



International Coastal Atlas Network Cookbook:
**Establishing a CSW metadata catalogue
with GeoNetwork opensource**



Table of Contents

Introduction.....	3
What is a metadata catalogue?	3
What is CSW?.....	3
CSW Application Profiles	3
CSW Servers.....	4
Installing and configuring GeoNetwork 2.6.4.....	5
Installing GeoNetwork 2.6.4	5
Configuring the database for GeoNetwork 2.6.4.....	6
Configure GeoNetwork for Tomcat (optional).....	8
Starting GeoNetwork 2.6.4 (required for Jetty).....	9
Stopping GeoNetwork 2.6.4 (required for Jetty).....	9
GeoNetwork 2.6.4 Administration.....	10
Adding metadata records in GeoNetwork 2.6.4	10
Examples of CSW operations.....	12
CSW GetRecordById operation request.....	12
CSW GetRecords operation request.....	13
Acknowledgements.....	14
Document Information	14

Introduction

This document provides a tutorial for those who wish to understand CSW (Catalog Services for the Web) metadata catalogues, with a focus on the ISO Metadata Application Profile of CSW. It is aimed specifically at members of the International Coastal Atlas Network community and more generally at scientists, data managers, and system developers. Included in this document is a description of a metadata catalogue, the CSW standard, and a list of selected CSW servers. The document also contains initial pointers to establishing a CSW server using GeoNetwork opensource, and examples of selected CSW query operations aimed at system developers. GeoNetwork is recommended and used by the NETMAR project.

What is a metadata catalogue?

A metadata catalogue stores and publishes collections or sets of metadata records describing data, services, and related information resources. A user can search for and discover these resources within the metadata catalogue. Distributed metadata catalogues enable searching of metadata catalogues across the Internet between organisations.

What is CSW?

Standards are required to enable interoperable searching of distributed metadata catalogues between organisations. This is achieved using the CSW (Catalog Services for the Web) standard¹. CSW is an OGC (Open Geospatial Consortium) specification that defines common interfaces and operations to query and retrieve metadata contained in metadata catalogues. It enables a client application to search or query metadata across organisational boundaries. Formal query languages supported by CSW are Filter encoding and CQL (Common Query Language). Filter Encoding is an XML encoded query language, while CQL is a text encoded query language which attempts to be more human readable. Filter encoding is a joint OGC² and ISO³ standard. A simple Filter encoding example is outlined in a later section of this document.

CSW Application Profiles

While the common interfaces and operations of OGC catalogue services are well defined, it is left up to the system developer to define a specific information model for the catalogue service implementation. This includes mandatory and optional metadata elements to be incorporated in the catalogue, supported query languages, available search terms, results, etc. Experience has shown there is no single solution for catalogue services that fits every user's needs. Therefore,

¹ OpenGIS Catalogue Service: <http://www.opengeospatial.org/standards/specifications/catalog>

² OpenGIS Filter Encoding: <http://www.opengeospatial.org/standards/filter>

³ ISO 19143:2010, Geographic information -- Filter encoding:
http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=42137

there is a need for application profiles⁴. OGC has developed these profiles against CSW version 2.0.2:

1. ISO Metadata Application Profile (version 1.0.0)
2. ebRIM Profile (version 1.0.1)
3. OWL Application Profile (version 0.3)

Metadata catalogues that use ISO 19115 and ISO 19119 as its underlying information model should use the ISO Metadata Application Profile of CSW. The Filter encoding query language is mandatory for this profile, with CQL optional. Upon CSW client query requests, such CSW catalogues return ISO 19115 or ISO 19119 metadata records implemented using an ISO 19139 XML schema encoding. Also upon CSW client query requests, the return of Dublin Core encoded metadata is supported by the ISO Metadata Application Profile of CSW. The CSW specification defines the mappings between ISO 19115/19119 and Dublin Core.

CSW Servers

A CSW server is software that implements a metadata catalogue with a CSW interface. A CSW server may also implement ancillary metadata management functionality such as metadata editing tools, harvesting and synchronisation of metadata between distributed catalogues, group and user management, etc.

Several CSW servers have been developed. Commercial CSW servers include:

1. ESRI ArcGIS Server (<http://www.esri.com/>)
2. Intergraph GeoMedia (<http://www.intergraph.com/>)
3. MapInfo Manager (<http://www.pbinsight.com/>)

Open-source/freeware CSW servers include:

1. Constellation (<http://constellation.codehaus.org/>)
2. Degree (<http://www.degree.org/>)
3. GeoNetwork opensource (<http://geonetwork-opensource.org/>)
4. GI-cat (<http://essi-lab.eu/cgi-bin/wiki/view/GIcat/>)

GeoNetwork opensource is recommended and used by the NETMAR project as a CSW server. It is open source, supports ISO 19115/19119 and ISO 19139, supports the OGC CSW ISO Metadata Application Profile, has a strong user community, and is used by many government agencies across Europe to support the implementation of the INSPIRE Directive. For static metadata it is recommended to use GeoNetwork as an editing tool, unless specific organisations have special requirements which require the use of an alternative tool. For more

⁴ OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile, Version 1.0.0, OGC 07-045

dynamic metadata which needs to be automated in a processing chain, it is recommended that a data provider use alternative tools or scripting/programming languages to automate metadata production. However, for such dynamic metadata production environments, GeoNetwork can still be utilised as a CSW server. Guidance notes regarding the installation and configuring of GeoNetwork are outlined in the next section. Methods for metadata ingestion into GeoNetwork are also included.

Installing and configuring GeoNetwork 2.6.4

This section contains some guidance notes regarding the installation and configuring of GeoNetwork opensource version 2.6.4 aimed at system developers. GeoNetwork is an open source implementation of the CSW 2.0.2 ISO Metadata Application Profile 1.0.0 standard. It can run on Microsoft Windows, Linux and Mac OS X. The latest version can be downloaded from:

- <http://sourceforge.net/projects/geonetwork/>

A detailed user manual can be found on the GeoNetwork opensource website:

- <http://www.geonetwork-opensource.org/manuals/2.6.4/eng/users/>

For software developers, a detailed developer manual can also be found on the GeoNetwork opensource website:

- <http://www.geonetwork-opensource.org/manuals/2.6.4/eng/developer/>

By default GeoNetwork comes embedded with the Jetty Servlet container. However, Apache Tomcat (version 5.5+) is also supported via a custom install. GeoNetwork requires an RDBMS database in order to store metadata and ancillary information. The default database in GeoNetwork 2.6.4 is an embedded McKoiDB RDBMS database. However, other RDBMS databases including MySQL, PostGreSQL, Oracle and generic JDBC (Java Database Connectivity) connections are supported. It is recommended to use a standalone database instead of McKoiDB in a production environment. These installation instructions deal with configuring the MySQL database.

Installing GeoNetwork 2.6.4

Before installing GeoNetwork 2.6.4, please make sure you have a Java Runtime Environment (JRE 1.5.0+) installed. You can use the GeoNetwork Windows installer (.exe file) for the Windows platforms, or the platform independent installer (.jar file) for any platform (also works on Windows). For example, to start the graphical installation wizard in Windows double click on:

```
geonetwork-install-2.6.4-0.exe
```

To run the platform independent graphical installation wizard double click on:

```
geonetwork-install-2.6.4-0.jar
```

To run the platform independent graphical installation wizard from the command line type:

```
java -jar geonetwork-install-2.6.4-0.jar
```

Once the wizard is running, simply follow the onscreen instructions.

Configuring the database for GeoNetwork 2.6.4

After GeoNetwork 2.6.4 is installed we next need to configure the GeoNetwork database. For these installation instructions we are using MySQL. You can download the open source MySQL Community Server from the MySQL website: <http://dev.mysql.com/downloads/>. Alternatively on UNIX systems, you may be able to download and install MySQL Community Server via a UNIX distribution stream. Once you have successfully installed the MySQL server, next you can create a dedicated database for GeoNetwork 2.6.4, e.g. by using the following commands:

```
mysql -u root -p
create database myDatabase;
CREATE USER 'myUsername'@localhost IDENTIFIED BY 'myPassword';
GRANT SELECT, INSERT, UPDATE, DELETE, CREATE, INDEX, DROP ON *.*
TO 'myUsername'@'localhost';
```

To start configuring the GeoNetwork database we need to run GeoNetwork's GAST software - GeoNetwork's Administrator Survival Tool. GAST is a standalone Java application. It performs administration tasks such as the configuration of the JDBC connection to the RDBMS database, database initialisation, etc. More detailed information on GAST can be found in the GeoNetwork user manual.

To start GAST on Windows, simply click:

```
Start GAST (under Start > Programs > GeoNetwork opensource)
```

GAST can also be started by double clicking on:

```
gast.jar (in the ../geonetwork/gast directory)
```

GAST can also be started from the command line by typing:

```
java -jar gast.jar (in ../geonetwork/gast directory)
```

Next, we need to configure and connect GeoNetwork to this new database. Once GAST is running, click the **DBMS** link (Figure 1). In this example, we configure GeoNetwork to connect to the MySQL database that we created previously. Then click the **Save** when finished.

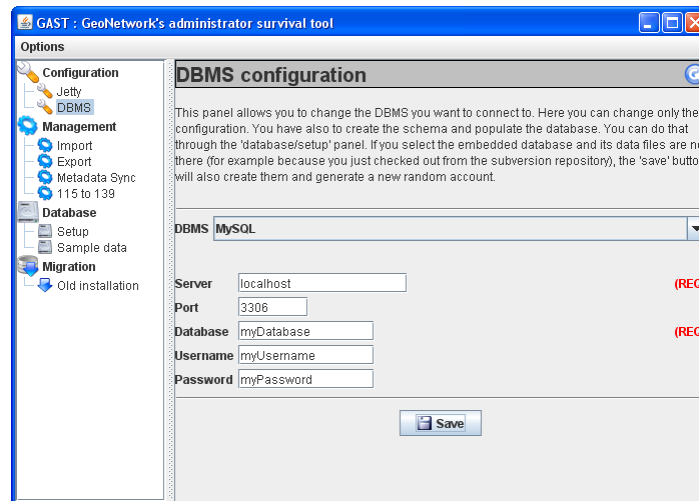


Figure 1 GeoNetwork GAST – DBMS configuration

Next, we need to initialise this new database for GeoNetwork 2.6.4 by creating tables that are required by GeoNetwork. Within the GAST, click the **Setup** link in the left menu, and then the **Setup** button within the main menu (Figure 2). This will recreate and reinitialise GeoNetwork's internal tables (note: this will overwrite and reinitialise existing tables if the database has previously been setup with GAST).

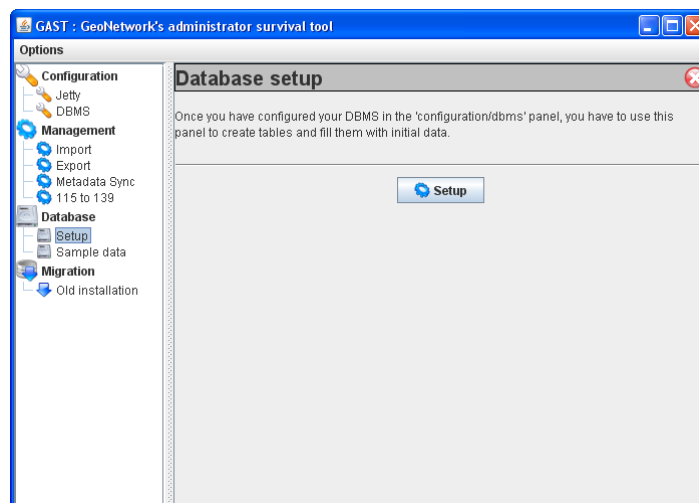


Figure 2 GeoNetwork GAST – Database setup

Note: in GeoNetwork 2.6.4 the following error message may appear in a dialog box when setting up the MySQL database for the first time via the GAST tool:

```
Error: Cyclic reference found:
[relations, categories, settings, languages, sources, isolanguages,
isolanguagesDec, Regions, RegionsDes, Users, Operations,
OperationsDes, Groups, GroupsDes, UsersGroups, CategorisesDes,
Metadata, MetadataCateg, OperationAllowed]
```

This error message can be ignored. Please click “OK” in this dialog box, and the GeoNetwork database setup will continue.

Configure GeoNetwork for Tomcat (optional)

If you wish to deploy GeoNetwork on Apache Tomcat as opposed to the embedded Jetty, then you have a number of options. If you installed GeoNetwork using the steps outlined in the previous sections, then you have two main options to configure for Tomcat:

1. Leave the existing GeoNetwork installation in the “../geonetwork/web/geonetwork/” directory and point a Tomcat context file (named **geonetwork.xml**) to this directory. An example geonetwork.xml file is:

```
<Context docBase="../../geonetwork/web/geonetwork/"
        privileged="true"
        antiResourceLocking="false"
        antiJARLocking="false">
</Context>
```

The context file is usually placed in the \$CATALINA_HOME/conf/Catalina/localhost directory. Make sure that the Tomcat user has the appropriate permissions to access the GeoNetwork installation directory.

2. An alternative option is to copy the installation from “../geonetwork/web/geonetwork/” to Tomcat’s webapps directory. Tomcat will then automatically pick up the GeoNetwork installation. Make sure that the Tomcat user has the appropriate permissions to access this copied installation directory.

If you did not install GeoNetwork using the steps outlined in the previous sections, then you have a third option to configure for Tomcat:

3. Download the **geonetwork.war** file from the GeoNetwork download website. This file can be placed in the Tomcat’s webapps directory where it will be automatically deployed by GeoNetwork. However, if you are configuring this deployment with a MySQL database then you still need to create a MySQL database. You also need to install the GAST tool independently to initialise this MySQL database. Once the database is setup you then need to modify the **config.xml** file located in the “../webapps/geonetwork/WEB-INF” directory.

The default McKoiDB RDBMS database must be removed from **config.xml**:

```
<resource enabled="true">
  <name>main-db</name>
  <provider>jeeves.resources.dbms.DbmsPool</provider>
  <config>
    <user>BayACrsQ</user>
    <password>Qf7Po9T0</password>
```



```

<driver>com.mckoi.JDBCdriver</driver>
<url>jdbc:mckoi://localhost:9157/</url>
<poolSize>10</poolSize>
</config>
<activator class="org.fao.geonet.activators.McKoiActivator">
  <configFile>WEB-INF/db/db.conf</configFile>
</activator>
</resource>

```

The following should be inserted in its place:

```

<resource enabled="true">
  <name>main-db</name>
  <provider>jeeves.resources.dbms.DbmsPool</provider>
  <config>
    <user>myUsername</user>
    <password>myPassword</password>
    <driver>com.mysql.jdbc.Driver</driver>
    <url>jdbc:mysql://localhost/myDatabase</url>
    <poolSize>10</poolSize>
  </config>
</resource>

```

This **config.xml** modification was tested in GeoNetwork 2.6.4. Please ensure Tomcat is shut down when making these changes.

Starting GeoNetwork 2.6.4 (required for Jetty)

If you installed GeoNetwork using the embedded Jetty and successfully configured the database, then you can start GeoNetwork. On Windows click:

```
Start Server (under Start > Programs > GeoNetwork opensource)
```

Or from the Windows command line by typing:

```
sh start-geonetwork.bat (in ../geonetwork/bin directory)
```

Under UNIX, you can start GeoNetwork from the command line by typing:

```
sh start-geonetwork.sh (in ../geonetwork/bin directory)
```

Once started (can take up to 1 minute), you can access GeoNetwork in your web browser at:

- <http://localhost:8080/geonetwork/>

Stopping GeoNetwork 2.6.4 (required for Jetty)

If you installed GeoNetwork using the embedded Jetty, you can stop GeoNetwork on Windows by clicking:

```
Start Server (under Start > Programs > GeoNetwork opensource)
```

Or from the Windows command line by typing:

```
sh stop-geonetwork.bat (in ../geonetwork/bin directory)
```

Under UNIX, you can stop GeoNetwork from the command line by typing:

```
sh stop-geonetwork.sh (in ../geonetwork/bin directory)
```

GeoNetwork 2.6.4 Administration

Once GeoNetwork is up and running, an important first time administration step is to change the default GeoNetwork administrator username/password which is admin/admin. The first time you start GeoNetwork, login as the administrator and select the “Administration” link in the tab bar. Then reset the administrator password via the “User management” link. For more information on all GeoNetwork 2.6.4 administration features, please consult the GeoNetwork manual.

Adding metadata records in GeoNetwork 2.6.4

There are a number of ways to add metadata records to the GeoNetwork 2.6.4 database. You must login with appropriate privileges to add metadata. These methods are documented in the user manual. A summary of these methods are outlined here:

- 1. Creating a new metadata record with the GeoNetwork metadata editor:**

You can choose this method by first selecting the “Administration” link in the tab bar, and then clicking “New metadata”. Next, you need to select an appropriate metadata template. The ISO 19139 templates for vector and raster data are preferred. Once you select the template you can start editing metadata (e.g. Figure 1). Note, when setting up GeoNetwork, you may wish to add the default metadata templates as they are not activated by default. This can be done via “Add templates” under the “Administration” tab. You can also create your own template via the metadata editor. In this case, when saving the new metadata template in the editor you choose “Template” from the “Type” dropdown list.

- 2. Uploading pre-existing metadata with the GeoNetwork metadata insert tool:**

You can choose this method by first selecting the “Administration” link in the tab bar, and then clicking “Metadata insert”. You can choose to upload the XML file directly or else to paste the XML contents into a textbox. This method is useful for uploading pre-existing metadata created by other tools or scripts. You can also choose to upload the XML as a metadata template by selecting “Template” from the “Type” dropdown list.

Firefox

GeoNetwork - The portal to spatial data ...

netmar.ucc.ie/geonetwork/srv/en/metadata.edit?id=95

Reset Save Save and close Check Other actions Cancel

Parent/child metadata:
 Add or update parent metadata section

Related service metadata:
 Link service metadata

Related feature catalogues:
 Link feature catalogue

Identification info

Title * General Bathymetric Chart of the Oceans GEBCO

Alternate title ⓘ British Oceanographic Data Centre record 1048GE

Date * 2011-10-04 Clear

Date type * Publication

Date * 2011-01-25 Clear

Date type * Creation

Date * 2011-01-25 Clear

Date type * Revision

Code EDMED4285

Codespace ⓘ http://www.bodc.ac.uk/ (Suggestions: ⓘ)

Abstract *
 The General Bathymetric Chart of the Oceans GEBCO_08 Grid is a global terrain model for oceans and land at 30 arc-second intervals. The bathymetric portion of the grid is largely based on a database of ship-track soundings with interpolation between soundings guided by satellite-derived gravity data. Existing gridded bathymetric data sets are included in some areas. The land portion of the grid is largely based on the US Geological Survey's SRMT30 data set, developed with data from the US

Point of contact ⓘ

Individual name ⓘ Unknown

Organisation name ⓘ British Oceanographic Data Centre

Position name ⓘ Unknown

Role * Originator

Delivery point ⓘ Joseph Proudman Building

Delivery point ⓘ 6 Brownlow Street

City ⓘ Liverpool

Administrative area ⓘ Merseyside

Postal code ⓘ L3 5DA

Country ⓘ United Kingdom United Kingdom

Electronic mail address ⓘ enquiries@bodc.ac.uk

Online resource

URL * http://www.bodc.ac.uk/

Point of contact ⓘ ⓘ

Organisation name ⓘ British Oceanographic Data Centre

Position name ⓘ Director

Role * Custodian

Delivery point ⓘ Joseph Proudman Building

Delivery point ⓘ 6 Brownlow Street

City ⓘ Liverpool

Figure 1: GeoNetwork opensource metadata editor

3. Batch uploading pre-existing metadata from a server-side directory:

You can choose this method by first selecting the “Administration” link in the tab bar, and then clicking “Batch Import”. Next, you enter the full directory path located on the server’s file system. Once activated, GeoNetwork will scan this directory and load all metadata records contained in the directory. This is a useful method where a user can batch upload numerous pre-existing metadata records in one go, where these metadata records have typically been created by another tool or script.

4. Harvesting pre-existing metadata:

You can choose this method by first selecting the “Administration” link in the tab bar, and then clicking “Harvesting management”. Click “Add” to configure a new metadata harvesting task. GeoNetwork is able to harvest from a number of remote sources

including: another GeoNetwork node, a CSW server, a WebDAV server, an OAI-PMH server, a local file system, etc. You can decide to schedule the harvesting task (e.g. every hour, every week, etc.) or just run a once-off harvest. Harvesting is a useful method for collecting remote metadata and storing them locally for faster access. It is also a useful method for the automatic pulling of local file system metadata into the catalogue, where this metadata is being dynamically produced by an independent processing chain.

5. XML services (advanced):

GeoNetwork provides access to several functionalities through the use of XML web services. These web services support the insertion of metadata records into the GeoNetwork database. There are two main methods: Metadata Service⁵ and CSW Service⁶. With the Metadata Service, metadata can be added using the “metadata.insert” operation. With the CSW Service, metadata can be added using the CSW “Transaction” operation. The CSW Service is more standards compliant as it uses the OGC CSW 2.0.2 specification. These two metadata insertion methods are useful for the automatic pushing of metadata into the catalogue by a processing chain, where this metadata is being dynamically produced by the same processing chain. These two methods require software developer skills to implement. Further details, with examples, are documented in the developer manual.

Examples of CSW operations

This section gives examples of the CSW GetRecordById and GetRecords operation requests aimed at system developers who wish to understand better, explore and test the engineering details of CSW in more detail using examples. Full details of these operations can be referenced in CSW 2.0.2 - ISO Metadata Application Profile v1.0.0.

CSW GetRecordById operation request

This operation enables a client CSW application to request one or more specific metadata records from the CSW server using their metadata identifiers (mapped to the fileidentifier of an ISO 19139 document). This example, tested against GeoNetwork 2.6.4, requests a full ISO 19139 metadata record with an identifier of “64c8493d6bd95d93b7e04fb868fd568e”:

```
http://netmar.ucc.ie/geonetwork/srv/en/csw?
SERVICE=CSW&
REQUEST=GetRecordById&
VERSION=2.0.2&
elementSetName=full&
outputSchema=csw:IsoRecord&
ID=64c8493d6bd95d93b7e04fb868fd568e
```

⁵ http://www.geonetwork-opensource.org/manuals/2.6.4/eng/developer/xml_services/metadata_xml_services.html (accessed 19 July 2012)

⁶ http://www.geonetwork-opensource.org/manuals/2.6.4/eng/developer/xml_services/csw_services.html (accessed 19 July 2012)

The example below contains the response with an extract of the requested ISO 19139 metadata record:

```
<csw:GetRecordByIdResponse>
  <gmd:MD_Metadata>
    <gmd:fileIdentifier>
      <gco:CharacterString>
        64c8493d6bd95d93b7e04fb868fd568e
      </gco:CharacterString>
    </gmd:fileIdentifier>
    .
    .
    .
  </gmd:MD_Metadata>
</csw:GetRecordByIdResponse>
```

CSW GetRecords operation request

This operation enables a client CSW application to search for metadata in the catalogue service. This example, tested against GeoNetwork 2.6.4, requests the CSW server to return full ISO 19139 metadata for any records which contain the word "GEBCO_08" (note: the XML Filter below needs to be URL encoded before pasting into a web browser):

```
http://netmar.ucc.ie/geonetwork/srv/en/csw?
SERVICE=CSW&
REQUEST=GetRecords&
VERSION=2.0.2&
resultType=results&
elementSetName=full&
outputSchema=http://www.isotc211.org/2005/gmd&
typeName=csw:Record&
constraintLanguage=FILTER&
constraint_language_version=1.1.0&
constraint=
  <Filter xmlns="http://www.opengis.net/ogc"
    xmlns:gml="http://www.opengis.net/gml">
    <PropertyIsLike>
      <PropertyName>any</PropertyName>
      <Literal>GEBCO_08</Literal>
    </PropertyIsLike>
  </Filter>
```

The example below contains the response with an extract of the one ISO 19139 metadata record that was discovered:

```
<csw:GetRecordsResponse>
  <csw:SearchStatus timestamp="2012-07-26T17:07:29"/>
  <csw:SearchResults numberOfRecordsMatched="1"
    numberOfRecordsReturned="1"
    elementSet="full"
    nextRecord="0">
```

```

<gmd:MD_Metadata>
  .
  .
  .
  <gmd:title>
    <gco:CharacterString>
      General Bathymetric Chart of the Oceans GEBCO_08 Grid
    </gco:CharacterString>
  </gmd:title>
  .
  .
  .
</gmd:MD_Metadata>
</csw:SearchResults>
</csw:GetRecordsResponse>

```

Acknowledgements

This cookbook was written for the International Coastal Atlas community under the auspices of the NETMAR (Open Service Network for Marine Environmental Data) project. NETMAR is partially funded by the European Commission under Theme ICT-2009.6.4 ICT for environmental services and climate change adaptation of the Information & Communication Technologies FP7 Programme.

Thanks to Roy Lowry of the British Oceanographic Data Centre for supplying the baseline GEBCO metadata example which has been tuned for the NETMAR project.

This document has been reviewed by, and incorporates comments from, Peter Walker of Plymouth Marine Laboratory; Torill Hamre of the Nansen Environmental and Remote Sensing Center; and John Helly of the San Diego Supercomputer Center. Thanks to the reviewers for their help in making the document clear and readable. Further feedback on this document is welcomed, and may be provided by contacting the author whose details are below.

Document Information

Author	Declan Dunne, Coastal and Marine Research Centre	
Contact	d.dunne@ucc.ie	
Version	2.1	
Date	2012 July 27	
Revisions	2.1	Responses to NETMAR internal review
	2.0	Comments from ICAN community
	1.0	2011 December 23