

**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	4
Installing GeoNetwork 2.6.4.....	5
Configuring the database for GeoNetwork 2.6.4	5
Starting GeoNetwork 2.6.4	7
Stopping GeoNetwork 2.6.4.....	7
GeoNetwork 2.6.4 Administration	7
Adding metadata records in GeoNetwork 2.6.4.....	8
Examples of CSW operations	9
CSW GetRecordById operation request.....	9
CSW GetRecordByMd operation request.....	10
Acknowledgements	11
Document Information	11

Introduction

This document provides a tutorial for those who wish to understand CSW (Catalog Services for the Web) metadata catalogues. 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 some CSW servers available. The document also contains initial pointers to establishing a CSW server using GeoNetwork opensource, and examples of some 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) standard. 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. Query languages supported by CSW are CQL (Common Query Language) and OGC FILTER.

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, there is a need for application profiles¹. 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)

¹ OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile, Version 1.0.0, OGC 07-045

Metadata catalogues that use the ISO 19115 and ISO 19119 as its underlying information model should use the ISO Metadata Application Profile of CSW. 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 also supported by the ISO Metadata Application Profile of CSW.

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/>)

Opensource/freeware CSW servers include:

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

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 standard. It can run on Microsoft Windows, Linux and Mac OS X. GeoNetwork is recommended and used by the NETMAR project. The latest version can be downloaded from: <http://sourceforge.net/projects/geonetwork/>. A more detailed user manual can be found on the GeoNetwork opensource website: <http://geonetwork-opensource.org/manuals/2.6.4/users/>.

By default GeoNetwork comes embedded with the Jetty Servlet container. However, Apache Tomcat (version 5.5+) is also supported via a custom install. However, these installation instructions deal with the default Jetty installation. 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

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 can 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 example 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 configuration tasks such as 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: <http://geonetwork-opensource.org/manuals/2.6.4/users/>

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.

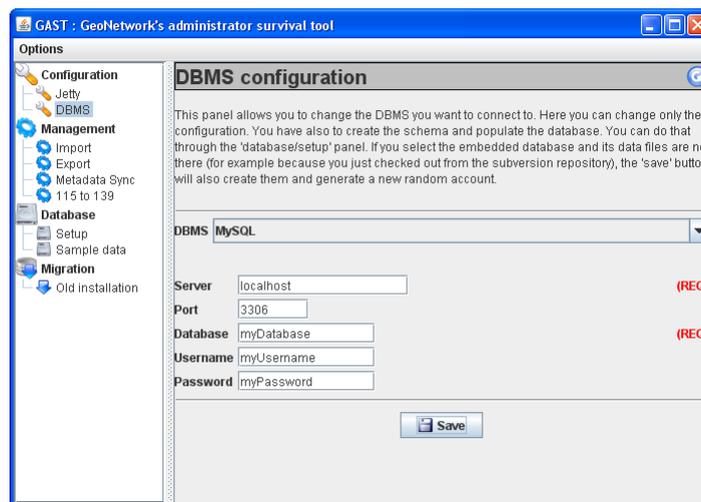


Figure 1 GeoNetwork GAST – DBMS configuration

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 internal tables if the database has previously been setup with GAST).

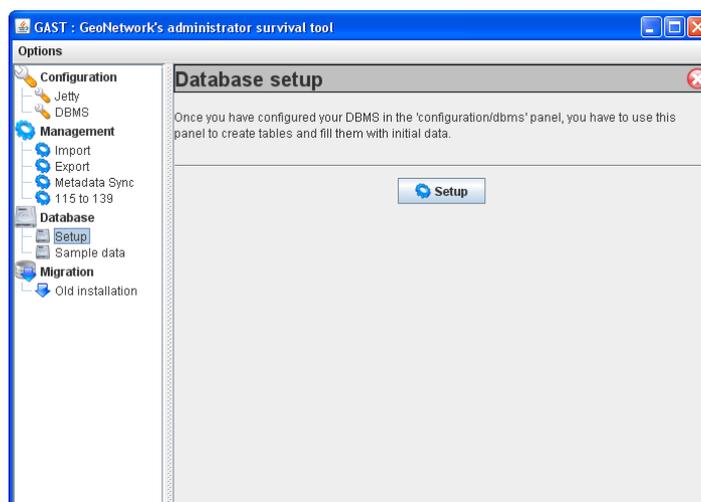


Figure 2 GeoNetwork GAST – Database setup

Note: in GeoNetwork 2.6.4 the following error message may appear in a dialog box when setting up the MySQL database 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.

Starting GeoNetwork 2.6.4

Once the database is initialised, 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

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

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

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. Most of these methods are documented in the user manual: <http://geonetwork-opensource.org/manuals/2.6.4/users/>. 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 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. The method is useful for uploading existing metadata created by other GIS applications. You can also choose to upload the XML as a metadata template by selecting “Template” from the “Type” dropdown list.

3. **Batch loading 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.

4. **Metadata harvesting:**

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 process for collecting remote metadata and storing them locally for faster access.

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 to the GeoNetwork database. There are two main methods: Metadata Service and CSW Service. 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. This XML services option for adding metadata requires software developer skills to implement. Further details are documented in the developer manual: <http://geonetwork-opensource.org/manuals/2.6.4/developer/>.

Figure 1: GeoNetwork opensource metadata editor

Examples of CSW operations

This section gives examples of the CSW GetRecordById and GetRecords operation requests aimed at system developers. Full details of these operations can be referenced in CSW 2.0.2 - ISO Metadata Application Profile v1.0.0².

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”:

² OpenGIS Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile, Version 1.0.0, OGC 07-045

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

CSW GetRecordById 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" (note: the XML Filter below needs to be URL encoded to work in 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</Literal>
    </PropertyIsLike>
  </Filter>
```

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.

This document has been reviewed by, and incorporates comments from, Pete Walker of Plymouth Marine Laboratory. 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	1.0
Date	2011 December 23
Revisions	