license period, Id-Passwords, and Roles, as administered by NODCs after registration submission by users.

- Shopping Basket – User Interface:
  Integrated in the CDI user interfaces and enabling users to prepare a user request of multiple data sets. To submit their shopping requests users have to log-in with their Id-Password. The component then handles the validation of users and their roles via checking the Central User Register, validates the combinations of user role and data access restrictions (via the access decisioning matrix) and routes requests with their resulting access status to the Request Status Manager.

- Request Status Manager – Web service + User Interface:
  Processes and administers all requests and data deliveries (downloads), for users in communication with the connected data centres via their installed Download Manager components. Data Centres use the RSM to oversee all transactions and to indicate their decisions for data requests that have received the status 'to be discussed' following the access decisioning matrix.

- Data Centre Download Manager – Java Tool:
  Is installed and configured at each Data Centre. It handles the communication with the Request Status Manager and takes care that requested files are made ready for downloading by User (if OK) in the agreed formats via a local website address at the Data Centre. The Download Manager makes use of the Coupling Table as a way to

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

10

make a uniq link between the CDI Metadata at the Portal and the locally managed data files. The data files must be in one of the SeaDataNet exchange formats (SDN ODV4, SDN MedAtlas, SDN NetCDF (CF), ..) and can be in stored in a local file management system in pre-processed formats. Alternatively data centres can include in the Coupling Table the database queries to have the Download Manager retrieve and convert to SDN ODV4 format the requested data sets from local databases.

2.4. Installation and configuration as SeaDataNet CDI provider

To connect to the SeaDataNet infrastructure as data provider via the CDI service requires that a local institute installs and configures the Download Manager software. This is a java software component that can be found at the SeaDataNet Extranet together with technical documentation and an installation guide. The Download Manager will handle the communication with the Request Status Manager (RSM) at the portal and takes care that requested files are made ready for downloading by users (if OK) in the agreed formats via a local website address at the Data Centre. The Download Manager makes use of the Coupling Table as a way to make a unique link between the CDI Metadata at the Portal and the locally managed data files or databases.

Data centres are advised to contact MARIS and its special support desk cdi-support@maris.nl for guidance with the installation and configuration. Part of the configuration is the registration of the new data centre and its Download Manager at the RSM portal, which will be undertaken by MARIS.

Data centres are advised to read and to follow up carefully the DM installation and configuration instructions, because it is a precise activity. The well-functioning of the Download Manager installation and configuration will be tested via the CDI Test service, which includes a test CDI database, test Shopping mechanism and test RSM. The test system can thus be set-up as a complete system,that can be switched to operation via a switch in the RSM configuration file. This enables also parallel functioning of the CDI test service and CDI production service, in case of Download Manager upgrades. In that situation the new system must be configured and tested, while the older production version is not hampered and can stay operational for users.

The Test Request Status Manager webinterface can be found at:
http://seadatanet.maris2.nl/request/welcome_test.html

The Test CDI User Interfaces can be found at:
http://seadatanet.maris2.nl/v_cdi_v3_test/search.asp
http://seadatanet.maris2.nl/v_cdi_v3_test/browse_step.asp

MARIS will support data centres that undertake the Download Manager installation and configuration. The DM has several components,that can be followed and tested one by one to achieve the required operation. First the communication set-up is tested between the RSM and the local DM. Thereafter the full cycle is tested from discovery to request to data download. This includes a good check on the Coupling Table which must give a unique relation between the CDI metadata and the associated local data files or data queries.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

11

3. Preparing and maintaining CDI entries

3.1. General considerations

The CDI entries have to be generated following the CDI metadata format. This format is based upon the ISO 19115 - 19139 standards and described in detail in the CDI documentation [4] which can be found at the SeaDataNet portal:
http://www.seadatanet.org/Standards-Software/Metadata-formats

MIKADO is an XML editor tool that is maintained by SeaDataNet and that can be freely downloaded from the SeaDataNet portal. It is dedicated to support the manual and automatic generation of CDI XML files. It has to be used automatically if a user has a high number of CDI descriptions, or to be used manually if only a small number of CDI descriptions have to be produced. Download MIKADO and other relevant tools from:
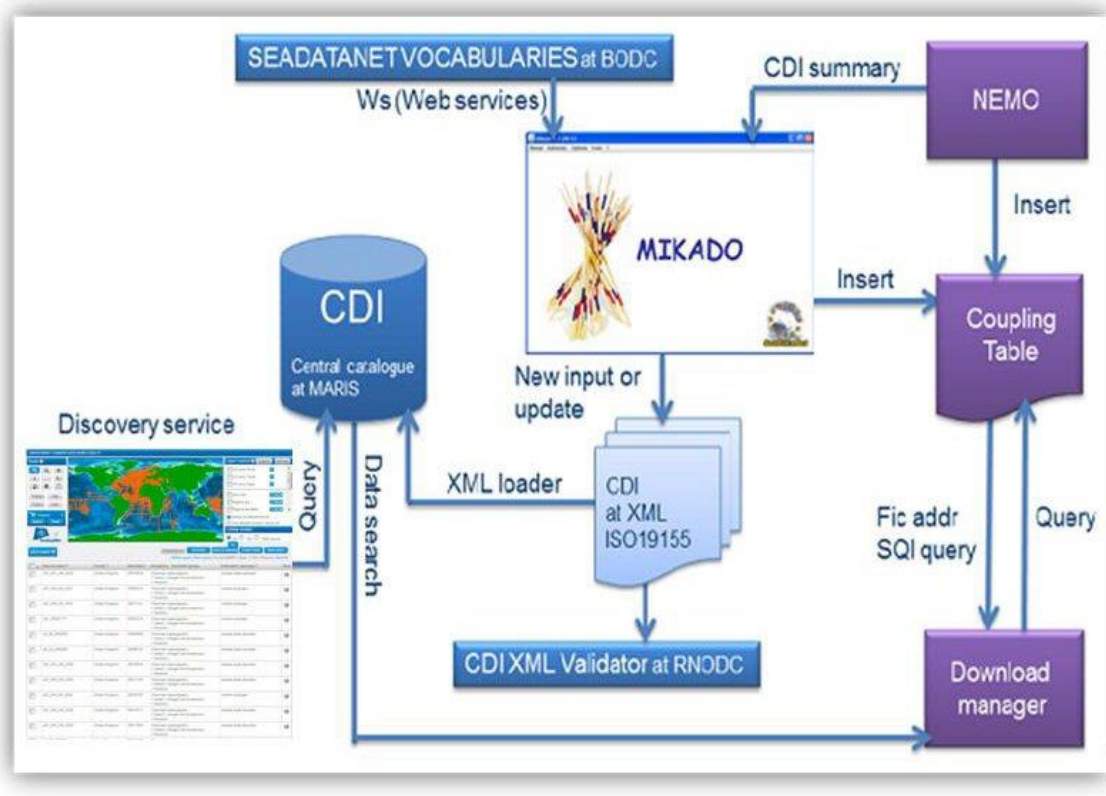http://www.seadatanet.org/Standards-Software/Software

How to use MIKADO is well explained in the MIKADO user manual which can be downloaded next to the MIKADO software. However the next paragraphs will explain the principles and the main components for preparing and submitting CDI entries in association with the related data sets.

The following figure represents the maintenance options and tools.

New CDI XML entries for the CDI service have to be delivered to MARIS as central task manager of the CDI service. Best is to contact the support desk: cdi-support@maris.nl.

For the maintenance of the CDI catalogue, the central tool is MIKADO which generates the XML descriptions of the data managed by NODCs or data focal points. MIKADO is linked to other important components of the SeaDataNet system like NEMO and the coupling table. These links are described in the following paragraphs.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

12

Maintenance options and tools

3.2. Short descriptions of the main tools and components

3.2.1. NEMO

NEMO is a software freely downloadable from the SeaDataNet web site, used for the reformatting of ASCII data files to SeaDataNet format files (http://www.seadatanet.org/Standards-Software/Software/NEMO). The NEMO user manual is available at the same address on SeaDataNet web site.

The SeaDataNet standard formats are ODV, MEDATLAS and NetCDF. The latest version supports to reformat ASCII files to any of these formats.

NEMO has to be installed on the user computer.

3.2.2. MIKADO

MIKADO is a software freely downloadable from SeaDataNet web site, used for the production of XML descriptions of SeaDataNet catalogues CSR, CDI, EDIOS, EDMED, EDMERP, O&M and SensorML (http://www.seadatanet.org/Standards-Software/Software/MIKADO). MIKADO user manual is available at the same address on SeaDataNet web site.

MIKADO has to be installed on the user computer.

3.2.3. XML Validation

The new CDI XML format is validated by the central manager MARIS. However data centres are requested to submit valid CDI XML entries. XML files, generated using the latest MIKADO software, will be valid and should parse to the associated Schema. However partners not making use of MIKADO but generating XML entries for CDI should perform a validation before

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

13

they can prepare and submit regular contributions. The Schema for CDI includes Schematron rules which allow to validate both the Syntax and the Semantics of CDI XML files, using an XML editor (e.g. Oxygene, XML Spy, ..) and the related Schema, which can be found at the SeaDataNet portal in the Standards and Formats section. If you are online, the Schema should be found by the editor automatically at the SeaDataNet namespace..

3.2.4. Download Manager

As explained in chapter 2, the Download Manager is a key software of SeaDataNet that should be implemented in each SeaDataNet node (NODC or data provider) and it is essential for the communication between the local system and the portal and the users.

The principle of this software is that it is a background service which waits for data requests coming from the SeaDataNet portal (via the RSM service). A data request contains a list of LOCAL_CDI_Ids ((local unique identifier of a CDI record) to be downloaded; each of these LOCAL_CDI_Ids are recorded in the coupling table to relate to asociated data sets or data queries.

The documentation about the Download Manager can be found at:
http://www.seadatanet.org/Standards-Software/Software/Download-Manager
The installation package is available on the project Extranet along with the installation manual. In case of problems contact the CDI Support desk at cdi-support@maris.nl

3.2.5. Coupling table / coupling file

This is also a key component of SeaDataNet system, and it must be installed in each SeaDataNet node (NODC or data provider). It is used by the Download Manager to retrieve and give access for downloading of data sets following their LOCAL_CDI_Ids. The coupling table may be a table in a database or a flat ASCII file. It contains the mapping between the LOCAL_CDI_Ids and the corresponding data file (if the data are archived in a file system) or data request (if the data are archived in a database).

For each LOCAL_CDI_Ids, the coupling table gives either the address of the corresponding data file in the local file system or the SQL order able to extract the corresponding data from the database.

The coupling table is described in the same documentation as the download manager:
http://www.seadatanet.org/Standards-Software/Software/Download-Manager

The coupling table / coupling file contains the following information:

- LOCAL_CDI_ID: the CDI local identifier, as included in the central CDI directory at the SeaDataNet portal,
- Management modus:
    - Value equals to 1 or 3 if it concerns a pre-processed data file
        - 1 for mono-station files,
        - 3 for multi-station files,
    - Value equals to 2 if data have to be retrieved from a local database
    - Value equals to 4 means that the file is available for viewing through High Resolution Viewer Service (HRSVS)
    - Value equals to 5 means that the file is available for viewing
- Format (ODV or MEDATLAS or NetCDF),
- For modus 1 or 3:
    - File name.
    - URL to the webservice used to retrieve the data file
- For modus 2:
    - SQL query which allows to retrieve from local database all metadata and data necessary for creating the ODV data file,

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

14

- o Database connection parameters (protocol, ip address, port, name, login, password),
  - o Mapping configuration XML file for the conversion of the original datasets to the SeaDataNet ODV format.
- o For modus 4: only for data accessible to the High Resolution Viewer Service
  - ▪ the local CDI ID and the text 'HRSVS' are necessary for the system to work with the HRSVS software to display the seismic data
- o For modus 5 Format for non-accessible data (typically magnetic bands) by High Resolution Seismic Visualization Service
  - o For each of the files : SensorML, OEM, UKOAP190, SEGY an separate line is required
- o Note: in case you want to use Modus 4 and/or 5., please contact first the CDI support desk at cdi-support@maris.nl because it requires special attention.

3.2.6. Link between NEMO and the Coupling table

While reformatting data files to SeaDataNet format, NEMO is able to generate the records that have to be inserted in the coupling table or file. NEMO manages an internal database which can be exported to a coupling file.

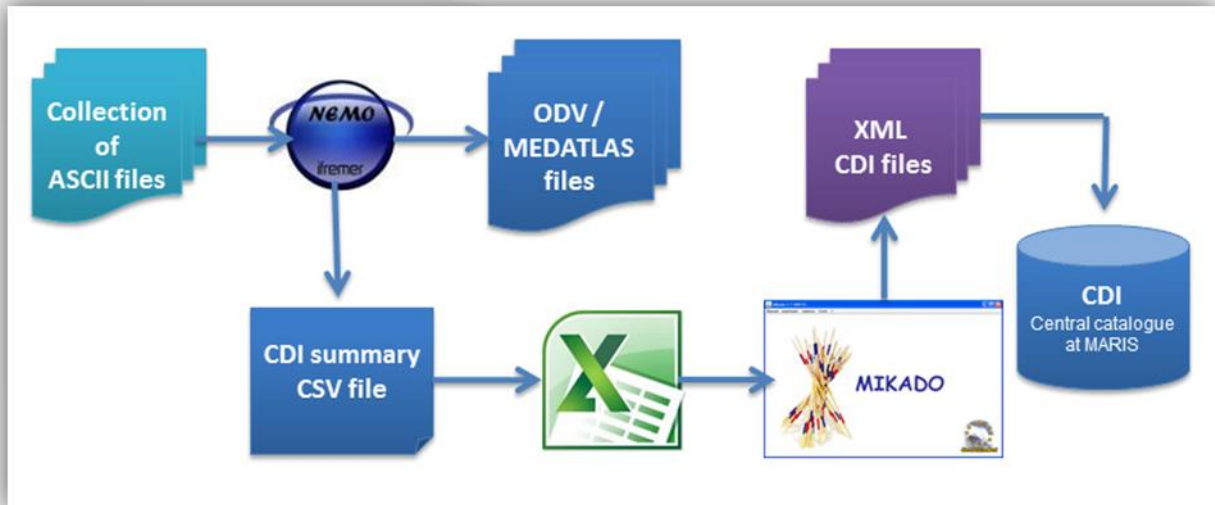### 3.2.7. Link between MIKADO and the coupling table

When MIKADO is used in an automatic way for the generation of XML files through a connection to a database or an Excel file, a user has to create SQL requests to retrieve the information in its database or its Excel file. In the same way, MIKADO is able to generate coupling files by using SQL requests.

3.2.8. Link between NEMO and MIKADO

The principle (See Figure below) is that while NEMO converts one file or a collection of files (like for example a collection of XBT files in a specific geographical area), it can also generate a " CDI summary CSV file" which contains all the minimum information necessary to create CDI records in the SeaDataNet catalogue. The content of the CSV summary file for CDI is described in the NEMO user manual.

This CSV file, once converted to EXCEL (or openOffice) **with a mandatory renaming of the Excel (OpenOffice) sheet into "STATION"**, can be read by MIKADO (automatic generation) using a JDBC driver for EXCEL (or OpenOffice) and then XML CDI files are generated (using predefined queries delivered in a XML configuration file with NEMO) and can be directly exported to the central SeaDataNet catalogue.

More information about this process can be found in the NEMO user manual downloadable at:
http://www.seadatanet.org/Standards-Software/Software/NEMO

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

15

From « raw » data files to SeaDataNet CDI catalogue using NEMO and MIKADO

3.3. Producing CDI entries in operational mode

Once the Download Manager is well installed and the data centres have prepared and submitted successfully some test CDI entries, associated data sets and the Coupling Table, and have tested the working of the shopping process, then it is time that data centres become operational and produce a full set of CDI metadata files, complete Coupling Table and local data sets with regular updating.

Once ready, the CDI metadata should be forwarded to MARIS for inclusion in the import version of the CDI service. This can be found at:

http://seadatanet.maris2.nl/v_cdi_v3_import/search.asp
http://seadatanet.maris2.nl/v_cdi_v3_import/browse_step.asp

During import automatic checks will be applied by MARIS on the syntax (= technical format) and semantics (=use of codes, conformity to CDI libraries, conformity to field definitions). Validated XML files will be incorporated in the import version. Partners will receive a report on the import, including possible rejected records and their errors. Moreover checks will take place for potential duplicates, within the new submission, and against the existing production database.These will also be included in the report and partners will be asked to check and explain.

The import version of the CDI service provides a means to do a last check of the metadata. Data centres can check the geographical distribution via the mapping functionality and are advised to use the Summary functionality on their full results, because that reveals very well the overall scope and consistency of the prepared metadata.

Note: it is not possible to order data in the import version! That is only possible in the test version and of course the production version.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

16

Once the import is checked and OK and once the data centre indicates that it is ready with its Coupling Table and files, then the data centre must request MARIS to move the CDI records from the import to the production version of the CDI service.

Thereafter MARIS will set the switch at the RSM for the specific data centre to production. From that point on the data centre will be operational and possibly receive data requests from registered users. This will be administered in the Request Status Manager (RSM) at the portal.

The production versions of the SeaDataNet CDI V3 user interfaces can be found at:

http://seadatanet.maris2.nl/v_cdi_v3/search.asp
http://seadatanet.maris2.nl/v_cdi_v3/browse_step.asp

Dedicated versions of the CDI user interface for related projects (Geo-Seas, Upgrade Black Sea SCENE, CaspInfo and EMODNET pilots) are operational from their respective websites.

Connected and operational data centres are requested to keep their CDI entries up-to-date by preparing and contributing regular updates to MARIS for inclusion.

For updating the principle is that only updated XML CDI files are transferred. These can contain new CDI entries in case of new data sets, but also modifications on existing CDI entries or even requested deletions of existing CDI entries.

At present it is advised to data centres to undertake a regular updating, e.g. at a fixed interval of 3 months, to contribute to an up-to-date SeaDataNet portal. It is still an indirect process, whereby data centres take the initiative and contact MARIS for the processing.

As part of SeaDataNet II activities are undertaking for implementing the OGC harvesting mechanism for metadata exchange between data centres and the portal.

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

17

4. How to complete a CDI entry content wise

4.1. Granularity and use cases

The core Common Data Index (CDI) is designed as a light metadata format, that can be generated automatically from data centres databases, after an initial mapping exercise to set translation properties. Each CDI should point to a data set, that can be accessed at or delivered by the specific data centre. The CDI is instrumental for data discovery purposes.

The definition of a dataset is arbitrary for each partner, because of the differences in granularity that partners are applying in their archive systems for storing and accessing datasets. A number of examples are given to illustrate the objectives and focus of the CDI.

Case: CTD measurements
CTD casts are collected at different geographical locations, e.g. during a scientific cruise. Each cast is represented by a data file / data set, which is stored and can be reproduced by a data centre. The CDI record reflects the metadata of a single CTD cast, covering multiple parameters, and including the information for accessing this specific CTD dataset or getting a copy of this specific CTD data set. The same approach can be applied for other in-situ and discrete measurements, such as sediment grabs, geological cores, water bottles, etc. In each case there is an individual measurement with one instrument at a geographical location which might have generated multiple parameters.

Case: Sea level / wave / current observations
Hydrodynamical observations are collected at different geographical locations, and might be part of a station, equipped with a number of instruments. Each instrument produces a timeseries of observation data, which is stored and can be reproduced by a data centre. The CDI record reflects the metadata of the resulting dataset of a single instrument at a single station, covering multiple parameters, and including the information for accessing this specific dataset or getting a copy of this specific dataset. This can cover a long timeseries. However in case there are large gaps in time coverage, the data centre might have decided to split the data set into a sub series of data sets, each covering a consistent observation period. In that case each sub serie is represented by a separate CDI record.

Case: Buoy with multiple sensors
In general practice a CDI will describe an observation data file from a single instrument. However nowadays there are various instruments with multiple sensors in use, providing their multiple parameter observations in single data files. A good example is a Poseidon buoy or a Ferrybox, that can be equipped with a combination of sensors, while the data are registered in a single data file. Such an observation should be described with a single CDI record, that gives the multiple parameters, but also the multiple sensors / instruments.
Note: in case of an offshore platform, equipped with separate instruments for e.g. waves and winds, that are recorded in separate data files, it is required to prepare separate CDI's for each measurement!

Case: Hydrographic measurements
The bathymetry of the seabed is measured by hydrographic surveys, which cover an specific area. Each survey can comprise a consistent timeperiod in which the area is sampled by sailing a number of tracks, during which the seabed bathymetry is recorded in singular tracks or zones by specific instruments. Each instrument during a specific survey produces a hydrographic survey dataset, which is stored and can be reproduced by a data centre. The CDI record reflects the metadata of the resulting data set of a single instrument during a single area survey, including the information for accessing this specific data set or getting a copy of this specific data set. The method of positioning during the survey gives important data quality information. Therefore position systems can be included in the CDI metadata by considering these as extra

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

18

entries of the measurement devices (= instruments). The same approach can be applied for other area measurements, such as seismic surveys, satellite images, etc. For these types of measurements it is also relevant to include metadata on the measurement resolution, such as grid size or frequency.

4.2. The ISO 19115 - 19139 standards

For purposes of standardisation, international exchange and interoperability with other systems and networks it was decided to adopt the International Metadata Standard for Geographic Information ISO 19115 and ISO 19139 as basis for the CDI format. This also fits in with the EU INSPIRE Directive..

The XML ISO 19115 Schema is defined and managed by the Technical committee TC211 of the International Organization for Standardization (ISO), who is responsible for making international standards on geographic information (www.isotc211.org). This standard defines more than 300 metadata elements, most of which can be applied optionally. It contains around ten elements, which are mandatory 'core' metadata. Moreover one can create profiles and add new elements.

ISO 19115 Profiles consist of :
•Core +
•some/all optional +
•zero-or-more extended elements
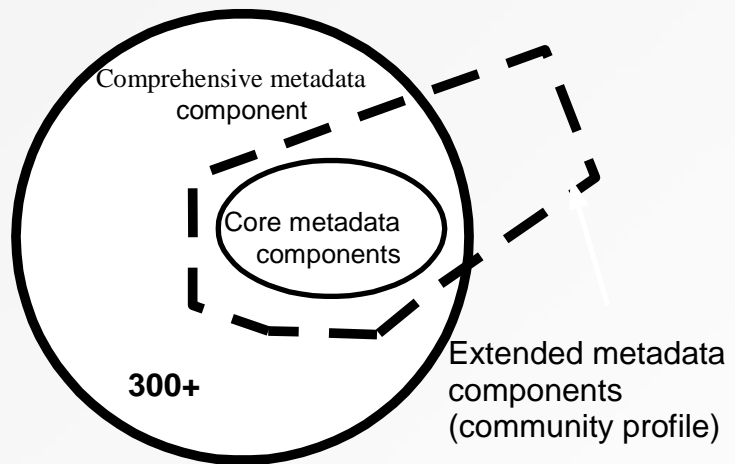
The CDI metadata are constructed as a dedicated subset of this standard. It also includes GML objects to alow for describing the coordinates of points, tracks and areas. More detail can be found in the CDI format documentation [4] which can be found at:
http://www.seadatanet.org/Standards-Software/Metadata-formats/CDI

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

19

5. References

5.1. Useful links

[1] SeaDataNet website: http://www.seadatanet.org
[2] SeaDataNet-I website: http://previous.seadatanet.org
[3] SeaDataNet-I training material: http://previous.seadatanet.org/Meetings#training and http://www.seadatanet.org/Events/Training-courses
[4] SeaDataNet CDI metadata format documentation: http://www.seadatanet.org/Standards-Software/Metadata-formats/CDI

5.2. On-line CDI User Interface Services

[5] Operational CDI User Interface Quick Search: http://seadatanet.maris2.nl/v_cdi_v3/browse_step.asp
[6] Operational CDI User Interface Extended Search: http://seadatanet.maris2.nl/v_cdi_v3/search.asp
[7] Import CDI User Interface Quick Search: http://seadatanet.maris2.nl/v_cdi_v3_import/browse_step.asp
[8] Import CDI User Interface Extended Search: http://seadatanet.maris2.nl/v_cdi_v3_import/search.asp
[9] Test CDI User Interface Quick Search: http://seadatanet.maris2.nl/v_cdi_v3_test/browse_step.asp
[10] Test CDI User Interface Extended Search: http://seadatanet.maris2.nl/v_cdi_v3_test/search.asp

5.3. On-line CDI Request Status Manager Service:

[11] Operational CDI RSM Service: http://www.seadatanet.org/Data-Access/Request-Status-Manager-RSM
[12] Test CDI RSM Service: http://seadatanet.maris2.nl/request/welcome_test.html

5.4. On-line Administration, Authentication and Authorisation, Service:

[13] Operational AAA Service: http://www.seadatanet.org/Standards-Software/AAA-services

5.5. Software and User Manuals:

See: http://www.seadatanet.org/Standards-Software/Software

[14] MIKADO User Manual: http://www.seadatanet.org/Standards-Software/Software/MIKADO
[15] NEMO User Manual: http://www.seadatanet.org/Standards-Software/Software/NEMO
[16] Med2MedSDN User Manual: http://www.seadatanet.org/Standards-Software/Software/Med2MedSDN
[17] EndsAndBends User Manual: http://www.seadatanet.org/Standards-Software/Software/EndsAndBends
[18] Download Manager Manuals: http://www.seadatanet.org/Standards-Software/Software/Download-Manager (note: DM Software should be downloaded from the Extranet).

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

20

5.6. Contacts for CDI Metadata Entries

[19] MARIS: Support (cdi-support@maris.nl)

5.7. Contacts for Common Vocabularies as used in CDI references

[20] BODC: Adam Leadbetter (alead@bodc.ac.uk)

5.8. Contacts for MEDATLAS Parameter Usage Vocabulary

[21] IFREMER: Michele Fichaut (Michele.Fichaut@ifremer.fr)

5.9. Contacts for EDMERP, EDMO and Bibliographic Metadata Entries as used in CDI references

[22] MARIS: Peter Thijsse (peter@maris.nl)

5.10. Contacts for EDMED Metadata Entries as used in CDI references

[23] BODC: Liz Bradshaw (elizb@bodc.ac.uk)

5.11. Contacts for CSR Metadata Entries as used in CDI references

[24] BSH: Anne Che-Bohnenstengel (Anne.Che-Bohnenstengel@bsh.de)

User manual and instructions for compiling CDI metadata, coupling table and associated data – 21/11/2013
sdn-userdesk@seadatanet.org – www.seadatanet.org

21