



The complete manual

By the developers

V 2.4

Copyright © 2007-2009 The Open Source Geospatial Foundation

Table of Contents

Preface	vi
I. User Guide	1
1. A Geographic Information Management System for all	2
1.1. Introduction	2
What is GeoNetwork opensource	2
Background and evolution	2
The use of International Standards	3
Harvesting geospatial data in a shared environment	3
1.2. GeoNetwork and the Open Source Community Development	3
2. Getting Started	5
2.1. Default Search	5
2.2. Searching by Categories	6
2.3. Advanced Search	7
2.4. Analysing Search Results	10
2.5. Privileges, roles and user groups	11
3. Viewing and Analysing the Data	13
3.1. Meta Data Description	13
Identification Section	13
Distribution Section	16
Reference System Section	16
Data Quality Section	17
Metadata Information Section	17
4. Adding new Data and Information	18
4.1. Creating a New Record using the Metadata Editor on line	19
The steps in more details	19
Switching Editing Views from Default to Advanced to XML View	20
Using basic commands of the editor	22
4.2. Entering Metadata for your Map	23
Entering Metadata For Your Map	23
Creating a Thumbnail	25
Linking data for download	26
Assigning Privileges for a Map	27
Assigning Categories for a Map	29
4.3. Uploading a New Record using the XML Metadata Insert Tool	29
5. Metadata in Spatial Data Management	32
5.1. What is Metadata?	32
5.2. What are Metadata Standards?	32
5.3. Why do we need Standardised Metadata?	32
5.4. Geographic Information Metadata Standard	32
5.5. Metadata profiles	33
5.6. Transition between metadata standards	33
6. Installing the software	34
6.1. New version - New functionalities	34
6.2. Where do I get the installer?	35
6.3. System requirements	35
Additional Software	35
Supported browsers	35

6.4. How do I install GeoNetwork opensource?	36
On Windows	36
Installation using the platform independent installer	37
Command line installation	37
II. Administrator Guide	39
7. Basic configuration	40
7.1. System configuration	40
Site General site parameters.	42
Catalogue services (CSW, Z39.50)	42
Proxy configuration	43
Email & notification	43
Removed metadata	44
Authentication	44
8. User and Group Administration	46
8.1. Creating new user Groups	46
8.2. Creating new Users	48
8.3. User Profiles	49
9. Import facilities	51
9.1. File import	51
XML file import	51
MEF file import	51
XML metadata insert	52
9.2. Batch import	52
Structured import	53
10. Harvesting	55
10.1. Introduction	55
10.2. Mechanism overview	55
10.3. Harvesting life cycle	55
10.4. Multiple harvesting and hierarchies	56
10.5. General notes and issues	56
General	56
GeoNetwork harvesting type	56
WebDAV harvesting type	57
CSW harvesting type	57
OAI-PMH harvesting type	57
OGC service harvesting type	57
10.6. The main page	57
Harvesting result tips	59
10.7. Adding new nodes	60
Adding a GeoNetwork node	61
Adding a Web DAV node	63
Adding a CSW node	65
Adding an OAI-PMH node	67
Adding an OGC Service (ie. WMS, WFS, WCS)	69
11. Metadata ownership	71
11.1. Introduction	71
11.2. Access policy	71
Visualisation	71
Editing	71
11.3. Privileges	71

11.4. Transfer Ownership	72
12. Thesaurus	74
12.1. Introduction	74
12.2. Thesaurus / SKOS format	74
12.3. Thesaurus administration	74
Creation of a new thesaurus	75
Import existing thesaurus	75
12.4. Editing/browsing thesaurus: add/remove/browse keywords	76
12.5. Metadata editing: adding keywords	76
12.6. Search criteria: keywords	77
13. GeoNetwork's Administrator Survival Tool - GAST	78
13.1. What is GAST?	78
13.2. Starting GAST	78
13.3. Operating modes	78
13.4. Tools subdivision	79
13.5. Server and Account configuration dialogue	79
14. Localisation	81
14.1. Localisation of dynamic user interface elements	81
15. Import / export tools	83
15.1. Introduction	83
15.2. Import	83
15.3. Export	84
III. Server Reference	85
16. Software development	86
16.1. System Requirements	86
16.2. Running the software with a servlet engine	86
16.3. Development	86
Compiling GeoNetwork	86
Source code documentation	87
Creating the installer	87
17. Harvesting	88
17.1. Structure	88
JavaScript code	88
Java code	88
XSL stylesheets	88
17.2. Data storage	89
17.3. Guidelines	89
18. Metadata Exchange Format v1.1	91
18.1. Introduction	91
18.2. File format	91
18.3. The info.xml file	92
Date format	93
19. XML Services	95
19.1. Calling specifications	95
Calling XML services	95
Exception handling	96
19.2. General services	98
xml.info	98
xml.forward	103
19.3. Harvesting services	104

Introduction	104
xml.harvesting.get	104
xml.harvesting.add	107
xml.harvesting.update	112
xml.harvesting.remove/start/stop/run	113
19.4. System configuration	114
Introduction	114
xml.config.get	114
xml.config.update	117
19.5. MEF services	120
Introduction	120
mef.export	120
mef.import	120
Metadata ownership	121
19.6. Relations	121
Introduction	121
xml.relation.get	121
19.7. Schema information	123
Introduction	123
xml.schema.info	123
20. Settings hierarchy	126
20.1. Introduction	126
20.2. The system hierarchy	126
20.3. Harvesting nodes	127
Nodes of type GeoNetwork	128
Nodes of type geonetwork20	129
Nodes of type WebDAV	130
Nodes of type CSW	130
A. Frequently Asked Questions	131
B. Glossary of Metadata Fields Description	134
C. ISO Topic Categories	138
D. Free and Open Source Software for Geospatial Information Systems	141
D.1. Web Map Server software	141
D.2. GIS Desktop software	141
D.3. Web Map Viewer and Map Server Management	141
Glossary	142
Index	146

Preface

About this Project. This document provides guidelines to install, configure, use and customise the GeoNetwork opensource software. The GeoNetwork project started out as a Spatial Data Catalogue System for the Food and Agriculture organisation of the United Nations (FAO)¹, the United Nations World Food Programme (WFP)² and the United Nations Environmental Programme (UNEP)³. At present the project is widely used as the basis of Spatial Data Infrastructures all around the world. The project is part of the Open Source Geospatial Foundation (OSGeo) and can be found at <http://geonetwork-opensource.org>.

License Information. Copyright (c) 2007-2009 Open Source Geospatial Foundation (OSGeo⁴).

You are free to Share — to copy, distribute and transmit the work. Under the following conditions:

- Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- No Derivative Works. You may not alter, transform, or build upon this work.
- For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- Nothing in this license impairs or restricts the author's moral rights.

You may obtain a copy of the License at Creative Commons⁵

The document is written in DocBook format for consistency and portability.

Author Information. This framework and documentation was written by the GeoNetwork opensource Developers. If you have questions, found a bug or have enhancements, please contact us through the GeoNetwork opensource Development Mailing list at [<geonetwork-devel@lists.sourceforge.net>](mailto:geonetwork-devel@lists.sourceforge.net)



¹ <http://www.fao.org>

² <http://vam.wfp.org>

³ <http://www.unep.org>

⁴ <http://www.osgeo.org>

⁵ <http://creativecommons.org/licenses/by-nd/3.0/>

Part I. User Guide

This part of the document focusses on the general use of GeoNetwork for visitors and for Spatial Data Managers that want to publish data and metadata.

1. A Geographic Information Management System for all

1.1 Introduction

What is GeoNetwork opensource

GeoNetwork opensource is a standard based and decentralised spatial information management system, designed to enable access to geo-referenced databases and cartographic products from a variety of data providers through descriptive metadata, enhancing the spatial information exchange and sharing between organisations and their audience, using the capacities and the power of the Internet. The system provides a broad community of users with easy and timely access to available spatial data and thematic maps from multidisciplinary sources, that may in the end support informed decision making. The main goal of the software is to increase collaboration within and between organisations for reducing duplication and enhancing information consistency and quality and to improve the accessibility of a wide variety of geographic information along with the associated information, organised and documented in a standard and consistent way.

Main Features

- Instant search on local and distributed geospatial catalogues
- Uploading and downloading of data, documents, PDF's and any other content
- An interactive Web map viewer that combines Web Map Services from distributed servers around the world
- Online map layout generation and export in PDF format
- Online editing of metadata with a powerful template system
- Scheduled harvesting and synchronisation of metadata between distributed catalogues
- Groups and users management
- Fine grained access control

Background and evolution

The prototype of the GeoNetwork catalogue was developed by the Food and Agriculture organisation of the United Nations (FAO) in 2001 to systematically archive and publish the geographic datasets produced within the organisation. The prototype was built on experiences within and outside the organisation. It used metadata content available from legacy systems that was transformed into what was then only a draft metadata standard, the ISO 19115. Later on, another UN agency, the World Food Programme (WFP) joined the project and with its contribution the first version of the software was released in 2003 and operational catalogues were established in FAO and WFP. The system was based on the ISO19115:DIS metadata standard and embedded the Web Map Client InterMap that supported Open Geospatial Consortium (OGC) compliant Web Map Services. Distributed searches were possible using the standard Z39.50 catalogue protocol. At that moment it was decided to develop the program as a Free and Open Source Software to allow the whole geospatial users community to benefit from the development results and to contribute to the further advancement of the software.

Jointly with the UN Environmental Programme (UNEP), FAO developed a second version in 2004. The new release allowed users to work with multiple metadata standards (ISO 19115, FGDC and Dublin Core) in a transparent manner. It also allowed metadata to be shared between catalogues through a caching mechanism, improving reliability when searching in multiple catalogues.

In 2006, the GeoNetwork team dedicated efforts to develop a DVD containing the GeoNetwork version 2.0.3 and the best free and open source software in the field of Geoinformatics. The DVD was produced and distributed in hard copy to over three thousand people and is now also available for download from the GeoNetwork Community website¹.

GeoNetwork opensource is the result of the collaborative development of many contributors. These include among others the Food and Agriculture organisation (FAO), the UN Office for the Coordination of Humanitarian Affairs (UNOCHA), the Consultative Group on International Agricultural Research (CSI-CGIAR), The UN Environmental Programme (UNEP), The European Space Agency (ESA) and many others. Support for the metadata standard ISO19115:2003 has been added by using the ISO19139:2007 implementation specification schema published in May 2007. The release also serves as the open source reference implementation of the OGC Catalogue Service for the Web (CSW 2.0.2) specification. Improvements to give users a more responsive and interactive experience have been substantial and include a new Web map viewer and a complete revision of search interface.

The use of International Standards

GeoNetwork has been developed following the principles of a Free and Open Source Software (FOSS) and based on International and Open Standards for services and protocols, like the ISO-TC211 and the Open Geospatial Consortium (OGC) specifications. The architecture is largely compatible with the OGC Portal Reference Architecture, i.e. the OGC guide for implementing standardised geospatial portals. Indeed the structure relies on the same three main modules identified by the OGC Portal Reference Architecture, that are focused on spatial data, metadata and interactive map visualisation. The system is also fully compliant with the OGC specifications for querying and retrieving information from Web catalogues (CSW). It supports the most common standards to specifically describe geographic data (ISO19139 and FGDC) and the international standard for general documents (Dublin Core). It uses standards (OGS WMS) also for visualising maps through the Internet.

Harvesting geospatial data in a shared environment

Within the geographic information environment, the increased collaboration between data providers and their efforts to reduce duplication have stimulated the development of tools and systems to significantly improve the information sharing and guarantee an easier and quicker access of data from a variety of sources without undermining the ownership of the information. The harvesting functionality in GeoNetwork is a mechanism of data collection in perfect accordance with both rights to data access and data ownership protection. Through the harvesting functionality it is possible to collect public information from the different GeoNetwork nodes installed around the world and to copy and store periodically this information locally. In this way a user from a single entry point can get information also from distributed catalogues. The logo posted on top each harvested record informs the user about the data source.

1.2 GeoNetwork and the Open Source Community Development

The community of users and developers of the GeoNetwork software has increased dramatically since the release of version 2.0 in December 2005 and the subsequent releases. At present, the user and

¹ <http://geonetwork-opensource.org>

developer mailing lists count well over 250 subscriptions each. Subscription to these lists is open to anyone interested. The archive of the mailing lists provides an important resource for users and can be freely browsed online. Members provide feedback within the community and provide translations, new functionalities, bug reports, fixes and instructions to the project as a whole. Building a self sustaining community of users and developers is one of the biggest challenges for the project. This community-building process relies on active participation and interaction of its members. It also relies on building trust and operating in a transparent manner, thereby agreeing on the overall objectives, prioritization and long term direction of the project. A number of actions have been taken by the project team to facilitate this process.

The foundation for the establishment of a GeoNetwork Advisory Board was laid at the 2006 workshop in Rome and membership criteria were defined.

A work plan is presented and discussed at the yearly GeoNetwork workshop; subsequently, the plan is maintained and updated throughout the year where needed. The project management team reports back to the advisory board about the reached developments and objectives during the annual workshops.

Two public Websites have been established. One focuses on the users of the software (<http://geonetwork-opensource.org>), while the other one is dedicated to the developers (<http://trac-osgeo.org/geonetwork>). Both can be updated and maintained online by trusted members of the community. They provide documentation, bug reporting and tracking, Wiki pages et cetera. A small part of the community connects through Internet Relay Chat (IRC) on a public #geonetwork² channel. But most interaction takes place on the user³ and the developer⁴ mailing lists.

During the 2006 workshop, the Project Advisory Board decided to propose the GeoNetwork opensource project as an incubator project to the newly founded Open Source Geospatial Foundation (OSGeo)⁵. This incubation process is currently ongoing but close to conclusions. The project Websites have been moved to servers accessible under the umbrella of the OSGeo foundation. Web pages have been updated to reflect the OSGeo principles and a source code review performed.

Source code is maintained in a publicly accessible code repository, hosted at an independent service provider, SourceForge.net⁶ that hosts thousands of FOSS projects. Developers and users have full access to all sections of the source code, while trusted developers can make changes in the repository itself. A special mailing list has been established to monitor changes in the code repository. This so-called "commit mailing list" delivers change reports by email to its subscribers.

The documentation is written in DocBook⁷ format to ensure versioning and support of multiple output formats (e.g. HTML and PDF).

² <irc://irc.freenode.net/geonetwork>

³ <https://lists.sourceforge.net/mailman/listinfo/geonetwork-users>

⁴ <https://lists.sourceforge.net/mailman/listinfo/geonetwork-devel>

⁵ <http://www.osgeo.org>

⁶ <http://sourceforge.net/projects/geonetwork>

⁷ <http://www.docbook.org>

2. Getting Started

Please make sure you have opened the home page of the GeoNetwork based catalogue.¹

There are many different ways to search the catalogue for maps and other geographic data. This guide will introduce you to the most popular search methods: default, advanced and by category. Whichever search you choose, remember that you will see results based on your privileges and assigned work group (Section 2.5, “Privileges, roles and user groups”).

Note that the term **data** in this application refers to datasets, maps, tables, documents, etc. that are linked to the metadata of a specific record.

2.1 Default Search

The default search allows you to search text within the entire record, such as keywords of the metadata and/or geographic location.

Free text search. *Type* a search term in the *What?* field. You can type anything here (free text). You can use quotes around text to find exact combinations of words.

Text and operators (and, or, not) are not case sensitive. (See Figure 2.1, “The free text field”).



Figure 2.1. The free text field

Geographic search. For the geographic search, two options are available for selecting a particular region to limit the search:

You can select a **region** from a predefined list (Figure 2.2, “The region field”);

¹If you installed the software on your local machine and started it, the default URL is <http://localhost:8080/geonetwork>



Figure 2.2. The region field

You can select your own *area of interest* in a more interactive way. A small global map is shown on the screen from which you can drag and drop the frame of your location area. Just click on the button on the upper right of the map screen (Figure 2.3, “Interactive Area Of Interest map”);

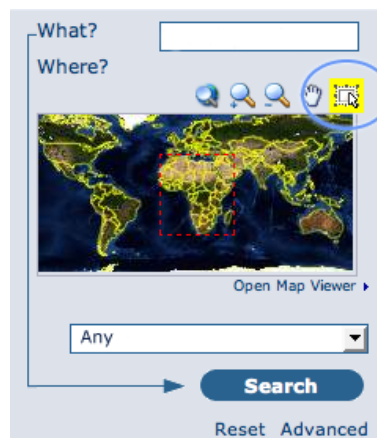


Figure 2.3. Interactive Area Of Interest map

Perform search. Both types of search, free text search and geographic search can be combined to restrict the query further.

Click the **Search** button to proceed and show the results. (Figure 2.4, “The Search button”).

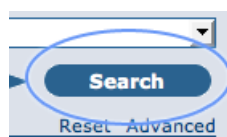


Figure 2.4. The Search button

2.2 Searching by Categories

An additional way to search data within the GeoNetwork database, from the home page, is searching by **Category**. A list of categories is provided to the user to identify data at a more generic level: **Applications, Audio/Video, Case study and best practises, Conference proceedings, Datasets, Directories, Interactive resources, Maps and graphics, Other information resources, Photo.**

To search only for maps, click on **Maps and Graphics** (Figure 2.5, “Search by Category”). A list of maps will be displayed from which you may view details of every single map; just clicking on the **Metadata** button of the map you wish to review.



Figure 2.5. Search by Category

2.3 Advanced Search

The advanced search option (Figure 2.6, “Advanced search options”) works similarly to the default search. However, you can be more specific in your search criteria as it offers different elements to look for data, each of them focusing one of the following aspects: **What?, Where?, When?**

Figure 2.6. Advanced search options

To perform an **advanced search**, from the home page **click Advanced** just below the search bottom (see Figure 2.7, “Show advanced search options”).

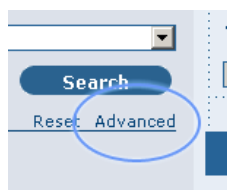


Figure 2.7. Show advanced search options

In the **WHAT?** section the elements are all related to the data content. Through them, in addition to searching only free keywords in the entire metadata content, you can also search directly in the title or abstract fields and add more keywords to customise your search further. You can also specify the level of accuracy you wish to reach in performing your search (Figure 2.8, ““What” section in the Advanced search”).

- To search by **Title**, **Abstract**, **Free Text**, or **Keyword(s)** type any text into the respective field. You can enter information in one or multiple field(s). If you do not want to search by a given field, simply leave it blank;
- You can choose the **accuracy of your search**, in terms of spelling words, from **Precise** = 1 to **Imprecise** = 0.2, through 3 more consecutive steps which are equal to 0.8, 0.6, 0.4.

Figure 2.8. "What" section in the Advanced search

The **WHERE?** parameters, which are related to the spatial extent, allow you, as in the default search, either to select your own area of interest or to select a predefined region from the drop-down list. In this section you can also type the geographic coordinates of a specific location that is not available from the above list. (Figure 2.9, "Where" section in the Advanced search")

- To select **your own area of interest**, drag and drop the frame of your area on the global map using the appropriate tool on the bottom left of the map screen;
- To use **free coordinates**, type the lat-long geographic references in the appropriate fields around the map screen, without any limitation of decimal figures;
- To use the coordinates of a **predefined region**, select the region from the drop-down list.

Figure 2.9. "Where" section in the Advanced search

Whatever type of geographic search you decide to perform, in the **Spatial search type** field, you can choose from different options: **is**, **overlaps**, **encloses**, **is fully outside of** (Figure 2.9, "Where" section in the Advanced search"). If you use this field, be cautious as this limits your output data as follows:

- If you choose **Spatial search type** *is* "Country", only maps for the selected country will be displayed. In other words, a city map within that country will not show in the output results.
- If you choose **Spatial search type** *overlaps* "Country", all maps with the bounding box overlapping that country will be displayed in the results, i.e. the neighbouring countries, the continent of which that country is part of and the global maps.
- If you choose **Spatial search type** *encloses* "Country" you will get, in the output results, maps of that country first and then all maps within its bounding box.

- Similarly, if you choose **Spatial search type** is *fully outside of* a selected region, only maps that follow that exact criteria will show in the output results.

The **WHEN?** section gives you the possibility to restrict your search in terms of temporal extent, indicating a specific range of time referred to the data creation or publication date (Figure 2.10, ““When” section in the Advanced search”).

- To specify a range of time, click on the date selector button next to **From – To** fields. Make use of the symbols > and >> on top of the calendar to select the month and the year first and then click on the exact day; a complete date will be filled in using the following standard order: YY-MM-DD.
- To clean the time fields, simply click on the white cross on their right; the box **Any** will be automatically selected and the search will be performed without any restriction on the time period.

Figure 2.10. “When” section in the Advanced search

Finally, the advanced search allows you to apply further restrictions on the basis of additional parameters as data source, data categories and data format (Figure 2.11, “Other options in the Advanced search”).

- To limit your queries to only one **Catalogue** out of those made available by the installation through the harvesting process, highlight the catalogue of preference or just keep **Any** selected to search all sites. (See more info about **data harvesting** in Section 4 Chapter 1 of these guidelines).
- To search for data organised by **Category**, such as Applications, Datasets, etc., simply highlight the category you wish to search in from the related drop-down list, otherwise we suggest to leave this field in **Any Category**.
- You can search for **Digital** or **Hard Copy** maps. To search in one or the other, simply check the box next to the one you wish to search. If no box is checked, all content will be searched.

At last, you can customise the number of output results per page in the **Hits Per Page** field. Simply highlight the number of records to be displayed or leave the field set on the default number (10).

- Click the **Search** button.

Figure 2.11. Other options in the Advanced search

2.4 Analysing Search Results

The output of a search provides you a list of the metadata records that should fit your request. For each record, the result page shows the title, an abstract and the keywords. According to the privileges that have been set for each metadata, a maximum of four sections can be consulted, as shown below. (Figure 2.12, “Search results”)



Figure 2.12. Search results

1. **Metadata:** The metadata section describes the dataset (e.g.: citation, data owner, temporal/spatial/methodological information) and could contain links to other web sites that could provide further information about the dataset.
2. **Download:** Depending on the privileges that have been set for each record, when this button is present, the dataset is available and downloadable. The process for retrieving data is simple and quick by just clicking the download button (Figure 2.13, “A single search result”) or by using the proper link in the specific metadata section for distribution info in the full metadata view (Figure 2.14, “Available services related to the resource”).

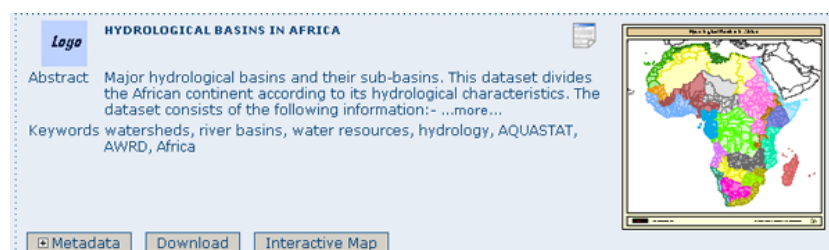


Figure 2.13. A single search result



Figure 2.14. Available services related to the resource

3. **Interactive Map:** The map service is also optional. When this button is shown, an interactive map for this layer is available and, by default, it will be displayed on the map screen of the simple search. To better visualise the map through the map viewer, **click on Open Map Viewer** on the bottom left of the map screen (Figure 2.15, “The interactive map viewer”).

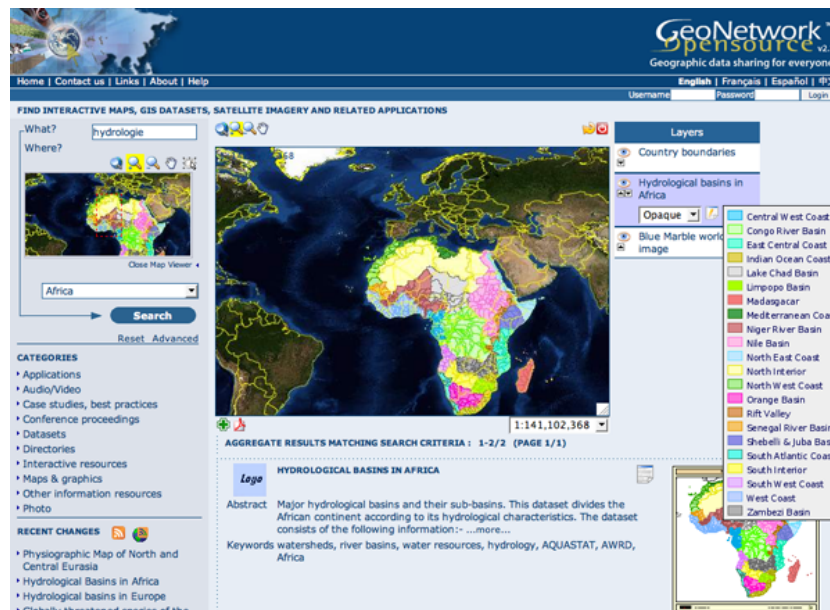


Figure 2.15. The interactive map viewer

4. **Graphic Overviews:** There are small and large overviews of the map used to properly evaluate usefulness of the data, especially if the interactive map is not available. Simply click on the small image to enlarge it. (Figure 2.16, “Large preview image”)

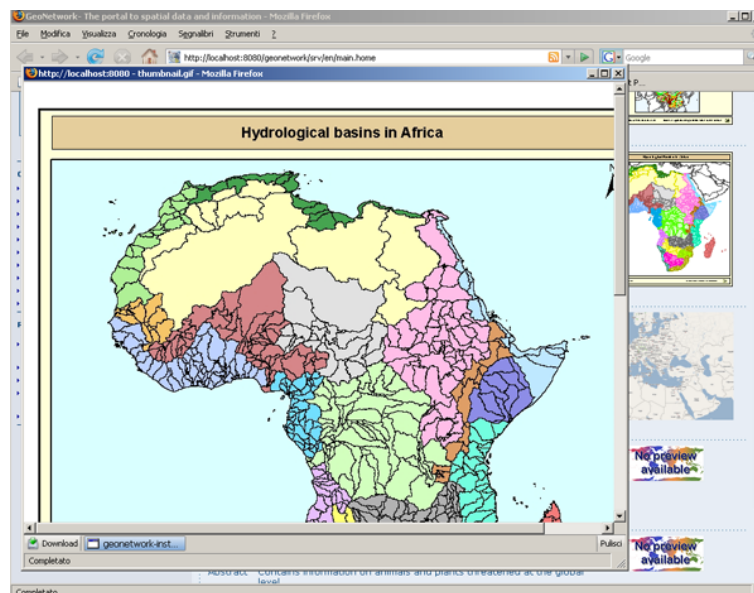


Figure 2.16. Large preview image

2.5 Privileges, roles and user groups

GeoNetwork uses a system of *Privileges*, *Roles* and *User groups*.

There are no restrictions for users to search and access **public information** in a GeoNetwork opensource based catalogue. To get access to **restricted information** or advanced functionality, an account to log in is required. This should be provided by the GeoNetwork administrator.

To log in, simply go to the home page and enter your username and password in the dedicated fields on the top right corner, then click the login button. (See Figure 2.17, “Login”)



English Français Español 中文		
Username	Password	Login

Figure 2.17. Login

Privileges. Depending on the privileges set on a metadata record and on your role as an authenticated user, you will be able to read about a resource and download or interactively browse data related to that resource.

Roles. Users with an *Editor* role can create, import and edit metadata records. They can also upload data and configure links to interactive map services.

User groups. Every authenticated user is assigned to a particular work group and is able to view data within that work group.

3. Viewing and Analysing the Data

Once you have completed your search, you view details of a particular record by clicking on the **Metadata** button.

The metadata profiles used by GeoNetwork opensource to present and describe geographic data and general documents stored in the catalogue are based on the **International Standard ISO 19115:2003**, encoded according to the **implementation schema 19139:2007**, the **FGDC** and the international standard **Dublin Core**.

In this guide the **ISO 19139** metadata implementation will be described in details since it is also suggested as profile for the creation of new metadata records.

3.1 Meta Data Description

The metadata ISO 19139 profile used by GeoNetwork opensource to describe the geographic data and services is based on the ISO standard 19115:2003 and provides information related to the identification, the maintenance and constraints, the spatial and temporal extent, the spatial representation and reference, the quality and distribution of a geographic dataset.

The metadata profile is organised in sections and the most important, illustrated in Figure 3.1, “Main metadata sections”, are the: **Identification Section, Distribution Section, Reference System Section, Data Quality Section and Metadata Section**. These sections are described here in details.

Identification info	Identification info basic information about the resource(s) to which the metadata applies mandatory
Title	
Date	
Date type	Publication
Abstract	The ISO19115 metadata standard is the preferred metadata standard to use. If unsure
Maintenance and update frequency	asNeeded
Language	English
Character set	utf8
Topic category code	boundaries
Distribution info	Distribution info provides information about the distributor of and options for obtaining the resource(s)
OnLine resource	
OnLine resource	http://localhost:8080/geonetwork/srv/en/resources.get?id=13&fname=&access=private
Reference system info	Reference system info description of the spatial and temporal reference systems used in the dataset
Code	
Data quality info	Data quality info provides overall assessment of quality of a resource(s)
Hierarchy level	
Metadata	Metadata info information about the metadata itself mandatory
File identifier	oca7b99
Language	English
Character set	utf8
Date stamp	2007-10-25T18:03:03
Metadata standard name	ISO 19115:2003/19139
Metadata standard version	1.0
Metadata author	

Figure 3.1. Main metadata sections

Identification Section

This section includes information on the citation of the resource (**title, date of creation or publication, edition, presentation form**), the **abstract**, the **purpose** and the present **status** of the resource that can be defined among the options: completed, historical archive, obsolete, ongoing, planned, required or under development. (Figure 3.2, “Identification information”).

Identification info
basic information about the resource(s) to which the metadata applies
mandatory

::Identification info *

Title *

Date *

Date type * Publication

Edition

Presentation form

Abstract *
The ISO19115 metadata standard is the preferred metadata standard to use. If unsure what templates to start with, use this one.

Purpose

Status

Figure 3.2. Identification information

This section also contains information about the person or organisation responsible for the data and who is considered to be a **point of contact** for the resource i.e. the dataset owner, originator, distributor, publisher, etc. and it provides information on **data maintenance** i.e. annually, monthly, daily, not planned, as needed, etc. (Figure 3.3, “Point of Contact”)

Point of contact
identification of, and means of communication with, person(s) and organizations(s) associated with the resource(s)

::Point of contact *

Individual name *

Organisation name *

Position name *

Voice

Facsimile

Delivery point

City

Administrative area

Postal code

Country

Electronic mail address

Role *

Maintenance and update frequency *

Figure 3.3. Point of Contact

Elements for **keywords** and for describing restrictions on **data access** and **use** are also included in this section in addition to **spatial representation** info like data type (vector, raster, text table, etc.) (Figure 3.4, “Descriptive keywords”).

Figure 3.4. Descriptive keywords

The identification section provides information about the **scale**, the **language** and **character set** used within the resource and the list of **ISO categories** through which your map could be classified (Figure 3.5, “Scale and other data properties”).

Figure 3.5. Scale and other data properties

Finally, the temporal and spatial extent are also defined in this section. The temporal extent is defined through the starting and ending date of data validation (Figure 3.6, “**Temporal extent**”);

Figure 3.6. **Temporal extent**

The spatial extent of the interested area is defined through geographic coordinates or through the selection of a country or region from a predefined list (Figure 3.7, “**Geographic bounding box**”). Free text **supplemental information** can be added to complete the data identification section.

Geographic bounding box
geographic position of the dataset

Geographic bounding box

World

North bound latitude *
90

West bound longitude *
-180

East bound longitude *
180

South bound latitude *
-90

Supplemental Information

You can customize the template to suit your needs. You can add and remove fields and fill out default information (e.g. contact details). Fields you can not change in the default view may be accessible in the more comprehensive (and more complex) advanced view. You can even use the XML editor to create

Dropdown list of countries, the used map belongs to.

Figure 3.7. **Geographic bounding box**

Distribution Section

This section provides metadata elements for **accessing** other useful **on-line resources** available **through the web**. The distribution elements allow for on-line access using an URL address or similar addressing scheme and provide the protocol for the proper connection **for accessing geographic data** or any other types of digital documents **using the download function**. Furthermore, it is possible **to link** a metadata with a **predefined map service** through the **on line resource** and see the map interactively (Figure 3.8, “Distribution information”).

Distribution info
provides information about the distributor of and options for obtaining the resource(s)

OnLine resource
information about online sources from which the resource can be obtained

OnLine resource

URL

Protocol *
Web address (URL)

Description *

OnLine resource

URL
http://localhost:8080/geonetwork/srv/en/resources.gd

Protocol *
File for download

File
Browse...

Description *

OnLine resource

URL

Protocol *
OGC Web Map Service (ver 1.1.1)

Name *

Description *

Browse to the folder where the file you want to upload is stored

Uploads the file

Dropdown list of the type of protocol you want to provide as link or to download files from

Upload

Figure 3.8. **Distribution information**

Reference System Section

The Spatial Reference System section **defines metadata** required **to describe the spatial reference system** of a dataset. It contains one element to identify the name of the reference system used (Figure 3.9, “Reference system”). Using elements **from the advanced form**, this section may be modified to provide more **details on data projection, ellipsoid and datum**. Note that if this information is provided, a reference system identifier is not mandatory.

Reference system info
description of the spatial and temporal reference systems used in the dataset

::Reference system info

Code * WGS 1984

Figure 3.9. Reference system

Data Quality Section

The Data Quality section provides a **general assessment** of the **quality of the data**. It describes the **different hierarchical levels of data quality**, namely a **dataset series**, **dataset**, **features**, **attributes**, etc. This section also contains information about **sources of the input data**, and a **general explanation of the production processes** (lineage) used for creating the data (Figure 3.10, “Data quality”).

Data quality info
provides overall assessment of quality of a resource(s)

::Data quality info

Hierarchy level * dataset

Statement * [Empty text area]

Figure 3.10. Data quality

Metadata Information Section

This section contains information about the metadata itself: the *Universally Unique Identifier (UUID)* assigned to the record (this is the ‘File identifier’), **language** and **character set** used, **date of last edit** (‘Date stamp’) and the **metadata standard** and **version name** of the record. It also contains information on the **metadata author** responsible for the metadata record; this person can also be a point of contact for the resource described. Information on the Metadata author is mandatory (Figure 3.11, “Metadata properties”).

Metadata author
party responsible for the metadata information
mandatory

::Metadata

File identifier 0a38e56d-9519-492f-ace1-20a155ca7b99

Language * English

Character set utf8

Date stamp 2007-10-24T15:07:21

Metadata standard name ISO 19115:2003/19139

Metadata standard version 1.0

::Metadata author

Individual name * [Empty text area]

Organisation name * [Empty text area]

Position name * [Empty text area]

Voice [Empty text area]

Facsimile [Empty text area]

Delivery point [Empty text area]

City [Empty text area]

Administrative area [Empty text area]

Postal code [Empty text area]

Country [Empty text area]

Electronic mail address [Empty text area]

Role * pointOfContact

Figure 3.11. Metadata properties

4. Adding new Data and Information

This section guides you through the process of adding new records (new data with information) into the GeoNetwork catalogue using either the online metadata editor or the advanced metadata insert tool, based on XML documents. In both cases you will use the template system, add thumbnails, upload data, link to services and set access privileges to the metadata and data.

To add or edit data and metadata, you must be **registered** as an **Editor** into the work group you want to add records to. Contact your administrator if you are not a registered editor for your work group.

For the metadata creation using the online editor, GeoNetwork provides a set of simplified metadata templates based on the cited standards: *ISO*, *FGDC* and *DC*. The templates for vector and raster based on the ISO 19139 are the preferred ones since they are devised in a way that hides the complexity of the ISO19115 standard in the default view. At the same time those templates are extensible with new elements to fit specialised needs through the advanced view.

To produce a good metadata record, always try to gather as much details as possible on the resource that you want to describe taking into account the metadata elements that have been presented in the previous chapter. The next step is to fill out properly the fields provided by the metadata templates, while at the same time avoiding duplication of information throughout the form.

The most important, mandatory fields that should not be skipped while compiling a standard based metadata record are the following:

1. Title
2. Date of Creation or Publication
3. Abstract
4. Language used for documenting data
5. Topic Category
6. Scale
7. Maintenance and Update Frequency
8. Metadata Author
9. Language Used for Documenting Metadata

In addition to the main mandatory fields, we recommend you to fill out these optional but critical fields (if information is available):

1. Purpose
2. Keywords
3. Presentation Form
4. Status
5. Spatial Representation Type
6. Geographic Location
7. Reference System Info

- 8. Temporal Extent
- 9. Data Quality Info
- 10. Access and Use Constraints
- 11. Point of Contact
- 12. Distribution Info: Online Resources

You should also prepare an image of your data that is required to be displayed in search results as thumbnail.

Next section will guide you through the process of metadata creation using the online editor.

4.1 Creating a New Record using the Metadata Editor on line

1. In the home page, click on the **Administration Tab**.
2. Select **New Metadata** from the List of the admin page.
3. Select the metadata standard **Template**, if possible, using the preferred ones (Figure 4.3, “Template selection”). GeoNetwork opensource comes by default with support for three metadata standards, ISO19139, FGDC and Dublin core. For the ISO standard, two templates have been developed; one for vector and one for raster data. Both contain a relevant set of elements to describe the respective types of data. More templates can be developed online.
4. Select the **Group** the metadata will belong to. These are the groups authorised to add metadata to by your administrator.
5. Click on **Create**.

The steps in more details

1. Enter your username and password and click on the login button (Figure 4.1, “Login”). The system will identify you and assign the correct privileges to work with.

Figure 4.1. Login

2. Open the Administration page by clicking the Administration button in the banner and then click on the New metadata link (Figure 4.2, “Administration panel”).

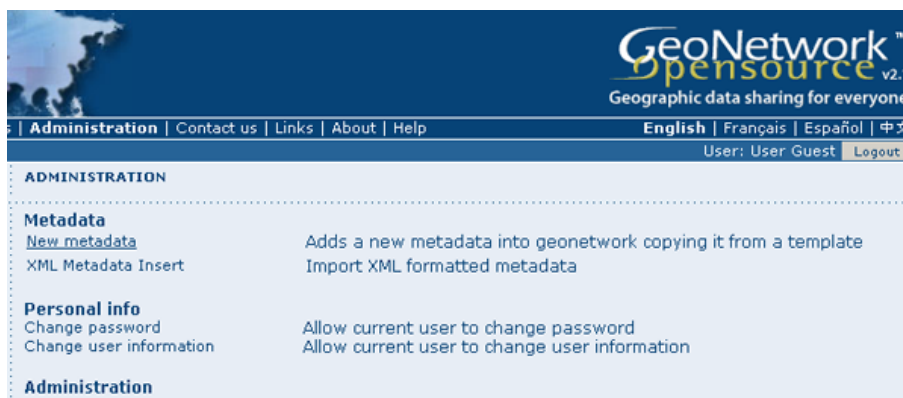


Figure 4.2. Administration panel

- From the metadata creation page, select the metadata standard to use from the dropdown list (Figure 4.3, “Template selection”)

The screenshot shows a web form titled "METADATA CREATION". It has two dropdown menus. The first is labeled "Template" and is currently open, showing four options: "Template for Dublin Core", "Template for FGDC", "Template for Raster data in ISO19139", and "Template for Vector data in ISO19139 (preferred!)". The second dropdown menu is labeled "Group" and is currently closed, showing "Sample group" as the selected option. Below the dropdowns are "Back" and "Create" buttons.

Figure 4.3. Template selection

- After selecting the correct template, you should identify which group of users the metadata will belong to (Figure 4.4, “Group selection”) and finally click on **Create**.

The screenshot shows the same "METADATA CREATION" form. The "Template" dropdown is now closed and shows "Template for Vector data in ISO19139 (preferred!)". The "Group" dropdown is open, showing two options: "Sample group" and "Sample group". The "Back" and "Create" buttons are visible below the dropdowns.

Figure 4.4. Group selection

A new metadata form based on the selected template will be displayed for you to fill out.

Switching Editing Views from Default to Advanced to XML View

Once you create a new record, you can choose between **Default**, **Advanced** or **XML View**. To switch view, simply click on the view you want to switch to on the left column of the page. The view in **bold** is the view you are currently using (Figure 4.5, “Metadata view options”).

The screenshot shows a web page for GeoNetwork. At the top right is the logo "GeoNetwork opensource" and "Geographic data sharing". Below the logo is a navigation bar with links: "Home", "Last results", "Administration", "Contact us", "Links", "About", "Help". To the right of the navigation bar are language options "English" and "Français" and a user profile "User: admin". On the left side, there is a vertical menu with three options: "Default view" (which is bolded), "Advanced view", and "XML view". To the right of this menu are three buttons: "Reset", "Save", and "Save and close". Below the menu is a yellow box with the text "::Identification info".

Figure 4.5. Metadata view options

In the previous chapter you have analysed the metadata structure as it is presented in the **Default View**. A selection of the main fields from different categories of information is shown in one single view. The minimum set of metadata required to serve the full range of metadata applications (data discovery, determination of data fitness for use, data access, data transfer and use of digital data) is defined here, along with optional metadata elements to allow for a more extensive standard description of geographic data, if required. However, if should be there a need to add more metadata elements, you can switch to the advanced view at any time while editing.

In the **Advanced View**, the ISO profile offers the possibility to visualise and edit the entire metadata structure organised in sections accessible through tabs from the left column. You can use this view to write more advanced metadata descriptions or templates to fit specialised needs. (Figure 4.6, “Advanced view”)

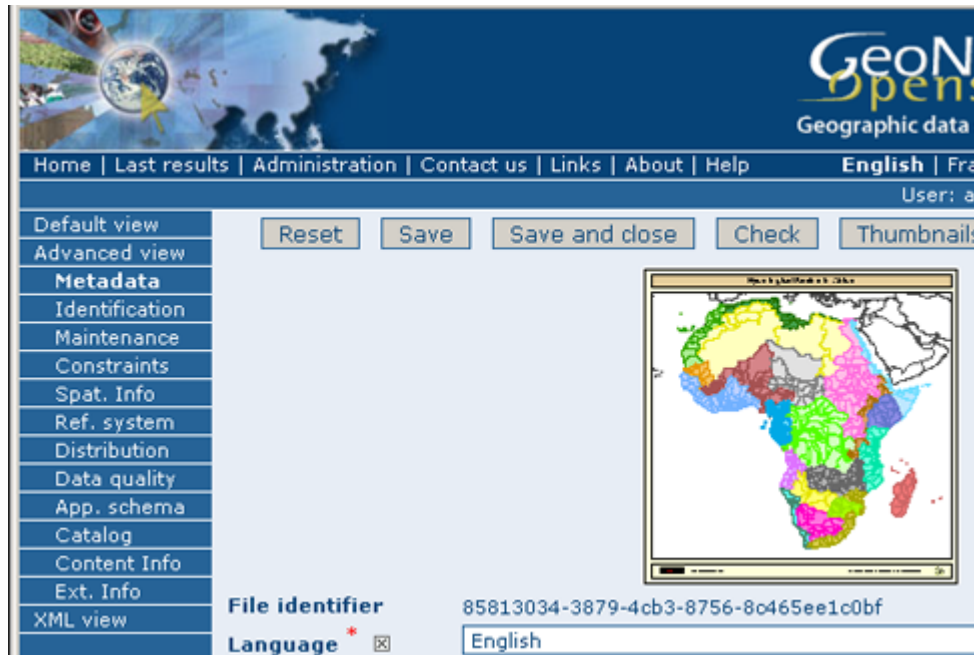


Figure 4.6. Advanced view

The **XML View** shows the entire content of the metadata in the original hierarchical structure; different colours allow to distinguish between an element's name and its value. The XML structure is composed of tags and to every tag must correspond a closing tag (Figure 4.7, "XML view"). The content is entirely contained within the two, i.e.

```
<gmd:language>
  <gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
```



Figure 4.7. XML view

Nevertheless, the use of the XML view requires some knowledge of the XML language.

Both the **Default** and the **Advanced Views** are composed of mandatory, conditional and optional metadata fields. The meaning of mandatory and optional is fairly intuitive; the mandatory

fields are required, like `Title` and `Abstract` for instance, whereas the optional fields can be provided but are not fundamental, depending on the metadata author. The conditional fields may be considered mandatory under certain circumstances: essentially a conditional requirement indicates that the presence of a specified data element is dependent on the value or presence of other data elements in the same section. For instance, the `Individual name` metadata element of the `Point of Contact`, which is a conditional element of the `Identification` section, becomes mandatory if another element of the same section, `organisation name` or `Position name` is not already defined (Figure 4.8, “Point of Contact”).

Figure 4.8. Point of Contact

The **mandatory fields** as well as those **highly recommended** are flagged with **red asterisk [*]**. The standard definition for each field can be read by passing the mouse on the element name.

The **Default View** is the preferred view as it provides a selection of the available metadata elements, facilitating both the user and the editor in reading and editing a metadata record, and at the same time it ensures that a geospatial data can be properly described, through:

- the minimum set of metadata required to serve the full range of metadata applications (data discovery, determination of data fitness for use, data access, data transfer, and use of digital data);
- optional metadata elements - to allow for a more extensive standard description of geographic data, if required;
- a method for extending metadata to fit specialised needs.

Using basic commands of the editor

Fields are either **free text fields** or **drop down lists**. Free text means you can type any text into that field. Drop down lists allow you to select only one option from the list. You can add multiple fields of the

same kind by clicking on the **[+]** symbol next to the element. Every new field that you will add in the advanced view will then be visible in the default view. You can also delete existing fields by clicking on the **[x]** symbol next to the element. Clearly, mandatory fields cannot be deleted. One example of the need to add multiple fields can arise if the content of your dataset has some text written in two different languages (Figure 4.9, “Describing multilingual data”).

Figure 4.9. Describing multilingual data

4.2 Entering Metadata for your Map

As we mentioned in the introduction to this guide, GeoNetwork opensource provides tools to describe any type of geographic data (vector layers, raster, tables, map services, etc.) as well as general document like reports, projects, papers, etc. For the purpose of this Quick Start Guide, an example of required and useful metadata elements to properly describe a thematic map will be provided hereafter. You should gather as much information as possible to identify and understand the map’s resource and characteristics you want to describe. Use the default view to start. If necessary, you can always switch to advanced view or come back later and edit the record with the additional information collected.

Entering Metadata For Your Map

Please follow these steps to enter your map’s metadata. Note that we will only go through the fields that have been identified as compulsory (i.e. those fields marked with the asterisk **[*]**, mandatory or highly recommended).

Title *: Under the **Identification Info** field, give your map a **name**. There will be a default name of your data. Use free text to describe your map here.

Date *: Indicate the exact **date** of **creation**, **publication** or **revision** on your map.

Presentation Form: Specify the type of presentation, i.e. **digital**, **hard copy**, **table**, etc.

Abstract *: Enter some description of the map.

Purpose: Enter a short summary of the purposes for your map to be developed.

Status: Specify the status of your map within the following options: completed, historical archive, obsolete, ongoing, planned, required, under development.

Point of Contact: Enter all mandatory information and others you have at hand for the contact of the person(s) associated with this resources of the map. Note that some fields are only conditionally mandatory, such as organisation Name if Individual Name and Position are not entered.

Maintenance and update frequency *: Specify the frequency with which you expect to make changes and additions to your map after the initial version is completed. If any changes are scheduled you can leave `As Needed` selected from the drop-down list.

Descriptive Keywords: Enter keywords that describe your map. Also specify the type of keyword you are entering, i.e. place, theme, etc. Remember that you can add another keyword field if you need to add different types of keywords.

Access Constraints: Enter an access constraint here, such as a copyright, trademark, etc. to assure the protection of privacy and intellectual property.

User Constraints: Enter a user constraint here to assure the protection of privacy and intellectual property.

Other Constraints *: Enter other constraint here to assure the protection of privacy and intellectual property. Note that this field is conditionally mandatory if Access and Use constraints are not entered.

Spatial representation type: Select, from the drop-down list the method used to spatially represent your data. The options are: vector, grid, text table, stereo model, video.

Scale Denominator *: Enter the denominator for an equivalent scale of a hard copy of the map.

Language *: Select the language used within your map

Topic category *: Specify the main ISO category/ies through which your map could be classified (see Annex for the complete list of ISO topic categories).

Temporal Extent *: Enter the starting and ending date of the validity period.

Geographic Bounding Box *: Enter the longitude and latitude for the map or select a region from the predefined drop-down list. Make sure you use degrees for the unit of the geographic coordinates as they are the basis for the geographic searches.

Supplemental Information: Enter any other descriptive information about your map that can help the user to better understand its content.

Distribution Info: Enter information about the distributor and about options for obtaining your map.

Online Resource: Enter information about online resources for the map, such as where a user may download it, etc. This information should include a link, the link type (protocol) and a description of the resource.

Reference System Info: Enter information about the spatial reference system of your map. The **default view contains** one element to provide the **alphanumeric value** identifying the reference system used. GeoNetwork opensource uses the **EPSG codes** which are numeric codes associated with coordinate system definitions. For instance, EPSG:4326 is Geographic lat-long WGS84, and EPSG:32611 is "UTM zone 11 North, WGS84". Using elements **from the advanced view**, you may add **more details on data projection, ellipsoid and datum**. Note that if this information is provided, a reference system identifier is not mandatory.

Data Quality: Specify the hierarchal level of the data (**dataset series, dataset, features, attributes**, etc.) and provide a **general explanation on the production processes** (lineage) used for creating the data. The statement element is mandatory if the hierarchical level element is equal to `dataset` or

series. Detailed information on **completeness**, **logical consistency** and **positional**, **thematic** and **temporal accuracy** can be directly added **into the advanced form**.

Metadata Author *: Provide information about the author of the map, including the person's name, organisation, position, role and any other contact information available.

After completion of this section, you may select the **Type** of document that you are going to save in the catalogue. You have three options: **Metadata**, **Template**, **Sub-template**. By default **Metadata** is set up.

When done, you may click *Save* or *Save and Close* to close the editing session.

Creating a Thumbnail

Next, you need to create a graphic overview of your map which will be for a double purpose; as small thumbnail will be displayed in search results and as large thumbnail with much more details, to allow users to properly evaluate the data usefulness. As for the latest, the image that you will use as source should be a significant reproduction of the real dataset, possibly inclusive of the legend.

To create a thumbnail, go to the editing menu for your map. If you are no longer in editing mode, retrieve the map from one of the search options then click on *Edit*. Then follow these simple steps:

- From the editing menu, click on the *Thumbnails* button on the top or bottom of the page. (Figure 4.10, "The thumbnail wizard button")

Figure 4.10. The thumbnail wizard button

- You will be taken to the *Thumbnail Management wizard* (Figure 4.11, "Thumbnail wizard").
- To create a small or large thumbnail, click on the *Browse* button next to either one. It is recommended that you use 180 pixels for small thumbnails and 800x600 for large thumbnails. Using the 'Large thumbnail' option allows you to create both a small and large thumbnail in one go.
- You can use GIF, PNG and JPEG images as input for the thumbnails.
- A pop up window will appear allowing you to browse your files on your computer. *Select* the file you wish to create a thumbnail with by double-clicking on it.
- Click on *Add*.
- Your thumbnail will be added and displayed on the following page.

- You can then click on **Back to Editing** and save your record (Figure 4.12, “Completed thumbnail wizard”).

The screenshot shows a 'THUMBNAIL MANAGEMENT' interface. It is divided into two main sections: 'Small thumbnail' and 'Large thumbnail'. Each section contains a 'Browse...' button and an 'Add' button. Below each 'Add' button are three radio button options: 'Scale image' (checked), 'Scale width', and 'Scale height'. The 'Scale image' option has a dropdown menu showing '180 pixel' for the small thumbnail and '800 pixel' for the large thumbnail. At the bottom of the interface is a 'Back to editing' button.

Figure 4.11. Thumbnail wizard

The screenshot shows the 'THUMBNAIL MANAGEMENT' interface after completion. The 'Small thumbnail' section now displays a preview of a map of Africa, the filename 'thumbnail_s.png', and a 'Delete' button. The 'Large thumbnail' section displays a larger preview of the same map, the filename 'thumbnail.png', and a 'Delete' button. A 'Back to editing' button is located at the bottom.

Figure 4.12. Completed thumbnail wizard

Linking data for download

Finally, you can upload the dataset stored on your local computer and then create a link between data and related description. Files in whatever format can be uploaded: doc, PDF, images, vector layers, etc. For the latter the distribution in a compressed file is recommended. You can include the vector data, the legend, any documentation that can help the interpretation of the data, related reports, detailed descriptions of the data processing, base data used to create the dataset specified and/or other relevant information. Follow these guidelines for uploading datasets:

- Make sure the total size of the compressed file is reasonable (less than 50 MB). Should your data be bigger than 50MB, consider a different mechanism to serve this data, e.g. through an FTP or HTTP server and then link the resource through an online resource 'Web address (URL)'.
- You can create several smaller files when appropriate and upload them sequentially.
- You add the size of the file at the end of the description field.

To Upload a Dataset, follow these steps (Figure 4.13, "An online resource"):

1. The **URL** field can be left empty when uploading a file. The system will automatically fill this field out;
2. Select the correct **protocol** to be used. If you do not see the buttons to browse and upload when File for download is selected, save the metadata and return to the upload section. Both buttons should appear;
3. Provide a short **description** of the data;
4. Click the **Browse** button and navigate to the folder where the file to be released is stored. Consider if you want to upload multiple files as one unique zip file or as multiple separate downloads. It is a good idea to add additional documentation with the datasets that provide the user with information related to the data described. Remind: the size of a single file to upload can't exceed 100 Megabytes;
5. Click **Upload** and then **Save**.

Figure 4.13. An online resource

Assigning Privileges for a Map

As an important step of entering metadata to your map, you need to assign privileges for each map. This means that you will identify which work groups have which privileges, i.e. view, download, etc. for your particular map.

For instance, you can define if the information and related services is visible to all (Internet users) or just to internal users only (Intranet). Privileges are assigned on a per group basis. Depending on the user profile (Registered User, Editor, Content Reviewer, User Administrator and Administrator) access to these functions may differ on a per user basis.

To assign privileges for your map, follow these steps:

- Find your map by using the search option. Whether you have multiple or single results from the search, on top of the individual record or next to the record you will always see a row of buttons including a Privileges button (Figure 4.14, "The editing toolbar with Privileges button").



Figure 4.14. The editing toolbar with Privileges button

- Click on the Privileges button. This will open the privileges settings window. You can assign certain privileges to specific groups by selecting or deselecting them from this page. Simply click on the small box next to the privilege to place or remove a checkmark. **Set All** and **Clear All** buttons allow you to place and remove the checkmarks all at once (Figure 4.15, “Privileges settings”).

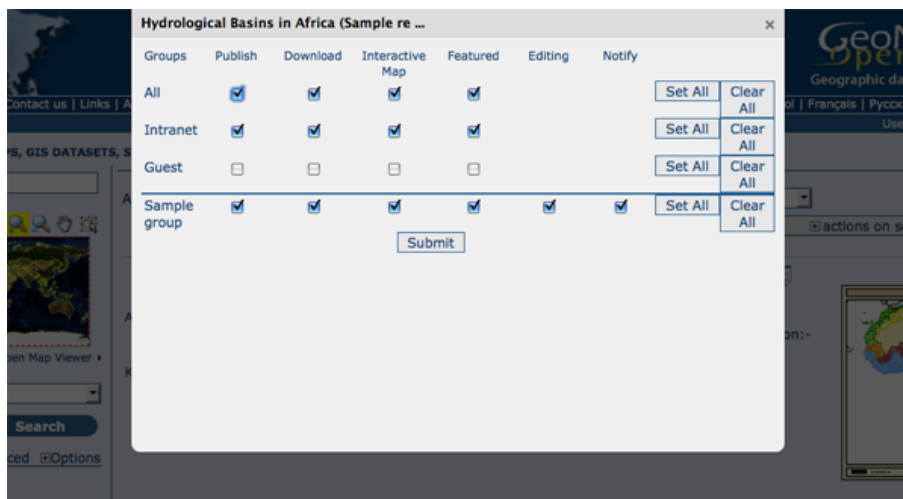


Figure 4.15. Privileges settings

Below is a brief description for each privilege to help you identify which ones you should assign to which group(s).

Publish: Users in the specified group/s are able to see the map, i.e. if searching with matching criteria.

Download: Users in the specified group/s are able to download the map.

Interactive Map: Users in the specified group/s are able to get an interactive map. **The interactive map** has to be created separately using a Web Map Server, which is part of the GeoNetwork opensource application.

Featured: When selected, the map is placed in the Features Maps of the home page and it appears there randomly.

Editing: When selected, the editors of the group(s) concerned can edit the respective metadata record.

Notify: A notification email is send to the emailaddress of the group, informing that the map has been downloaded.

Assigning Categories for a Map

As a final step to entering metadata for a map, you should assign categories for it. The assigned categories will determine the categories the map will display under on the home page. To assign categories for a map, follow these steps:

- Find your map by using the search option. Whether you have multiple or single results from your search, on top of the individual record or next to the record, you will always see a row of buttons including a **Categories** button.
- Click on the **Categories** button. This will open the category selection window. You can assign one or multiple categories selecting or deselecting them from this page. Simply click on the small box next to the category to place or remove a checkmark. (Figure 4.16, “Category management”)

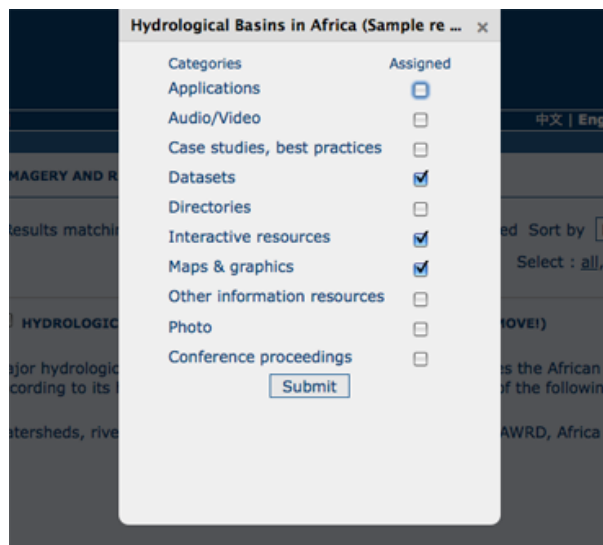


Figure 4.16. Category management

4.3 Uploading a New Record using the XML Metadata Insert Tool

A more advanced procedure to upload a new metadata record in the GeoNetwork system is using an XML document. This procedure is particularly useful for users who already have metadata in XML format, for instance created by some GIS application. To this regard, it has to be noted that the metadata must be in one of the standards used by GeoNetwork: ISO19115, FGDC and Dublin Core.

To start the metadata uploading process through the **XML Metadata Insert** tool, you should log in (see Step. 1. in paragraph 7.1.1) and select the appropriate option from the Administration page (Figure 4.17, “Administration panel”).

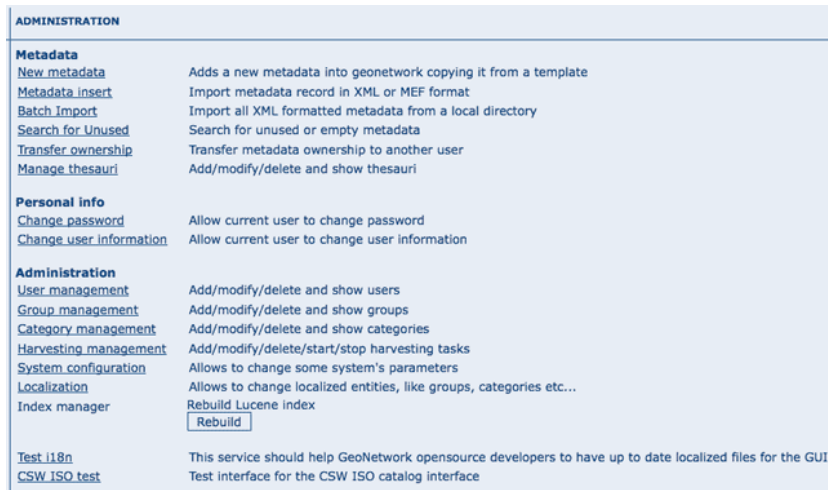


Figure 4.17. Administration panel

The main part of the page **Import XML Formatted Metadata** that is displayed (Figure 4.18, “XML metadata import tool”) is the **Metadata** text area, where the user can paste the XML metadata to import. Below this, there is the **Type** choice, which allows you select the type of record that you are going to create (Metadata, Template and Subtemplate). Then you can apply a stylesheet to convert your metadata input from ArcCatalog8 to ISO1915 or from ISO19115 to ISO19139, if required. Otherwise you can just leave *none* selected. The **Destination schema** list provides you with four options to choose the final standard layout for your metadata (ISO19115, ISO19139, FGDC and Dublin Core). Finally you should select the **Group** as main group in charge of the metadata and the **Category** that you want to assign to your metadata. By clicking the **Insert** button the metadata is imported into the system; please note that all links to external files, for instance to thumbnails or data for download, have to be removed from the metadata input, to avoid any conflict within the data repository.

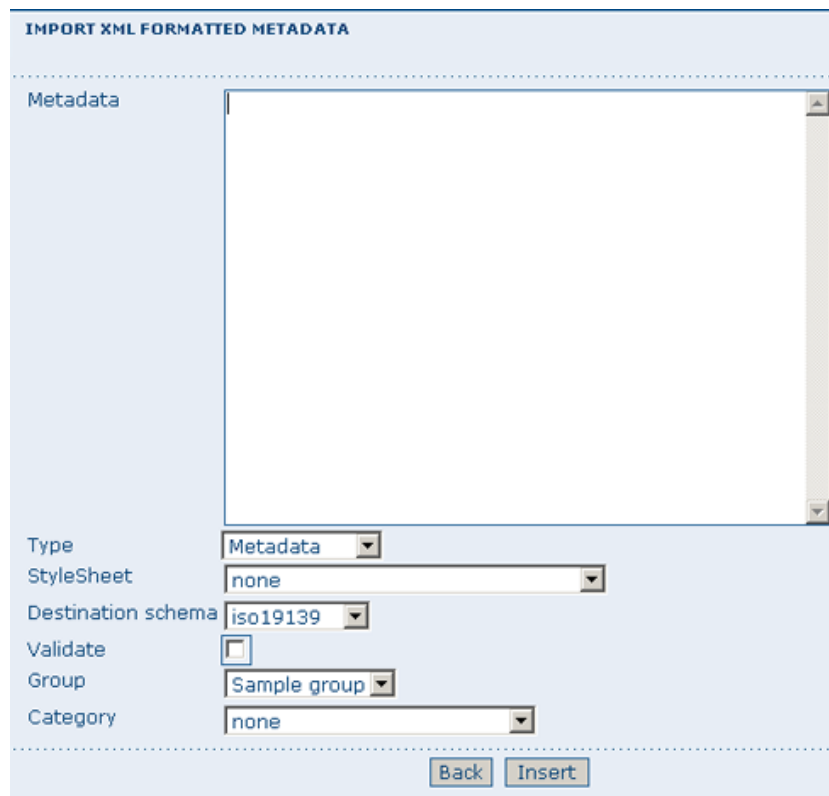


Figure 4.18. XML metadata import tool

If your metadata is already in ISO19115 format, the main actions to be performed are the following (Figure 4.19, "XML metadata import 2"):

1. Paste the XML file that contains the metadata information in the **Metadata** text area;
2. Select Metadata as **type** of record that you are going to create
3. Select the metadata schema ISO19139 that will be the final **destination schema**;
4. Select the **validate** check box if you want your metadata to be validated according to the related schema.
5. Select the **group** in charge of the metadata from the drop down list;
6. Select **Maps and Graphics** from the list of categories;
7. Click the **Insert** button and the metadata will be imported into the system.

IMPORT XML FORMATTED METADATA

Metadata

```
<?xml version="1.0" encoding="UTF-8"?>
<gmd:MD_Metadata xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:gml="http://www.opengis.net/gml"
xmlns:gts="http://www.isotc211.org/2005/gts"
xmlns:gco="http://www.isotc211.org/2005/gco">
<gmd:fileIdentifier>
<gco:CharacterString></gco:CharacterString>
</gmd:fileIdentifier>
<gmd:language>
<gco:CharacterString>eng</gco:CharacterString>
</gmd:language>
<gmd:characterSet>
<gmd:MD_CharacterSetCode codeListValue="utf8"
codeList="/resources/codeList.xml#MD_CharacterSetCode"/>
</gmd:characterSet>
<gmd:contact>
<gmd:CI_ResponsibleParty>
<gmd:individualName>
<gco:CharacterString/>
```

Type: Metadata

StyleSheet: none

Destination schema: iso19139

Validate:

Group: Sample group

Category: Maps & graphics

Back Insert

Figure 4.19. XML metadata import 2

5. Metadata in Spatial Data Management

5.1 What is Metadata?

Metadata, commonly defined as "data about data" or "information about data", is a structured set of information which describes data (including both digital and non-digital datasets) stored in administrative systems. Metadata may provide a short summary about the content, purpose, quality, location of the data as well as information related to its creation.

5.2 What are Metadata Standards?

Metadata standards provide data producers with the format and content for properly describing their data, allowing users to evaluate the usefulness of the data in addressing their specific needs.

The standards provide a documented, common set of terms and definitions that are presented in a structured format.

5.3 Why do we need Standardised Metadata?

Standardised metadata support users in effectively and efficiently accessing data by using a common set of terminology and metadata elements that allow for a quick means of data discovery and retrieval from metadata clearinghouses. The metadata based on standards ensure information consistency and quality and avoid that important parts of data knowledge are lost.

5.4 Geographic Information Metadata Standard

Geographic data, which can be defined as any data with a geographic component, is often produced by one individual or organisation, and may address the needs of various users, including information system analysts, programme planners, developers of geographic information or policy makers. Proper standard documentation on geographic data enable different users to better evaluate the appropriateness of data to be used for data production, storage, update.

The metadata standards supported by GeoNetwork opensource are the **ISO 19115:2003** - approved by the international community in April 2003 as a tool to define metadata in the field of geographic information - and the **FGDC** - the metadata standard adopted in the United States by the Federal Geographic Data Committee. In addition, GeoNetwork opensource supports also the international standard **Dublin Core** for the description of general documents.

This ISO Standard precisely defines how geographic information and related services should be described, providing mandatory and conditional metadata sections, metadata entities and metadata elements. This standard applies to data series, independent datasets, individual geographic features and feature properties. Despite ISO 19115:2003 was designed for digital data, its principles can be extended to many other forms of geographic data such as maps, charts, and textual documents as well as non-geographic data.

The underlying format of an ISO19115:2003 compliant metadata is XML. GeoNetwork uses the ISO Technical Specification 19139 Geographic information - Metadata - XML schema implementation for the encoding of this XML.

5.5 Metadata profiles

GeoNetwork supports multiple metadata profiles. Profiles can take the form of `Templates` that you can create using the metadata editor. Using the Advanced view of the editor, virtually all of the metadata elements are accessible to the user.

Support for extensions to a metadata standard can also be achieved relatively quick by an experienced XML/XSL software engineer.

5.6 Transition between metadata standards

With the ISO19115:2003 Metadata standard for Geographic Information now being the preferred common standard, many have a need to migrate legacy metadata into the new standard.

GeoNetwork provides import (and export) functionality and has a number of transformers in place. It is an easy process for a system administrator to install custom transformers based on XSLT.

6. Installing the software

6.1 New version - New functionalities

The new GeoNetwork opensource comes with substantial upgrades of different components for a more intuitive and responsive user-system interaction. Web 2.0 technologies have been adopted, in particular AJAX techniques, to allow for more interactive and faster services in the web interface and for the integration of the existing web map viewer in the home page. Similar functionalities have been implemented in the administrative part of the system, to provide an easier access to the configuration pages related to site settings, catalogue harvesting, scheduling and maintenance.

The search interface has been completely overhauled to provide highly interactive searching capabilities. Furthermore, the new version of GeoNetwork embeds GeoServer as map server. Users can now not only overlay OGC web map services available on the web, but also create their own map services for other users to browse without having to download additional plugins. Maps created with web map services can be now saved as PDF and sent to others.

The metadata catalogue handles the latest ISO19115:2003 geographic metadata format based on the ISO19139:2007 schemas, as well as the older ISO19115 final draft format, FGDC and Dublin Core. The metadata editor is able to handle the majority of these complex standards, providing default, advanced and XML editing online tools.

The new version has a number of different harvesting interfaces allowing users to connect their own server to many other catalogues around the world. This is the result of the implementation of the open source reference for the web catalogue services according to OGC specifications. Harvesting in the new version is fully compatible with GeoNetwork 2.0 and higher nodes.

We have added advanced online and offline administration functionalities to configure, backup and migrate the application. We have also added a convenient import and export format "MEF" or Metadata Exchange Format, that allows the users to move metadata, previews and even data in a convenient single file. GeoNetwork can be easily expanded with plugins to export/import metadata to/from other software supporting MEF.



Figure 6.1. Standard home page of GeoNetwork opensource

6.2 Where do I get the installer?

You can find the software on the Internet at the GeoNetwork opensource Community website¹. The software is also distributed through the SourceForge.net Website at <http://sourceforge.net/projects/geonetwork>.

Use the platform independent installer (.jar) if you need anything more than a plain Windows installation.

6.3 System requirements

GeoNetwork can run either on **MS Windows**, **Linux** or **Mac OS X**.

Some general system requirements for the software to run without problems are listed below:

Processor: 2 GHz or higher

Memory (RAM): 1 GB or higher

Disk Space: 200 MB minimum. However, it is suggested to have a minimum of 1 GB of free disk space. Additional space is required depending on the amount of spatial data that you expect to upload into the internal geodatabase.

Other Software requirements: A Java Runtime Environment (JRE 1.5.0). For server installations, Apache Tomcat and a dedicated JDBC compliant DBMS (MySQL, Postgresql, Oracle) can be used instead of Jetty and McKoiDB respectively.

Additional Software

The software listed here is not required to run GeoNetwork, but can be used for custom installations.

1. MySQL DBMS v5.5+ (All)²
2. Postgresql DBMS v7+ (All)²
3. Apache Tomcat v5.5+ (All)²
4. Druid v3.8 (All)² to inspect the database

Supported browsers

GeoNetwork should work normally with the following browsers:

1. Firefox v2+ (All)²
2. Internet Explorer v6+ (Windows)
3. Safari v3+ (Mac OS X Leopard)²

¹ <http://geonetwork-opensource.org>

6.4 How do I install GeoNetwork opensource?

Before running the GeoNetwork installer, make sure that all system requirements are satisfied, and in particular that the Java Runtime Environment version 1.5.0 is set up on your machine.

On Windows

If you use Windows, the following steps will guide you to complete the installation (other FOSS will follow):

1. Double click on **geonetwork-install-2.4.0.exe** to start the GeoNetwork opensource desktop installer
2. Follow the instructions on screen (Figure 6.2, "Installer"). You can choose to install sample data, install the embedded map server (based on GeoServer³ and the CSW 2.0.2 test client. Developers may be interested in installing the source code and installer building tools. Full source code can be found in the GeoNetwork SubVersion code repository.
3. After completion of the installation process, a 'GeoNetwork desktop' menu will be added to your Windows Start menu under 'Programs'
4. Click Start > Programs > GeoNetwork desktop > Start server to start the GeoNetwork opensource Web server. The first time you do this, the system will require about 1 minute to complete startup.
5. Click Start > Programs > GeoNetwork desktop > Open GeoNetwork opensource to start using GeoNetwork opensource, or connect your Web browser to <http://localhost:8080/geonetwork/>

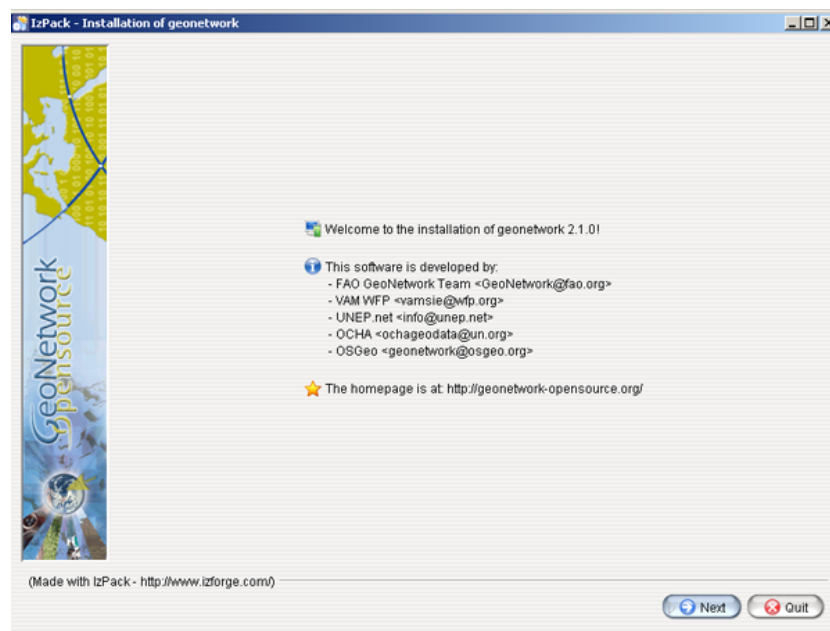


Figure 6.2. Installer

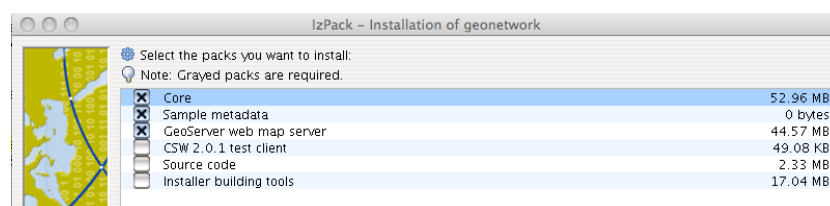


Figure 6.3. Packages to be installed

Installation using the platform independent installer

If you downloaded the platform independent installer (a .jar file), you can in most cases start the installer by simply double clicking on it.

Follow the instructions on screen (see also the section called “On Windows”).

At the end of the installation process you can choose to save the installation script (Figure 6.4, “Save the installation script for command line installations”).

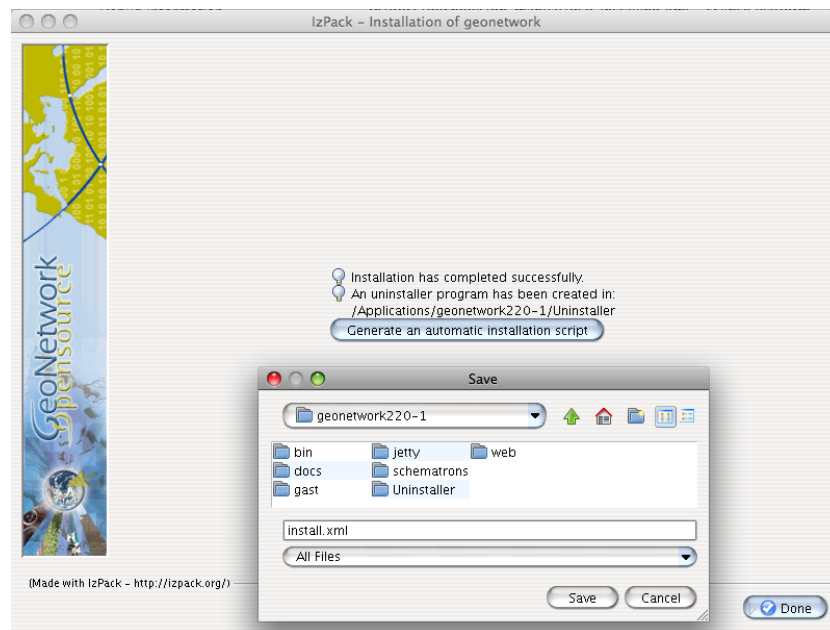


Figure 6.4. Save the installation script for command line installations

Command line installation

If you downloaded the platform independent installer (a .jar file), you can perform command line installations on computers without a graphical interface. You first need to generate an install script (see Figure 6.4, “Save the installation script for command line installations”). This install script can be edited in a text editor to change some installation parameters.

To run the installation from the command line, issue the following command in a terminal window and hit enter to start:

```
java -jar geonetwork-install-2.4.0-0.jar install.xml
[ Starting automated installation ]
[ Starting to unpack ]
[ Processing package: Core (1/3) ]
[ Processing package: Sample metadata (2/3) ]
[ Processing package: GeoServer web map server (3/3) ]
[ Unpacking finished ]
[ Writing the uninstaller data ... ]
[ Automated installation done ]
```

You can also run the installation with lots of debug output. To do so run the installer with the flag -DTRACE=true:

```
java -DTRACE=true -jar geonetwork-install-2.4.0-0.jar
```

Part II. Administrator Guide

This part explains how to configure and administer GeoNetwork based applications using its web and desktop tools. The tools are subdivided into 2 categories:

1. Web tools: these are directly accessible from the web interface and allow the user to tune almost all aspects of GeoNetwork.
 2. The GAST tool: this application provides special functions that cannot be executed while GeoNetwork is running. It also performs special tasks that are better done in a separate tool.
-

7. Basic configuration

7.1 System configuration

Most of the GeoNetwork system configuration parameters can be changed using the web interface. Those parameters that cannot be changed through the web interface can usually be changed using the GAST application.



Important

Configuration of these parameters is critically important for a proper functioning of the GeoNetwork catalogue in an operational context. Failing to properly change these settings may result in a system that does not function as expected. For example, downloads may fail to be correctly processed, or metadata harvesting from other servers may not work.

To get to the System configuration, you must be logged on as administrator first. Open the Administration page and select System configuration (Figure 7.1, “The link to the System configuration page”. The link is surrounded with a red rectangle).



Important

New installations of GeoNetwork use `admin` for both username and password. It is important to change this from the Administration page once you logged on!

ADMINISTRATION	
Metadata	
New metadata	Adds a new metadata into geonetwork copying it from a template
Metadata insert	Import metadata record in XML or MEF format
Batch Import	Import all XML formatted metadata from a local directory
Search for Unused	Search for unused or empty metadata
Transfer ownership	Transfer metadata ownership to another user
Manage thesauri	Add/modify/delete and show thesauri
Personal info	
Change password	Allow current user to change password
Change user information	Allow current user to change user information
Administration	
User management	Add/modify/delete and show users
Group management	Add/modify/delete and show groups
Category management	Add/modify/delete and show categories
Harvesting management	Add/modify/delete/start/stop harvesting tasks
System configuration	Allows to change some system's parameters
Localization	Allows to change localized entities, like groups, categories etc...
Index manager	Rebuild Lucene index <input type="button" value="Rebuild"/>
Test i18n	This service should help GeoNetwork opensource developers to have up to date localized files for the GUI.
CSW ISO test	Test interface for the CSW ISO catalog interface

Figure 7.1. The link to the System configuration page

Clicking the page’s link you will get the set of parameters that you can change (Figure 7.2, “The configuration options”). Here follows a detailed description of them:

SYSTEM CONFIGURATION	
SITE	
Name	<input type="text" value="GeoNetwork site"/>
Organization	<input type="text" value="My Organisation"/>
SERVER	
Host	<input type="text" value="somehost.org"/>
Port	<input type="text" value="80"/>
INTRANET	
Network	<input type="text" value="127.0.0.1"/>
Netmask	<input type="text" value="255.0.0.0"/>
Z39.50 SERVER	
Enable	<input checked="" type="checkbox"/>
Port	<input type="text" value="2100"/>
CSW	
Enable	<input checked="" type="checkbox"/>
Contact	<input type="text" value="admin (Jefferson D Bussey)"/>
Title	<input type="text" value="Chief"/>
Abstract	<input type="text" value="This catalog delivers info on the US Civil War"/>
Fees	<input type="text" value="None"/>
Access Constraints	<input type="text" value="None"/>
PROXY	
Use	<input type="checkbox"/>
FEEDBACK	
E-Mail	<input type="text" value="admin@somehost.org"/>
SMTP Host	<input type="text" value="smtp.somehost.org"/>
SMTP Port	<input type="text" value="25"/>
REMOVED METADATA	
Directory	<input type="text" value="../data/removed"/>
AUTHENTICATION	
Login uses:	
<input checked="" type="radio"/> GeoNetwork Database	
<input type="checkbox"/> Enable User Self-Registration	
<input type="radio"/> LDAP	
Allow other users to login from:	
<input type="checkbox"/> Shibboleth	

Figure 7.2. The configuration options

At the bottom of the page there are some buttons with the following purpose:

Back Simply returns to the main administration page. **Save** Saves the current options. If some options are invalid, the system will show a dialogue with the wrong parameter and will focus its text field on the page. Once the configuration is saved a success dialogue will be shown. **Refresh** This button simply refreshes the displayed options taking the new values from the server. This can be useful if some options get changed dynamically (for example by another user).

Public host and port usage

Up to now, the server's host and port are used in these cases:

1. During an editing session, when adding data links to a metadata. The host and port will be used to build download links to store inside the metadata.
2. During CSW requests. The GetCapabilities operation returns an XML document with HTTP links to the CSW services. These links are dynamically built using the host and port values.

Site General site parameters.

Name The name of the GeoNetwork's installation. This name will be used to identify the node in operations like the harvesting.

organisation The organisation the node belongs to. Just for informative reasons.

Server Here you have to enter the address of your GeoNetwork's node. This address is important because it will be used to access the node.

Host The node's address or IP number. If your node is publicly accessible from the Internet, you have to use the machine's domain/address. If your node is hidden into your private network and you have a firewall or web server that redirects incoming calls to the node, you have to enter the public address of the firewall or web server. A typical configuration is to have an Apache web server on address A that is publicly accessible and redirects the requests to a Tomcat server on a private address B. In this case you have to enter A in the host parameter.

port The node's port (usually 80 or 8080). If the node is hidden, you have to enter the port on the public firewall or web server.

Intranet A common need for an organisation is to discriminate between internal anonymous users (users that access the node from within the organisation) and external ones (users from the Internet). Node's administrators can specify different privileges for internal and external anonymous users and, in order to do so, they have to specify the parameters of the internal network.

Network The internal network's address in IP form.

Netmask The network's mask.

Catalogue services (CSW, Z39.50)

OGC CSW configuration

When using Open Geospatial Catalogue Service for the Web (OGC-CSW) service, a client will ask for a description of the service. This description, provided in the form of a GetCapabilities document, describes the main service's properties. The administration section allows configuration of the following CSW properties:

Enable: This option allows you to start or stop the CSW services. If this option is disabled, other catalogues cannot connect to the node using CSW protocol.

Contact: The main contact who is defined in the GetCapabilities document of the CSW service. This contact is one user of the catalogue.

Title: The title of your CSW service.

Abstract: The abstract of your CSW service.

Fees

Access constraints

The service description also contains the main keywords of the catalogue. The list of keywords is generated by the catalogue based on metadata content.

Inserted metadata is public: If this setting is selected the metadata inserted using CSW transaction is viewable for all users. This setting is disabled by default.

Z39.50 configuration

Z39.50: GeoNetwork can act as a Z39.50 server, which is an OGC communication protocol to query and retrieve metadata.

Enable: Check this option to start the Z39.50 submodule. Please, notice that GeoNetwork must be restarted in order to make this change active.

port: This is the port on which GeoNetwork will be listening for incoming Z39.50 requests. Usually, the value of 2100 is a standard one, but to have multiple GeoNetwork nodes on the same machine you have to change this value in order to avoid port conflicts between the different nodes.

Proxy configuration

Proxy: In some occasions (like harvesting) GeoNetwork must be able to connect to remote sites and this may be denied if an organisation uses proxy servers. In this cases, GeoNetwork must be configured to use the proxy server in order to route outgoing requests.



The screenshot shows a configuration panel titled 'PROXY'. It contains a 'Use' checkbox which is checked. Below this are four input fields: 'Host', 'Port', 'Username', and 'Password', each with a corresponding text box.

Figure 7.3. The proxy configuration options

Host: The proxy's name or address to use (usually an IP address).

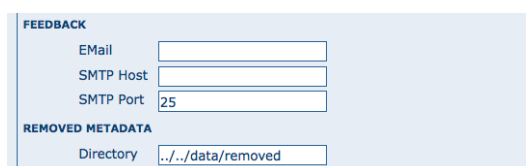
Port: The proxy's port to use.

Username (optional): a username should be provided if the proxy server requires authentication.

Password (optional): a password should be provided if the proxy server requires authentication.

Email & notification

Feedback GeoNetwork can sometimes send email, for example if a metadata is downloaded or if a user provides feedback using the online form. You have to configure the mail server GeoNetwork should use in order to enable it to send email.



The screenshot shows two configuration sections. The first is 'FEEDBACK' with three input fields: 'EMail', 'SMTP Host', and 'SMTP Port' (which has the value '25'). The second section is 'REMOVED METADATA' with one input field: 'Directory' (which has the value '../../data/removed').

Figure 7.4. The mail server configuration options

Email: This is the email address that will be used to send the email (the From address).

SMTP host: the mail server address to use when sending email.

SMTP port: the mail server SMTP port (usually 25).

Removed metadata

Defines the directory used to store a backup of metadata and data after a delete action. This directory is used as a backup directory to allow system administrators to recover metadata and possibly related data after erroneous deletion. By default the `removed` directory is created under the data folder

Authentication

In this section you define the source against which GeoNetwork will authenticate users and passwords.

AUTHENTICATION

Login uses:

GeoNetwork Database
Enable User Self-Registration

LDAP

Allow other users to login from:

Shibboleth

Figure 7.5. Authentication configuration options

By default, users are authenticated against info held in the GeoNetwork database. When the GeoNetwork database is used as the authentication source, the user self-registration function can be enabled.

You may choose to authenticate logins against either the GeoNetwork database tables or LDAP (the lightweight directory access protocol) but not both. The next section describes how to authenticate against LDAP.

In addition to either of these options, you may also configure other authentication sources. At present, Shibboleth is the only additional authentication source that can be configured. Shibboleth is typically used for national access federations such as the Australian Access Federation. Configuring shibboleth authentication in GeoNetwork to use such a federation would allow not only users from a local database or LDAP directory to use your installation, but any user from such a federation.

LDAP Authentication

The section defines how to connect to an LDAP authentication system.

AUTHENTICATION

Login uses:

GeoNetwork Database

LDAP

Host

Port

Default profile

Distinguished names

Base

Users

User's attributes

Name

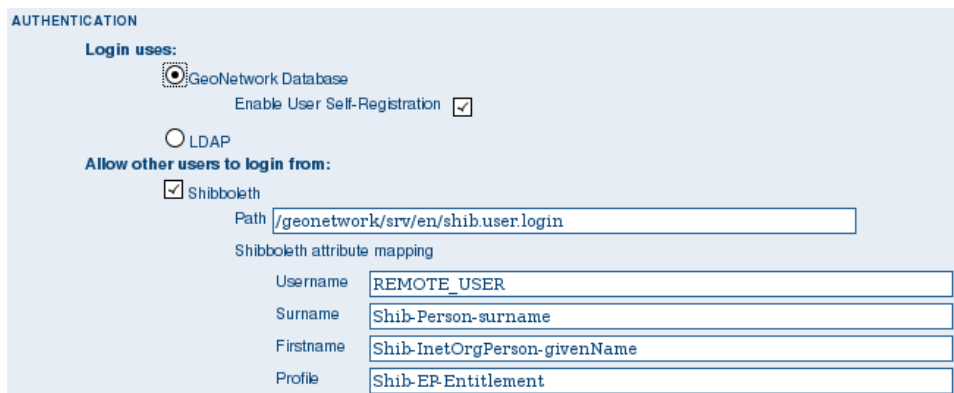
Profile

Figure 7.6. The LDAP configuration options

Typically all users must have their details in the LDAP directory to login to GeoNetwork. However if a user is added to the GeoNetwork database with the Administrator profile then they will be able to login without their details being present in the LDAP directory.

Shibboleth Authentication

When using either the GeoNetwork database or LDAP for authentication, you can also configure shibboleth to allow authentication against access federations.



The screenshot shows a configuration panel titled "AUTHENTICATION". Under "Login uses:", there are two radio buttons: "GeoNetwork Database" (selected) and "LDAP". Below this is a checkbox for "Enable User Self-Registration" which is checked. Under "Allow other users to login from:", there is a checked checkbox for "Shibboleth". Below this, there is a text input field for "Path" containing the value "/geonetwork/srv/en/shib.user.login". Under "Shibboleth attribute mapping", there are four text input fields: "Username" with "REMOTE_USER", "Surname" with "Shib-Person-surname", "Firstname" with "Shib-InetOrgPerson-givenName", and "Profile" with "Shib-EP-Entitlement".

Figure 7.7. The Shibboleth configuration options

Shibboleth authentication requires interaction with Apache web server. In particular, the apache web server must be configured to require Shibboleth authentication to access the path entered in the configuration. The apache web server configuration will contain the details of the shibboleth server that works out where a user is located (sometimes called a 'where are you from' server).

The remainder of the shibboleth login configuration describes how shibboleth authentication attributes are mapped to GeoNetwork user database fields as once a user is authenticated against shibboleth, their details are copied to the local GeoNetwork database.

8. User and Group Administration

GeoNetwork uses the concept of Users, Groups and User Profiles. A User can be part of several Groups. A User also has a User Profile¹. The combination of User Profile and Group defines what tasks the User can perform on the system or on specific metadata records.

8.1 Creating new user Groups

The administrator can create new groups of users. User groups can correspond to logical units within an organisation. For example groups for Fisheries, Agriculture, Land and Water, Health etcetera.

To create new groups you should be logged on with an account that has administrative privileges. To log in, simply go to the home page and enter your username and password in the top right corner fields, then click on the login button (Figure 8.1, "Login form").



Important

New installations of GeoNetwork use `admin` for both username and password. It is important to change this from the Administration page once you logged on!

English Français Español 中文		
Username	Password	Login

Figure 8.1. Login form

- Select the `Administration` button in the menu. On the Administration page, select `Group management` (Figure 8.2, "Administration page").

ADMINISTRATION	
Metadata	
New metadata	Adds a new metadata into geonetwork copying it from a template
Metadata insert	Import metadata record in XML or MEF format
Batch Import	Import all XML formatted metadata from a local directory
Search for Unused	Search for unused or empty metadata
Transfer ownership	Transfer metadata ownership to another user
Manage thesauri	Add/modify/delete and show thesauri
Personal info	
Change password	Allow current user to change password
Change user information	Allow current user to change user information
Administration	
User management	Add/modify/delete and show users
Group management	Add/modify/delete and show groups
Category management	Add/modify/delete and show categories
Harvesting management	Add/modify/delete/start/stop harvesting tasks
System configuration	Allows to change some system's parameters
Localization	Allows to change localized entities, like groups, categories etc...
Index manager	Rebuild Lucene index <input type="button" value="Rebuild"/>
Test i18n	This service should help GeoNetwork opensource developers to have up to date localized files for the GUI.
CSW ISO test	Test interface for the CSW ISO catalog interface

Figure 8.2. Administration page

1. Select **Add a new group**²;

¹A User can only have one User Profile associated.

GROUP MANAGEMENT

Important! This form allows you to add a Key value to the database. The key can not have spaces. After adding your key values, use the "Localization" form in the Administration panel to provide the actual name you want to be displayed on the website for the different languages.

Name	Description	Operation
sample		<input type="button" value="Edit"/> <input type="button" value="Delete"/>

Figure 8.3. Group management

2. Fill out the details. The email address will be used to send feedback on data downloads when they occur for resources that are part of the Group.

ADD A NEW GROUP

Important! This form allows you to add a Key value to the database. The key can not have spaces. After adding your key values, use the "Localization" form in the Administration panel to provide the actual name you want to be displayed on the website for the different languages.

Name:

Description:

Download Email:

Figure 8.4. Group edit form

3. Click on **Save**



Important

The Name should *NOT* contain spaces! You can use the Localisation functions to provide localised names for groups.

Access privileges can be set per metadata record. You can define privileges on a per Group basis. Privileges that can be set relate to visibility of the Metadata (*Publish*), data Download, Interactive Map access and display of the record in the Featured section of the home page.

Editing defines the groups for which editors can edit the metadata record.

Notify defines what groups are notified when a file managed by GeoNetwork is downloaded.

Below is an example of the privileges management table related to a dataset (Figure 8.5, "Privilege settings").

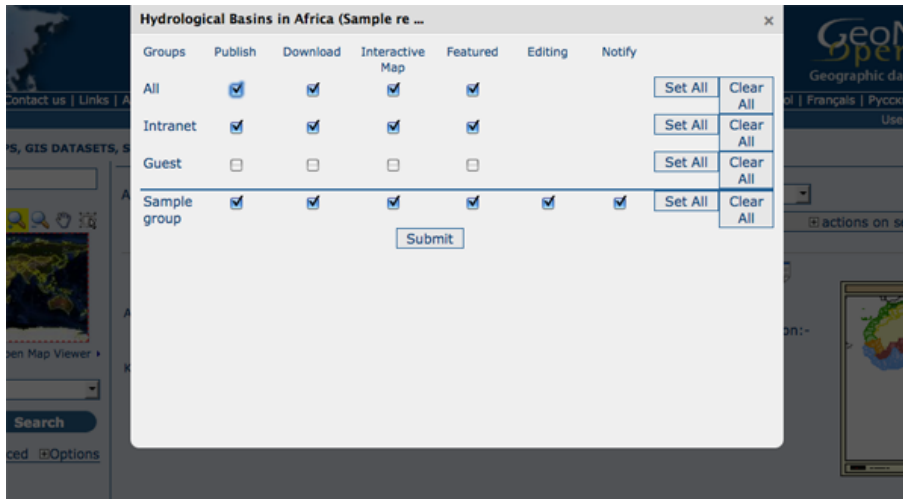


Figure 8.5. Privilege settings

8.2 Creating new Users

To add a new user to the GeoNetwork system you do the following:

1. Select **User Management** from the Administration link in the toolbar (Figure 8.2, “Administration page”);
2. Click the button **Add a new user** (Figure 8.6, “User administration form”);

USER MANAGEMENT				
Username	Last Name	First Name	Profile	Operation
admin	admin	admin	Administrator	<input type="button" value="Edit"/>
editor	Guest	Editor	Editor	<input type="button" value="Edit"/> <input type="button" value="Delete"/>
reviewer	Guest	Reviewer	Reviewer	<input type="button" value="Edit"/> <input type="button" value="Delete"/>
user	Guest	User	RegisteredUser	<input type="button" value="Edit"/> <input type="button" value="Delete"/>
useradmin	Guest	Useradmin	UserAdmin	<input type="button" value="Edit"/> <input type="button" value="Delete"/>

Figure 8.6. User administration form

3. Provide the **information** required for the new user (Figure 8.7, “User information form”);

The image shows a web form titled "INSERT" for adding a new user. It includes the following fields and controls:

- Username (*): Text input field
- Password (*): Text input field
- Confirm password (*): Text input field
- Last Name: Text input field
- First Name: Text input field
- Address: Text input field
- State: Text input field
- Country: Dropdown menu
- Email: Text input field
- Organisation / department: Text input field
- Kind: Dropdown menu (currently set to "Government")
- Profile: Dropdown menu (currently set to "Editor", with a list of options: Administrator, User administrator, Content reviewer, Editor, Registered user)
- Groups: List box (empty)

At the bottom of the form are two buttons: "Back" and "Save".

Figure 8.7. User information form

4. Assign the correct **profile** (Section 8.3, "User Profiles");
5. Assign the user to a **group**;
6. Click on **Save**.

8.3 User Profiles

Users can have different profiles depending on their role in the GeoNetwork system. A profile defines what tasks the user can perform.

User profiles are hierarchical and based on inheritance. This means that a user with an Editor profile can create and modify new metadata records, but can also use all functions a Registered user can use.

Rights associated with the profiles are illustrated in detail in the list below:

1. Administrator Profile

The Administrator has special privileges that give access to all available functions. These include:

- Full rights for creating new groups and new users
- Rights to change users/groups' profiles
- Full rights for creating/editing/deleting new/old metadata
- Perform system administration and configuration tasks.

2. User Administrator Profile

The User Administrator is the administrator of his/her own group with the following privileges:

- Full rights on creating new users within the own group
- Rights to change users profiles within the own group

- Full rights on creating/editing/ deleting new/old data within the own group

3. Content Reviewer Profile

The content reviewer is the only person allowed to give final clearance on the metadata publication on the Intranet and/or on the Internet:

- Rights on reviewing metadata content within the own group and authorising its publication

4. Editor Profile

The editor works on metadata with following privileges:

- Full rights on creating/editing/ deleting new/old data within the own group

5. Registered User Profile

The Registered User has more access privileges than non-authenticated Guest users:

- Right to download protected data

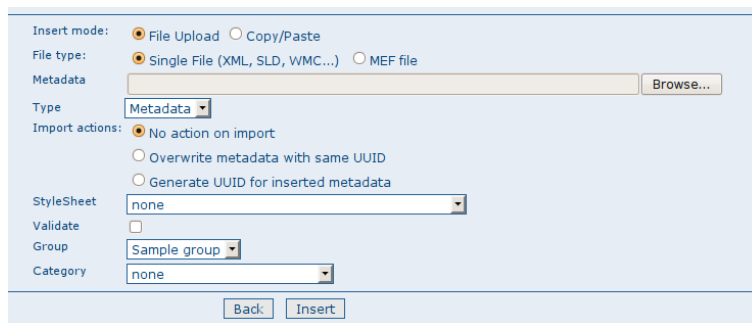
9. Import facilities

9.1 File import

The file import facility allows you to import metadata records in XML format or MEF format or copy/paste XML metadata to insert. In order to use this facility, you have to be logged in as an administrator. After the login step, go to the administration page and select the Metadata insert link.

Clicking the link, you will reach the metadata import file's page as illustrated in Figure 9.1, "The XML file import options". You have to specify a set of parameters to make the import working. Depending on the insert type you will perform:

XML file import



The screenshot shows a web form for XML file import. It includes the following fields and options:

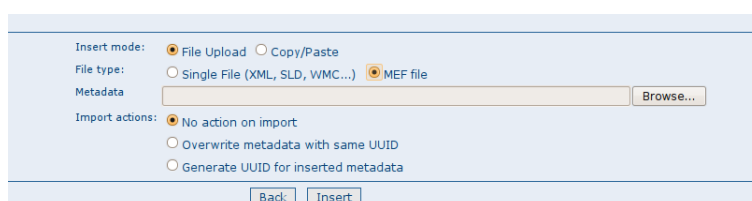
- Insert mode:** Radio buttons for File Upload and Copy/Paste.
- File type:** Radio buttons for Single File (XML, SLD, WMC...) and MEF file.
- Metadata:** A text input field with a "Browse..." button.
- Type:** A dropdown menu set to "Metadata".
- Import actions:** Radio buttons for No action on import, Overwrite metadata with same UUID, and Generate UUID for inserted metadata.
- StyleSheet:** A dropdown menu set to "none".
- Validate:** An unchecked checkbox.
- Group:** A dropdown menu set to "Sample group".
- Category:** A dropdown menu set to "none".
- Buttons:** "Back" and "Insert" buttons at the bottom.

Figure 9.1. The XML file import options

For both, XML and MEF file import, you could choose actions to be performed at import. Import actions options are: **No action on import** (nothing done during import process), **Overwrite metadata with same UUID** (this option will delete existing metadata with the same UUID, and add the new one), **Generate UUID for inserted metadata** (this option will force a new UUID for each metadata inserted).

Stylesheet option: This is a powerful option because allows you to specify a stylesheet for an XSL transformation. The drop down control is filled with files taken from the web/xsl/conversion/import folder: all XSL files you put there will be made available. This is a dynamic process so you don't have to restart GeoNetwork. The purpose of this option is to allow the conversion of a metadata into a suitable format that is supported by GeoNetwork. **Validate option:** This is a simple validation step that you can choose to perform. The metadata is validated against its schema. **Group option:** You have to select a group to associate to the imported metadata. Usually the group is the creator of the metadata set. **Category option:** You can specify one category to associate to your metadata in order to simplify the search.

MEF file import



The screenshot shows a web form for MEF file import. It includes the following fields and options:

- Insert mode:** Radio buttons for File Upload and Copy/Paste.
- File type:** Radio buttons for Single File (XML, SLD, WMC...) and MEF file.
- Metadata:** A text input field with a "Browse..." button.
- Import actions:** Radio buttons for No action on import, Overwrite metadata with same UUID, and Generate UUID for inserted metadata.
- Buttons:** "Back" and "Insert" buttons at the bottom.

Figure 9.2. The MEF file import options

As presented before, only actions options could be selected during a MEF import.

XML metadata insert

Figure 9.3. The XML insert options

Choosing the radio button Copy/Paste, you could perform an XML insert, directly copy/pasting your metadata record. Please refer to XML import for options you could use. Note that it is not possible to perform actions during import with this method.

9.2 Batch import

The batch import facility allows you to import a set of metadata into the system all at once. In order to use this facility, you have to be logged in as an administrator. After the login step, go to the administration page and select the batch import's link (Figure 9.4, "How to reach the batch import page". The link is surrounded with a red rectangle).

ADMINISTRATION	
Metadata	
New metadata	Adds a new metadata into geonetwork copying it from a template
Metadata insert	Import metadata record in XML or MEF format
Batch import	Import all XML formatted metadata from a local directory
Search for Unused	Search for unused or empty metadata
Transfer ownership	Transfer metadata ownership to another user
Manage thesauri	Add/modify/delete and show thesauri
Personal info	
Change password	Allow current user to change password
Change user information	Allow current user to change user information
Administration	
User management	Add/modify/delete and show users
Group management	Add/modify/delete and show groups
Category management	Add/modify/delete and show categories
Harvesting management	Add/modify/delete/start/stop harvesting tasks
System configuration	Allows to change some system's parameters
Localization	Allows to change localized entities, like groups, categories etc...
Index manager	Rebuild Lucene index
	<input type="button" value="Rebuild"/>
Test I18n	This service should help GeoNetwork opensource developers to have up to date localized files for the GUI.
CSW ISO test	Test interface for the CSW ISO catalog interface

Figure 9.4. How to reach the batch import page

Clicking the link, you will reach the batch import's page as illustrated in Figure 9.5, "The batch import options". You have to specify a set of parameters to make the import working. They are:

Directory This is the full path on the server's file system of the directory to scan. GeoNetwork will look for and try to import all XML files present into this directory. It is important to notice that this is the directory on the *server* machine and *not* on the client of the user that is doing the import. All metadata files present into the import directory *must* have the same schema format. Schema GeoNetwork supports only some metadata formats so you have to specify the schema of the metadata you want to import. If a metadata does not belong to the selected schema, the entire operation will be aborted. **Validate** This is a simple validation step that you can choose to perform. The metadata is validated against its schema. **Group** You have to select a group to associate to the imported metadata. Usually the group is the creator of the metadata set. **Category** You can specify one category to associate to your metadata in order to simplify the search. **Stylesheet** This is a powerful option because allows you to specify a stylesheet for an XSL transformation. The drop down control is filled with files taken from the `web/xsl/conversion/import` folder: all XSL files you put there will be made available. This is a dynamic process so you don't have to restart GeoNetwork. The purpose of this option is to allow the conversion of a metadata into a suitable format that is supported by GeoNetwork. Therefore, it is important that the result of the transformation matches the schema format selected above.

Below the page, there are the following buttons:

Back Goes back to the administration form. **Upload** Starts the import process. When the process ends, the total count of imported metadata will be shown. Please notice that the import is transactional: the metadata set will be fully imported or fully discarded (there are no partial imports). Files that starts with '.' or that do not end with '.xml' are ignored.

Figure 9.5. The batch import options

Structured import

An hidden feature of the batch import is the possibility to specify some import parameters in more detail. This feature is triggered when the specified folder contains the `import-config.xml` file. When this happen, this file is read and the standard import switches to the structured one.

The `import-config.xml` file has a config root element with the following children:

1. `categoryMapping [1]`: this element specifies the mapping of directories to categories.
 - a. `mapping [0..n]`: This element can appear 0 or more times and maps one directory name to a category name. It must have a `dir` attribute that indicates the directory and a `to` attribute that indicates the category name.
 - b. `default [1]`: This element specifies a default mapping of categories for all directories that do not match the other mapping elements. It must have only the `to` attribute.

2. schemaMapping [1]: this element specifies the mapping of directories to metadata schemas.
 - a. mapping [0..n]: This element can appear 0 or more times and maps one directory to the schema name that must be used when importing. The provided schema must match the one used by the metadata contained into the specified directory, which must all have the same schema. It must have a dir attribute that indicates the directory and a to attribute that indicates the schema name.
 - b. default [1]: default behaviour to use when all other mapping elements do not match. It must have only the to attribute.

Here is an example of the `import-config.xml` file:

```
<config>
  <categoryMapping>
    <mapping dir="1" to="maps" />
    <mapping dir="3" to="datasets" />
    <mapping dir="6" to="interactiveResources" />
    <mapping dir="30" to="photo" />
    <default to="maps" />
  </categoryMapping>
  <schemaMapping>
    <mapping dir="3" to="fgdc-std" />
    <default to="dublin-core" />
  </schemaMapping>
</config>
```

The import procedure starts by scanning the provided directory. This can contain, beside the `import-config.xml` file, only subdirectories which name will be ignored but used only as a container. Inside each directory, there is another level made only by directories that represent a metadata grouping for categories. Each directory name will be used as the `dir` attribute in the mapping scheme previously described.

10. Harvesting

10.1 Introduction

Since the beginning of the project, there has been the need to share metadata among several GeoNetwork nodes. Usually, each node takes care of a region of interest so it is important to be able to perform a search over all these nodes at the same time. This is called distributed search and exploits the Internet connectivity. In our cases, this distributed search can be heavy to perform if there are many maps with associated thumbnails. Furthermore, GeoNetwork is usually employed in areas (like Africa, Asia) where the connectivity can be limited, making the use of distributed search not feasible.

Harvesting is the process of collecting remote metadata and storing them locally for a faster access. This is a periodic process to do, for example, once a week. Harvesting is not a simple import: local and remote metadata are kept aligned. Using some *magic*, one GeoNetwork node is capable of discovering metadata that have been added, removed or updated in the remote node.

GeoNetwork is able to harvest from the following sources (for more details see below):

1. Another GeoNetwork node (version 2.1 or above).
2. An old GeoNetwork 2.0 node.
3. A *WebDAV* server.
4. A *CSW 2.0.1* or *2.0.2* catalogue server.
5. An *OAI-PMH* server.
6. An *OGC* service using its *GetCapabilities* document. These include *WMS*, *WFS*, *WPS* and *WCS* services.

10.2 Mechanism overview

The harvesting mechanism is based on the concept of a *universally unique identifier (UUID)*. This is a special id because it is not only unique locally to the node that generated it but it is unique across all the world. It is a combination of the network interface's *MAC address*, the current date/time and a random number. Every time you create a new metadata in GeoNetwork, a new UUID is generated and assigned to it.

Another important concept behind the harvesting is the *last change date*. Every time you change a metadata, its last change date is updated. Just storing this parameter and comparing it with a new one allows any system to find out if the metadata has been modified since last update.

These two concepts allow GeoNetwork to fetch a remote metadata, check if it has been updated and remove it locally if it has been removed remotely. Furthermore, thanks to UUIDs, a hierarchy of harvesting nodes can be built where B harvests from C and A harvests from B. Even loops can be created because harvested metadata cannot be modified.

10.3 Harvesting life cycle

When a harvesting node is set, there is no harvested metadata. During the first run, all remote matching metadata are retrieved and stored locally. After the first run, only changed metadata are retrieved. Harvested metadata are not editable for the following reasons:

1. The harvesting is periodic so any local change to harvested metadata will be lost during the next run.
2. The change date is used to keep track of changes so if it gets changed outside the originator site, the harvesting mechanism is compromised.

Beside the metadata itself, this implies that users cannot change all other metadata properties (like categories, privileges etc...).

The harvesting process goes on until one of the following situations arises:

1. An administrator stops (deactivates) the node.
2. An exception arises. In this case the node is automatically stopped.

When a harvesting node is removed, all harvest metadata are removed too.

10.4 Multiple harvesting and hierarchies

Catalogues that provide UUIDs for metadata (for example GeoNetwork and a CSW server) can be harvested several times without having to take care about metadata overlap. This allows the possibility to perform a thematic search and a metadata belonging to multiple searches is harvested only once and not duplicated.

This mechanism allows the GeoNetwork harvesting type to be combined with other GeoNetwork nodes to perform hierarchical harvesting. This way a metadata can be harvested from several nodes. For example, consider this scenario:

1. Node (A) has created metadata (a)
2. Node (B) harvests (a) from (A)
3. Node (C) harvests (a) from (B)
4. Node (D) harvests from both (A), (B) and (C)

In this scenario, Node (D) will get the same metadata (a) from all 3 nodes (A), (B), (C). The metadata will flow to (D) following 3 different paths but thanks to its UUID only one copy will be stored. When (a) will be changed in (A), a new version will flow to (D) but, thanks to the change date, the copy in (D) will be updated with the most recent version.

10.5 General notes and issues

General

1. The harvesting engine does not store harvesting results. This implies that if the server is restarted the last results are lost.
2. Changes to the harvesting parameters (for example privileges and categories) are taken into account in the next harvesting run.

GeoNetwork harvesting type

1. During harvesting, site icons are harvested and local copies updated. Icons are propagated to new nodes as soon as these nodes harvest from this one.
2. The metadata UUID is taken from the info.xml file of the MEF bundle. Any UUID stored inside the metadata will be overwritten with this one.

WebDAV harvesting type

1. The same metadata could be harvested several times by different harvesting nodes. Anyway, this is not a good practise because every copy of the metadata will have a different UUID and the system will fill with different copies of the same metadata.

CSW harvesting type

1. If the dct:modified element is missing from the GetRecords response the metadata will be always harvested.
2. Any exception during getRecordById operation is discarded and the metadata skipped.

OAI-PMH harvesting type

1. The id of the remote server must be a UUID. If not, metadata can be harvested but during hierarchical propagation id clashes could corrupt harvested metadata.
2. During harvesting, GeoNetwork will try to auto detect the schema of each metadata. If the schema is not supported the metadata is skipped.

OGC service harvesting type

1. Every time the harvester runs, it will remove previously harvested information and create new ones. GeoNetwork will generate the id for all metadata (both service and datasets). Therefore, for datasets, if the metadata is created using a remote XML document (ie. if a MetadataUrl tag is in the GetCapability document), the UUID of the document is used.
2. Thumbnails are generated only for Web Map Service (WMS). The service should also support the WGS84 projection

10.6 The main page

To access the harvesting main page you have to be logged in as an administrator. From the administration page, click the link shown in Figure 10.1, “How to access the harvesting main page” with a red rectangle.

ADMINISTRATION	
Metadata	
New metadata	Adds a new metadata into geonetwork copying it from a template
Metadata insert	Import metadata record in XML or MEF format
Batch import	Import all XML formatted metadata from a local directory
Search for Unused	Search for unused or empty metadata
Transfer ownership	Transfer metadata ownership to another user
Manage thesauri	Add/modify/delete and show thesauri
Personal info	
Change password	Allow current user to change password
Change user information	Allow current user to change user information
Administration	
User management	Add/modify/delete and show users
Group management	Add/modify/delete and show groups
Category management	Add/modify/delete and show categories
Harvesting management	Add/modify/delete/start/stop harvesting tasks
System configuration	Allows to change some system's parameters
Localization	Allows to change localized entities, like groups, categories etc...
Index manager	Rebuild Lucene index
	<input type="button" value="Rebuild"/>
Test I18n	This service should help GeoNetwork opensource developers to have up to date localized files for the GUI.
CSW ISO test	Test interface for the CSW ISO catalog interface

Figure 10.1. How to access the harvesting main page

Figure 10.2, “The harvesting main page” shows the harvesting main page. The page shows a list of harvesting nodes that have been created so far. On the bottom side there is a set of buttons to manage these nodes. The meaning of each column is as follows:

1. *Select* This is just a check box to select one or more nodes. The selected nodes will be affected by the first row of buttons (start, stop, run, remove). For example, if you select 3 nodes and press the Remove button, these 3 nodes will be removed.
2. *Name* This is the node’s name provided by the administrator.
3. *Type* The node’s harvesting type chosen when the node was created (GeoNetwork, web folder etc...).
4. *Status* This is an icon that reflects the node’s current status. See Table 10.1, “Possible status icons” for all different icons and status description.
5. *Errors* This column reflects the status of the last harvesting run, which could have succeeded or not. The result is reflected on this icon and a tool tip will show detailed information. See Table 10.2, “Possible error icons” for all different icons.
6. *Every* The time (in days, hours, minutes) between two consecutive harvesting from this node.
7. *Last run* The date, in ISO 8601 format, of the most recent harvesting run.
8. *Operation* A list of buttons for all possible operations on a node.
9. Selecting *Edit* will allow you to change the parameters for a node.

The screenshot shows a web interface titled "HARVESTING MANAGEMENT". It contains a table with the following columns: Select, Name, Type, Status, Errors, Every, Last run, and Operation. Below the table are two rows of buttons: the first row has Activate, Deactivate, Run, and Remove; the second row has Back, Add, and Refresh.

Select	Name	Type	Status	Errors	Every	Last run	Operation
<input type="checkbox"/>	Dav test	Web DAV			0:5:5		Edit
<input type="checkbox"/>	FAO site	GeoNetwork			0:1:30		Edit
<input type="checkbox"/>	CSW	CSW/2.0			0:1:30		Edit

Buttons: Activate, Deactivate, Run, Remove, Back, Add, Refresh

Figure 10.2. The harvesting main page

The bottom side of the page contains two rows of buttons. The first row contains buttons that can operate on a set of nodes. You can select the nodes using the check box on the first column and then press the proper button. When the button finishes its action, the check boxes are cleared. Here is the meaning of each button:

1. *Activate* When a new harvesting node is created, it’s status is *inactive*. Use this button to make it *active* and thus to start harvesting from the remote node.
2. *Deactivate* Stops harvesting from a node. Please notice that this does not mean that a currently running harvesting will be stopped but it means that this node will be ignored during future harvesting.
3. *Run* This button simply tells the harvesting engine to start harvesting immediately. This is useful for testing during the harvesting setup.
4. *Remove* Remove all currently selected nodes. A dialogue will ask the user to confirm the action.

The second row contains general purpose buttons. Here is the meaning of each button:

1. *Back* Simply returns to the main administration page.
2. *Add* This button allows to create new harvesting nodes.
3. *Refresh* Refreshes the current list of nodes querying the server. This can be useful to see if the harvesting list has been altered by someone else or if any harvesting process started.

Table 10.1. Possible status icons






Icon	Status	Description
	Inactive	The harvesting from this node is stopped.
	Active	The harvesting engine is waiting for the timeout specified for this node. When the timeout is reached, the harvesting starts.
	Running	The harvesting engine is currently running, fetching metadata from remote nodes. When the process will be finished, the status will be switched to active.

Table 10.2. Possible error icons

Icon	Description
	The harvesting was OK, no errors were found. In this case, a tool tip will show some harvesting results (like the number of harvested metadata etc...).
	The harvesting was aborted due to an unexpected condition. In this case, a tool tip will show some information about the error.

Harvesting result tips

If the harvesting succeeds, a tool tip will show detailed information about the harvesting process. This way you can check if the harvester worked as expected or if there is something to fix to the harvesting parameters or somewhere else. The result tip is like a table, where each row refers to

Total This is the total number of metadata found remotely. Metadata with the same id are considered as one. Added Number of metadata added to the system because they were not present locally. Removed Number of metadata that have been removed locally because they are not present in the remote server anymore. Updated Number of metadata that are present locally but that needed to be updated because their last change date was different from the remote one. Unchanged Local metadata left unchanged. Their remote last change date did not change. Unknown schema Number of skipped metadata because their format was not recognised by GeoNetwork. Unretrievable Number of metadata that were ready to be retrieved from the remote server but for some reason there was an exception during the data transfer process. Bad Format Number of skipped metadata because they did not have a valid XML representation. Does not validate Number of metadata which did not validate against their schema. These metadata were harvested with success but skipped due to the validation process. Usually, there is an option to force validation: if you want to harvest these metadata anyway, simply turn it off.

Table 10.3. Result information supported by harvesting types

Result vs harvesting type	GeoNetwork	WebDAV	CSW	OAI-PMH	OGC Service
Total	x	x	x	x	x
Added	x	x	x	x	x
Removed	x	x	x	x	x
Updated	x	x	x	x	
Unchanged	x	x	x	x	
Unknown schema	x	x	x	x	x
Unretrievable	x	x	x	x	x
Bad Format		x		x	
Does Not Validate		x		x	
Thumbnails / Thumbnails failed					x
Metadata URL attribute used					x

10.7 Adding new nodes

The Add button in the main page allows you to add new harvesting nodes. It will open the form shown in Figure 10.3, “Adding a new harvesting node”. When creating a new node, you have to choose the harvesting protocol supported by the remote server. The supported protocols are:

1. *GeoNetwork 2.1 remote node* This is the standard and most powerful harvesting protocol used in GeoNetwork. It is able to log in into the remote node, to perform a standard search using the common query fields and to import all matching metadata. Furthermore, the protocol will try to keep both remote privileges and categories of the harvested metadata if they exist locally. Please notice that since GeoNetwork 2.1 the harvesting protocol has been improved. This means that it is not possible to use this protocol to harvest from version 2.0 or below.
2. *Web DAV server* This harvesting type uses the web DAV (Distributed Authoring and Versioning) protocol to harvest metadata from a DAV server. It can be useful to users that want to publish their metadata through a web server that offers a DAV interface. The protocol allows to retrieve the contents of a web page (a list of files) with their change date.
3. *Catalogue Services for the Web 2.0* The Open Geospatial Consortium Catalogue Services for the Web and it is a search interface for catalogues developed by the Open Geospatial Consortium. GeoNetwork implements version 2.0 of this protocol.
4. *GeoNetwork v2.0 remote node* GeoNetwork 2.1 introduced a new powerful harvesting engine which is not compatible with GeoNetwork version 2.0 based catalogues. Old 2.0 servers can still harvest from 2.1 servers but harvesting metadata from a v2.0 server requires this harvesting type. This harvesting type is deprecated.

5. *Z3950 Remote search* Not implemented. This is a placeholder.
6. *OAI Protocol for Metadata Harvesting 2.0* This is a good harvesting protocol that is widely used among libraries. GeoNetwork implements version 2.0 of the protocol.

The drop down list shows all available protocols. Pressing the Add button you will reach an edit page whose content depends on the chosen protocol. The Back button will go back to the main page.



The screenshot shows a light blue header with the text "Harvesting management". Below the header is a form area with a dotted border. Inside the form, there is a label "Type" followed by a dropdown menu showing "Geonetwork remote node". Below the dropdown are two buttons: "Back" and "Add".

Figure 10.3. Adding a new harvesting node

Adding a GeoNetwork node

This type of harvesting allows you to connect to a GeoNetwork node, perform a simple search as in the main page and retrieve all matched metadata. The search is useful because it allows you to focus only on metadata of interest. Once you add a node of this type, you will get a page like the one shown in Figure 10.4, "Adding a GeoNetwork node". The meaning of the options is the following:

HARVESTING MANAGEMENT

SITE

Name

Host

Port

Servlet

Use account

Username

Password

SEARCH CRITERIA

Criteria

Free text

Title

Abstract

Keywords

Digital

Hardcopy

Source

OPTIONS

Every : : (days : hours : minutes)

One run only

PRIVILEGES

Remote group

all

sample

Copy policy

CATEGORIES

Figure 10.4. Adding a GeoNetwork node

Site Here you put information about the GeoNetwork's node you want to harvest from (host, port and servlet). If you want to search protected metadata you have to specify an account. The name parameter is just a short description that will be shown in the main page beside each node. Search In this section you can specify search parameters: they are the same present in the main page. Before doing that, it is important to remember that the GeoNetwork's harvesting can be hierarchical so a remote node can contain both its metadata and metadata harvested from other nodes and sources. At the beginning, the Source drop down is empty and you have to use the Retrieve sources button to fill it. The purpose of

this button is to query GeoNetwork about all sources which it is currently harvesting from. Once you get the drop down filled, you can choose a source name to constrain the search to that source only. Leaving the drop down blank, the search will spread over all metadata (harvested and not). You can add several search criteria for each site through the Add button: several searches will be performed and results merged. Each search box can be removed pressing the small button on the left of the site's name. If no search criteria is added, a global unconstrained search will be performed. Options This is just a container for general options.

Every This is the harvesting period. The smallest value is 1 minute while the greatest value is 100 days. One run only If this option is checked, the harvesting will do only one run after which it will become inactive. Privileges Here you decide how to map remote group's privileges. You can assign a copy policy to each group. The Intranet group is not considered because it does not make sense to copy its privileges. The All group has different policies from all the others:

1. Copy: Privileges are copied.
2. Copy to Intranet: Privileges are copied but to the Intranet group. This way public metadata can be made protected.
3. Don't copy: Privileges are not copied and harvested metadata will not be publicly visible.

For all other groups the policies are these:

1. Copy: Privileges are copied only if there is a local group with the same (not localised) name as the remote group.
2. Create and copy: Privileges are copied. If there is no local group with the same name as the remote group then it is created.
3. Don't copy: Privileges are not copied.

On the bottom side of the page there are some buttons:

Back Simply return to the main harvesting page. Save Saves the current node information and returns to the main harvesting page. When creating a new node, the node will be actually created only when you press this button.

Adding a Web DAV node

In this type of harvesting, metadata are retrieved from a remote web page. The available options are shown in Figure 10.5, "Adding a web DAV node" and have the following meaning:

HARVESTING MANAGEMENT

SITE

Name

URL

Icon **REMOTE**

Use account

Username

Password

OPTIONS

Every : : (days : hours : minutes)

One run only

Validate

Recurse

PRIVILEGES

Groups

Group View Interactive map Featured

All Remove

CATEGORIES

Figure 10.5. Adding a web DAV node

Site Here are the connection information. The available options are:

Name This is a short description of the node. It will be shown in the harvesting main page. **URL** The remote URL from which metadata will be harvested. Each file found that ends with .xml will indicate a metadata and will be retrieved, converted into XML and imported. **Icon** Just an icon to assign to harvested metadata. The icon will be used when showing search results. **Use account** Account credentials for a basic HTTP authentication toward the remote URL. **Options** General harvesting options:

Every This is the harvesting period. The smallest value is 1 minute while the greatest value is 100 days. **One run only** If this option is checked, the harvesting will do only one run after which it will become inactive. **Validate** If checked, the metadata will be validate during import. If the validation does not pass, the metadata will be skipped. **Recurse** When the harvesting engine will find folders, it will recursively descend into them. **Privileges** Here it is possible to assign privileges to imported metadata. The Groups area lists all available groups in GeoNetwork. Once one (or more) group has been selected, it can be added through the Add button (each group can be added only once). For each added group, a row of privileges is created at the bottom of the list to allow privilege selection. To remove a row simply

press the associated Remove button on its right. Categories Here you can assign local categories to harvested metadata.

At the bottom of the page there are the following buttons:

Back Go back to the main harvesting page. The harvesting is not added. Save Saves node's data creating a new harvesting node. Then it will go back to the main harvesting page.

Adding a CSW node

This type of harvesting is capable of connecting to a remote CSW server and retrieving all matching metadata. Please, note that in order to be harvested metadata must have one of the schema format handled by GeoNetwork. Figure 10.6, "Adding a Catalogue Services for the Web harvesting node" shows the options available, whose meaning is the following:

HARVESTING MANAGEMENT

SITE

Name

Capabilities URL

Icon **CSW**

Use account

Username

Password

SEARCH CRITERIA

Search criteria

Free text

Title

Abstract

Subject

OPTIONS

Every : : (days : hours : minutes)

One run only

PRIVILEGES

Groups

Group	View	Interactive map	Featured	
All	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="button" value="Remove"/>
Sample group	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="button" value="Remove"/>

CATEGORIES

Figure 10.6. Adding a Catalogue Services for the Web harvesting node

Site Here you have to specify the connection parameters which are similar to the web DAV harvesting. In this case the URL points to the capabilities document of the CSW server. This document is used to discover the location of the services to call to query and retrieve metadata. **Search** Using the Add button, you can add several search criteria. You can query only the fields recognised by the CSW protocol. **Options** General harvesting options:

Every This is the harvesting period. The smallest value is 1 minute while the greatest value is 100 days. **One run only** If this option is checked, the harvesting will do only one run after which it will become inactive. **Privileges** Please, see web DAV harvesting. **Categories** Please, see web DAV harvesting.

At the bottom of the page there are the following buttons:

Back Go back to the main harvesting page. The harvesting is not added. Save Saves node's data creating a new harvesting node. Then it will go back to the main harvesting page.

Adding an OAI-PMH node

An OAI-PMH server implements a harvesting protocol that GeoNetwork, acting as a client, can use to harvest metadata. If you are requesting the oai_dc output format, GeoNetwork will convert it into its Dublin Core format. Other formats can be harvested only if GeoNetwork supports them and is able to autodetect the schema from the metadata. Figure 10.7, “ Adding an OAI-PMH harvesting node ” shows all available options, which are:

HARVESTING MANAGEMENT

SITE

Name

URL

Icon **REMOTE**



Use account



Username

Password

SEARCH CRITERIA

Criteria

From  

Until  

Set

Prefix

OPTIONS

Every : : (days : hours : minutes)

One run only

Validate

PRIVILEGES


Groups


Group	View	Interactive map	Featured	
Sample group	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Remove
All	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Remove


CATEGORIES

Figure 10.7. Adding an OAI-PMH harvesting node

Site All options are the same as web DAV harvesting. The only difference is that the URL parameter here points to an OAI-PMH server. This is the entry point that GeoNetwork will use to issue all PMH commands. Search This part allows you to restrict the harvesting to specific metadata subsets. You can specify several searches: GeoNetwork will execute them sequentially and results will be merged to avoid the harvesting of the same metadata. Several searches allow you to specify different search criteria. In each search, you can specify the following parameters:

From You can provide a start date here. All metadata whose last change date is equal to or greater than this date will be harvested. You cannot simply edit this field but you have to use the  icon to

popup a calendar and choose the date. This field is optional so if you don't provide it the start date constraint is dropped. Use the  icon to clear the field. Until Works exactly as the from parameter but adds an end constraint to the last change date. The until date is included in the date range, the check is: less than or equal to. Set An OAI-PMH server classifies its metadata into hierarchical sets. You can request to return metadata that belong to only one set (and its subsets). This narrows the search result. Initially the drop down shows only a blank option that indicate *no set*. After specifying the connection URL, you can press the Retrieve Info button, whose purpose is to connect to the remote node, retrieve all supported sets and prefixes and fill the search drop downs. After you have pressed this button, you can select a remote set from the drop down. Prefix Here prefix means metadata format. The oai_dc prefix is mandatory for any OAI-PMH compliant server, so this entry is always present into the prefix drop down. To have this drop down filled with all prefixes supported by the remote server, you have to enter a valid URL and press the Retrieve Info button.

You can use the Add button to add one more search to the list. A search can be removed clicking the  icon on its left. Options Most options are common to web DAV harvesting. The validate option, when checked, will validate each harvested metadata against GeoNetwork's schemas. Only valid metadata will be harvested. Invalid one will be skipped. Privileges Please, see web DAV harvesting. Categories Please, see web DAV harvesting.

At the bottom of the page there are the following buttons:

Back Go back to the main harvesting page. The harvesting is not added. Save Saves node's data creating a new harvesting node. Then it will go back to the main harvesting page.

Please note that when you edit a previously created node, both the *set* and *prefix* drop down lists will be empty. They will contain only the previously selected entries, plus the default ones if they were not selected. Furthermore, the set name will not be localised but the internal code string will be displayed. You have to press the retrieve info button again to connect to the remote server and retrieve the localised name and all set and prefix information.

Adding an OGC Service (ie. WMS, WFS, WCS)

An OGC service implements a GetCapabilities operation that GeoNetwork, acting as a client, can use to produce metadata. The GetCapability document provides information about the service and the layers/feature types/coverages served. GeoNetwork will convert it into ISO19139/119 format. Figure 10.8, "Adding an OGC service harvesting node " shows all available options, which are:

Figure 10.8. Adding an OGC service harvesting node

Site: Name is the name of the catalogue and will be one of the search criteria. The type of OGC service indicates if the harvester has to query for a specific kind of service. Supported type are WMS (1.0.0 and 1.1.1), WFS (1.0.0 and 1.1.0, WCS (1.0.0) and WPS (0.4.0 and 1.0.0). The service URL is the URL of the service to contact (without parameters like "REQUEST=GetCapabilities", "VERSION=", ...). It has to be a valid URL like `http://your.preferred.ogcservice/type_wms`. The metadata language has to be specified. It will define the language of the metadata. It should be the language used by the web service administrator. The ISO topic category is used to populate the metadata. It is recommended to choose on as the topic is mandatory for the ISO standard if the hierarchical level is "datasets".

The type of import allows to define if the harvester has to produce only one metadata for the service or if it should loop over datasets served by the service and produce also metadata for each datasets. For each dataset the second checkbox allow to generate metadata for the dataset using an XML document referenced in the MetadataUrl attribute of the dataset in the GetCapability document. If this document is loaded but it is not valid (ie. unknown schema, bad XML format), the GetCapability document is used. For WMS, thumbnails could be created during harvesting.

Icons and privileges are defined as in the other harvester types.

Metadata for the harvested service is linked to the category selected for the service (usually "interactive resources" should be the best category). For each dataset, the "category for datasets" is linked to each metadata for datasets.

11. Metadata ownership

11.1 Introduction

Starting from release 2.1.0, GeoNetwork has a new metadata access policy. The old edit and admin privileges have been removed and the concept of reviewer has been introduced. The purpose of this new profile is to control when a metadata can be published outside or not. In previous releases, all users belonging to a group with edit privileges could edit the same metadata. Now, a metadata is only visible to its creator, to a reviewer which has access to the group owner and to an administrator.

11.2 Access policy

A public metadata is a metadata that has the view privilege for the group named all.

Visualisation

An administrator can view any metadata.

A reviewer can view a metadata if:

1. The metadata owner is member of one of the groups assigned to the reviewer.
2. She/he is the metadata owner.

A user administrator or an editor can view:

1. All metadata that has the view privilege selected for one of the groups she/he is member of.
2. All metadata created by her/him.

A registered user can view:

1. All metadata that has the view privilege selected for one of the groups she/he is member of.

Public metadata can be viewed by any user (logged in or not).

Editing

An administrator can edit any metadata.

A reviewer can edit a metadata if:

1. The metadata owner is member of one of the groups assigned to the reviewer.
2. She/he is the metadata owner.

A User Administrator or an Editor can only edit metadata she/he created.

11.3 Privileges

The Privileges administration page is accessible only by:

1. All Administrators
2. All Reviewers that are member of one of the groups assigned to the metadata owner.

3. The Owner of the metadata

Privileges for the All and Intranet groups can only be edited by Administrators and Reviewers.

11.4 Transfer Ownership

When metadata ownership needs to be transferred from one user to another for all or specific metadata records, the Transfer Ownership option is available. It is located in the Administration page (Figure 11.1, “How to open the Transfer Ownership page”) and once selected, leads to the page shown in Figure 11.2, “The Transfer Ownership page”.

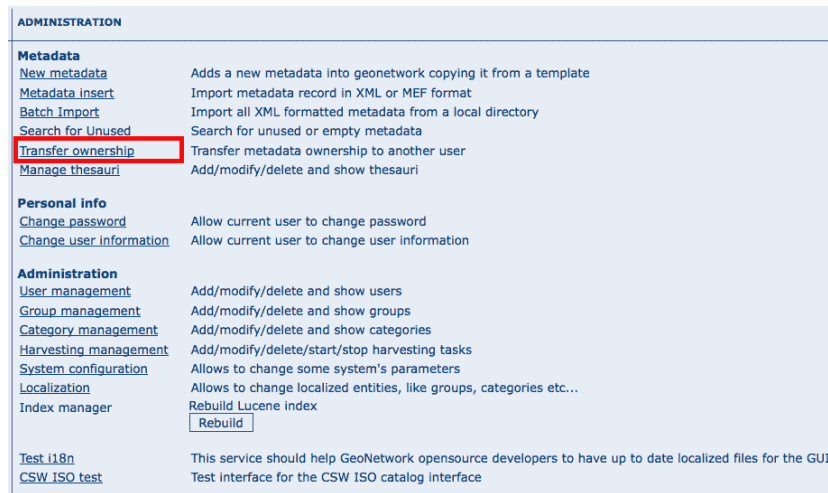


Figure 11.1. How to open the Transfer Ownership page

Initially, the page shows only a dropdown for a Source editor (the current metadata owner). The dropdown is filled with all GeoNetwork Users that have the Editor role and own some metadata. Selecting an Editor will select all metadata that is managed by that Editor. An empty dropdown means that there are no Editors with metadata associated and hence no transfer is possible.

Note The drop down will be filled with all Editors visible to you. If you are not an Administrator, you will view only a subset of all Editors.

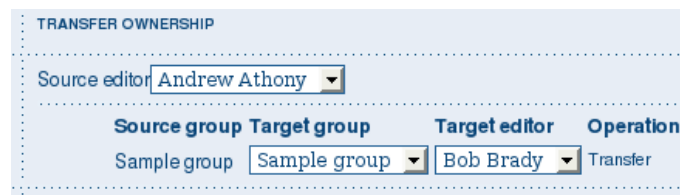


Figure 11.2. The Transfer Ownership page

Once a Source Editor has been selected, a set of rows is displayed. Each row refers to the group of the Editor for which there are privileges. The meaning of each column is the following:

1. *Source group*: This is a group that has privileges in the metadata that belong to the source editor. Put in another way, if one of the editor’s metadata has privileges for one group, that group is listed here.
2. *Target group*: This is the destination group of the transferring process. All privileges relative to the source group are transferred to the target group. The target group drop down is filled with all groups visible to the logged user (typically an administrator or a user administrator). By default, the Source group is selected in the target drop down. Privileges to groups All and Intranet are not transferable.

3. *Target editor*: Once a Target group is selected, this drop down is filled with all editors that belong to that Target group.
4. *Operation*: Currently only the Transfer operation is possible.

By selecting the Transfer operation, if the Source group is different than the Target group, the system performs the Transfer of Ownership, shows a brief summary and removes the current row because now there are no privileges to transfer anymore.

12. Thesaurus

12.1 Introduction

Thesaurus support in GeoNetwork allows:

- *Metadata editing*: controlled vocabulary on the metadata editing interface for ISO and Dublin Core
- *Administration interface* allows import/export/creation/browse thesaurus
- *Search interface*: a list of keyword is proposed for the keyword criteria

On a node, thesaurus types could be defined as:

- *External*: When a thesaurus is imported, it is flagged to "external" which means that users are not allowed to edit the thesaurus. This thesaurus is managed by an external organisation.
- *Local*: When a thesaurus is created, it is flagged to "local" which means that users are allowed to edit the thesaurus.

12.2 Thesaurus / SKOS format

The Simple Knowledge Organisation Systems (*SKOS*) <http://www.w3.org/2004/02/skos/> is an area of work developing specifications and standards to support the use of knowledge organisation systems (KOS) such as thesauri, classification schemes. This format is used by GeoNetwork to store thesaurus information.

A concept is defined by an identifier, a preferred label, a definition and links with other concepts. Labels and definitions could be stored in multiple languages (using the `xml:lang` attributes). Three type of links between concepts have been defined in the SKOS format:

- related terms
- broader terms
- narrower terms

For example, a concept "ABLETTE" could be defined as follow with a label in French and English, linked to broader concept:

```
<skos:Concept rdf:about="http://www.oieau.org/concept#c4fc54576dc00227b82a709287ac3681">
  <skos:prefLabel xml:lang="fr">ABLETTE</skos:prefLabel>
  <skos:prefLabel xml:lang="en">BLEAK</skos:prefLabel>
  <skos:broader rdf:resource="http://www.oieau.org/concept#9f25ece36d04776e09492c66627cccb9"/>
</skos:Concept>
```

GeoNetwork support multilingual thesaurus (e.g. Agrovoc). Search and edition are made based on current user interface language (i.e. if the interface is in English, when editing metadata, GeoNetwork will only search for concept in English).

12.3 Thesaurus administration

To reach the thesaurus administration page you have to be logged in as an administrator. From the administration page, click the link "Manage thesauri". Figure 5.3 shows the list of thesaurus available

in the GeoNetwork node. The page shows a list of thesaurus that have been created or imported. The upper part of the page allows user to edit/add/modify/consult thesaurus. The lower part allows upload of thesaurus in SKOS format.

Creation of a new thesaurus

To create a new thesaurus, click the "+" sign in the category you want your thesaurus to be in. Once created, the thesaurus could be updated through the edit interface. The meaning of each column is as follows:

Type The type allows to classify thesaurus according to its type. First, is defined the type of the thesaurus following ISO category list, then the type indicates if the thesaurus is a local one or an external one. Name This is the thesaurus's name provided by the administrator on creation or filename on upload. When creating a thesaurus, the name of the thesaurus will be the filename of the thesaurus.

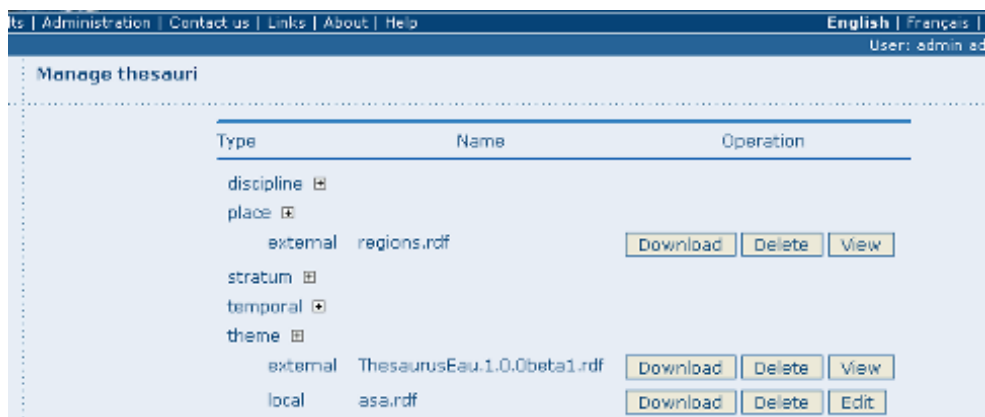


Figure 12.1. Administration interface for thesaurus

For each thesaurus the following buttons are available:

Download Link to the RFD file. Delete Remove thesaurus from the current node. View If type is external, the view button allows to search and view concepts. Edit If type is local, the edit button allows to search, add, remove and view concepts.

Import existing thesaurus

GeoNetwork allows thesaurus import in SKOS format. Once uploaded, an external thesaurus could not be updated. Select the category, browse for the thesaurus file and click upload. The file is located in / web/xml/codelist/external/thesauri/category/.

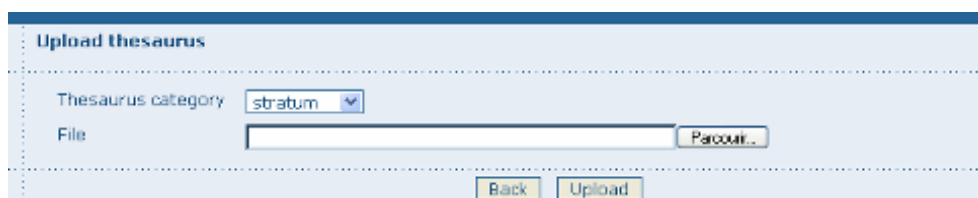


Figure 12.2. Upload interface for thesaurus

At the bottom of the page there are the following buttons:

1. *Back*: Go back to the main administration page.

2. *Upload*: Upload the selected RFD file to the node. Then it will list all thesaurus available on the node.

12.4 Editing/browsing thesaurus: add/remove/browse keywords

From the thesaurus administration interface, click on the edit button for a local thesaurus or the view button for an external thesaurus. This interface allows:

- keywords search
- add/remove keywords for local thesaurus.

Use the textbox and the type of search in order to search for keywords.



Figure 12.3. Browse interface for thesaurus

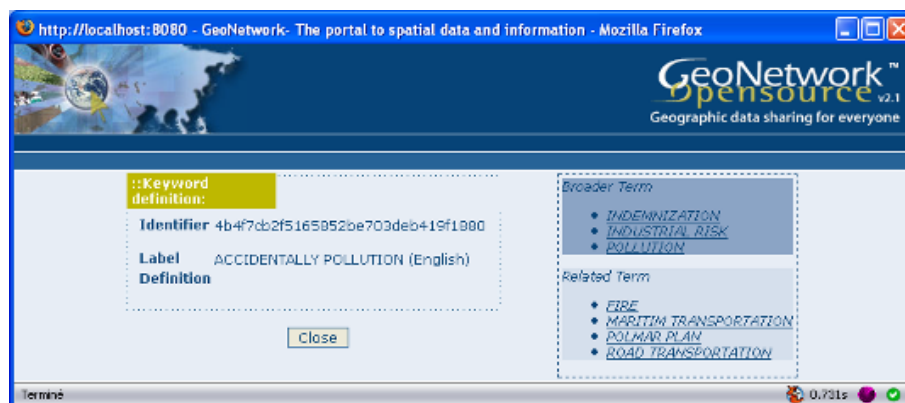


Figure 12.4. Keyword description

12.5 Metadata editing: adding keywords

When editing metadata in ISO or Dublin core, the keyword fields auto-complete when editor fill the fields. Keywords available in all thesaurus know by the current node are returned. Editor could select one of the list or could type any other keywords.

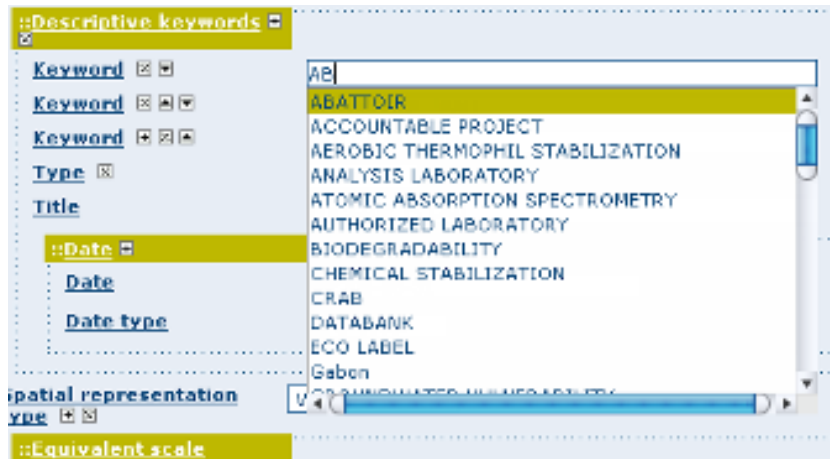


Figure 12.5. Auto-complete in keywords editor

12.6 Search criteria: keywords

In the advanced search interface, the keyword field will proposed all keywords used in the metadata. These keywords are indexed by Lucene on creation/update of metadata. The number of metadata linked to all keywords available in the index are display. User could type in the keyword field or click the icon to get the list of keywords available.

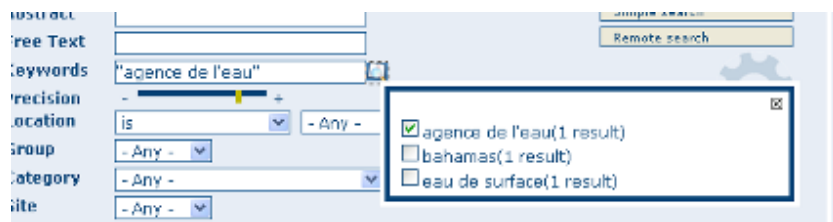


Figure 12.6. Thesaurus search interface

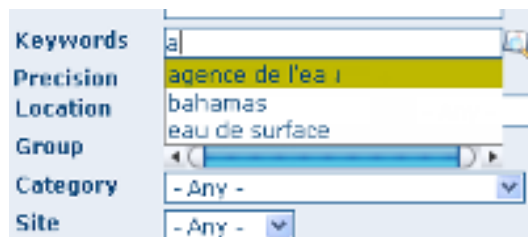


Figure 12.7. Auto-complete function in thesaurus search interface

13. GeoNetwork's Administrator Survival Tool - GAST

13.1 What is GAST?

GAST stands for GeoNetwork's Administrator Survival Tool and is a standalone application whose purpose is to simplify some low level tasks like change of the servlet, configuration of the *JDBC* account, setup the database and so on. Most of the GAST's facilities work only for the GeoNetwork's installation where GAST is in. This implies that if you are using a servlet container other than Jetty (like Tomcat) you will not be able to change some options (like the servlet's name). Other facilities work for any servlet container but you have to specify the GeoNetwork's URL into the GAST's configuration dialogue.

13.2 Starting GAST

GAST belongs to the core components so it is installed by default.

On Windows computers, simply select the `Start GAST` option under the GeoNetwork opensource program group under `Start > Programs > GeoNetwork opensource`

Other options to start GAST are either to use a Java command **from a terminal window** or just click its jar's icon. To issue the Java command you have to:

1. change directory to the GeoNetwork installation folder
2. issue the command `java -jar gast/gast.jar`

GAST will be in current system language if any translation is available. If you want to force GAST GUI language, you could start GAST using the `-Duser.language` option (e.g. `./gast.sh -Duser.language=fr` or `java -Duser.language=fr -jar gast/gast.jar`).

You can also try to simply open the GeoNetwork installation folder, go to the `gast` folder and double click on the `gast.jar` file. If you have Java installed, GAST should start in a few seconds.

To run, GAST requires Java 1.5. It will not work on Java 1.4 and it should run on Java 1.6 (this has not been tested!).

13.3 Operating modes

When you start GAST, you get an application window like the one in Figure 13.1, "GAST's main window with a tool selected". On the left side you have a panel with the tools you can use. After selecting a tool, on the right side you get the tool's options panel.

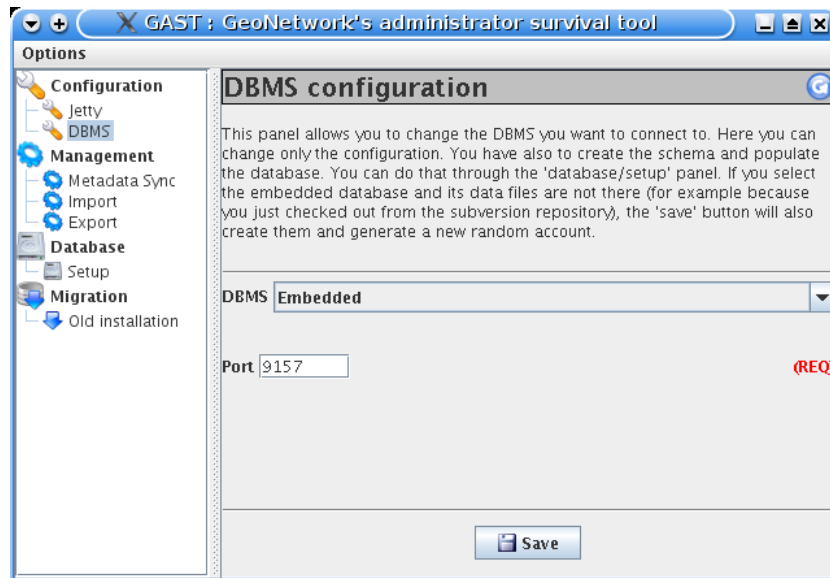





Figure 13.1. GAST's main window with a tool selected

Every function has an operating mode, which defines the condition under which the tool can be used. The tool's mode is shown with an icon on the right side of the tool's name. The operating modes, with their icons are summarised in the following table:

Mode	Icon	Description
Restarted		The tool can be always used, but GeoNetwork must be restarted in order to make the change effective.
Running		The tool can be used only if GeoNetwork is running.
Stopped		The tool can be used only if GeoNetwork is stopped. This is important because some tools change the database's account or create the database from scratch. These are sensitive operations that cannot be performed while GeoNetwork is running.

13.4 Tools subdivision

All GAST tools present into the left panel are logically subdivided into groups. Each group represents a GeoNetwork's aspect for which GAST allows you a graphic interface. The groups are:

Configuration You can change some configuration parameters, like the servlet's name, JDBC account etc... **Management** General purpose tools related to the site's administration. **Database Operations** that regard the database. Here you can find tools to create a database from scratch, creating the schema and filling it with proper data. **Migration** Tools that allow you to migrate metadata from old installation.

13.5 Server and Account configuration dialogue

Some of the GAST's tools access a running GeoNetwork application. Usually, GAST connects to GeoNetwork using the connection parameters it finds on the installation folder but you can specify other parameters in order to connect to other instances. This is required when the GeoNetwork instance is not

running on the embedded Jetty server. In addition to that, some tools require authentication so account parameters must be provided.

To provide these parameters, you have to use the GAST's configuration dialogue. To open the dialogue, select `Options >> Config` from the menu bar. You will get the dialogue shown in Figure 13.2, "The configuration dialogue".

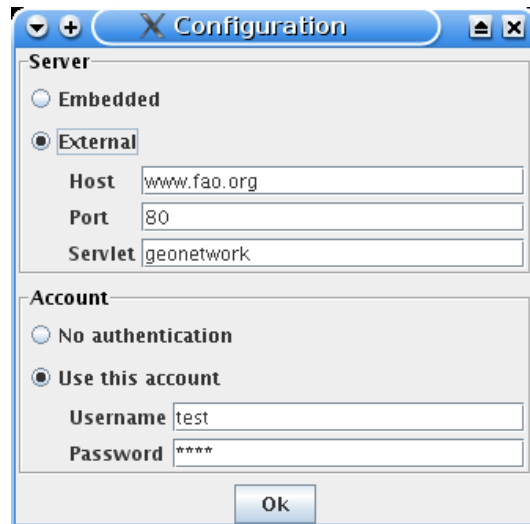


Figure 13.2. The configuration dialogue

The dialogue is subdivided into 2 areas: **Server** Tells GAST how to connect to a running GeoNetwork. If you select the embedded option, GAST will get the connection parameters from the installation directory. Alternatively, if you use Tomcat or an external servlet container you have to choose the external option and provide the connection parameters yourself. Remember that this will work only for tools which operating mode is *Running*. For all the others, GAST will access the parameters from the installation directory. **Account** Some tools require authentication. To authenticate, simply select the Use this account option and provide the username and password of a valid account. These parameters will work for both the embedded instance and for any external instance.

14. Localisation

14.1 Localisation of dynamic user interface elements

The user interface of GeoNetwork can be localised into several languages through XML language files. But beside static text, there is also more dynamic text that can be added and changed interactively. This text is stored in the database and is translated using the Localisation form that is part of the administrative functions (Figure 14.1, “How to open the Localisation form”).

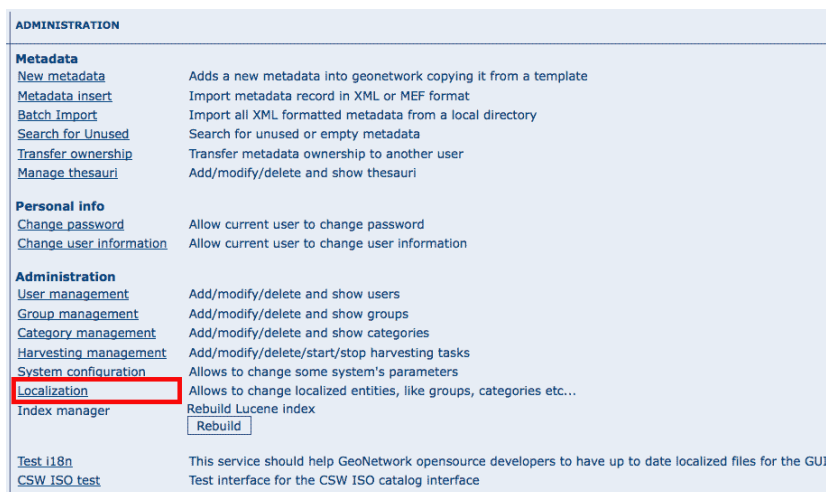


Figure 14.1. How to open the Localisation form

The form allows you to localise the following entities: *Groups*, *Categories*, *Operations* and *Regions*. The form (Figure 14.2, “The Localisation form”) subdivided in a left and a right panel.

The left panel allows you to choose which elements you want to edit. On the top, a dropdown let you choose which entity to edit. All elements of the selected type are shown in a list.

When you select an element from the list, the right panel will show the text as it will be displayed in the user interface. The text in the source language is read only while you can update the text in the target language field.



Note

You can change the source and target languages to best suit your needs. Some users may for instance prefer to translate from French to Spanish, others prefer to work with English as the source language.

Use the Save button to store the updated label and move to the next element.



Important

If the user changes a label and chooses another target language without saving, the label change is lost.

Localization

Entity Source language

maps
datasets
interactiveResources
applications
caseStudies
proceedings
photo
audioVideo
directories
otherResources

Case studies, best practices

Target language

Etude de cas, meilleurs pratiques

Figure 14.2. The Localisation form

15. Import / export tools

15.1 Introduction

Using GAST, you can import and export metadata at will. It allows you to:

1. Create a **backup** of the entire metadata set. Each metadata has its own file, including maps and other data files. Once you have the backup, you can decide to import all or only some of them.
2. **Move your metadata** from one GeoNetwork catalogue to another. This can be done to mirror your metadata or to upgrade an old installation. In the last case, you export your metadata from your old installation and then re-import them into the new one.
3. Fill the system with **test data**. Using the 'skip UUID' option, you can re-import the same metadata over and over again. This is useful, for example, if you want to perform stress tests.

Metadata are exported using the *MEF* format.



Ownership

Please, consider that the MEF format version 1.0 does not take into account user and group ownership. When exporting metadata, you lose this information. When importing metadata, the new owner becomes the user that is performing the import while the group ownership is set to null.

15.2 Import

This tool is located under Management tools on the left panel and allows you to import a set of metadata that have been previously exported using the export facility (see Section 15.3, "Export"). Selecting the Import tool opens the option panel on the right (Figure 15.1, "The metadata import panel").

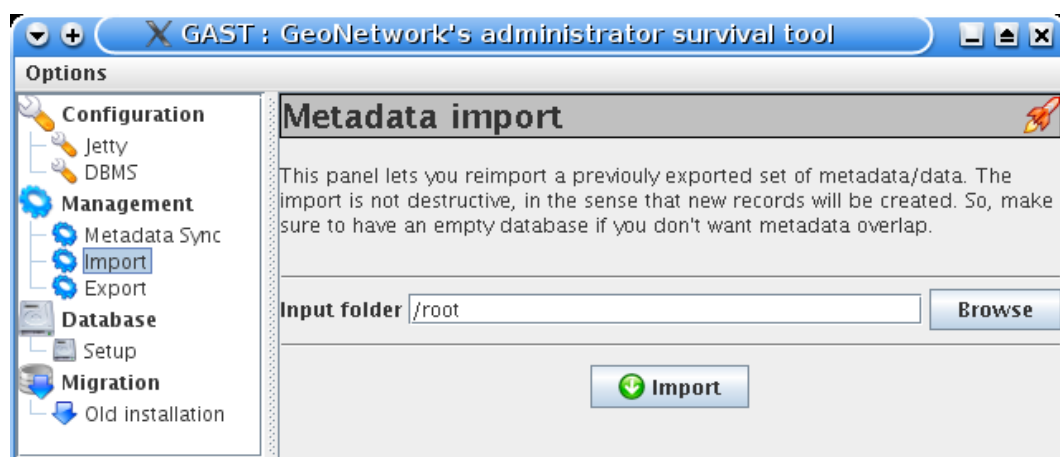


Figure 15.1. The metadata import panel

- **Input folder.** the source folder in your system that GAST will scan to collect metadata to import. GAST will try to import all files with the MEF extension.



Note

Sub-folders are not scanned.

- **Browse button.** Navigate through your file system to choose an output location or enter it manually into the text field.
- **Import.** This will start the process. A progress dialogue will be opened to show the import status.

15.3 Export

This tool is located under the Management tool on the left panel and allows you to export a set of metadata using the MEF format. Selecting the Export tool opens the option panel on the right (Figure 15.2, “The metadata export panel”).

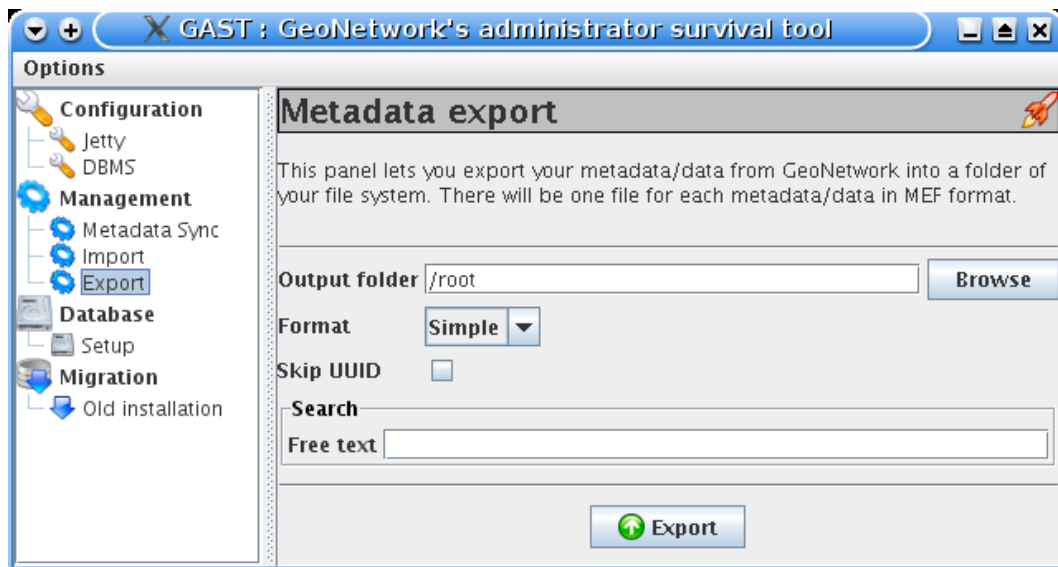


Figure 15.2. The metadata export panel

- **Output folder.** The target folder in your file system where GAST will put the exported metadata. You can either select the Browse button to navigate through your file system to choose a better location or enter it manually in the text field.
- **Format.** Here you can specify the metadata's output format. See the MEF specification for more information.
- **Skip UUID.** Normally this option is not required (see Warning). If you select it, you will lose the metadata's unique identifier (UUID) but you will be able to re-import that metadata over and over again. This is useful to fill the system with test data.
- **Search.** Allows to specify free text search criteria to limit the set of exported records.



Note

the export result will depend on the metadata visible to the searching user. If you do not authenticate, you will get only public metadata.

- **Export.** This will start the export process. A progress dialogue will be opened to show the export status.



Warning

Skipping the UUID on import or export can cause metadata to be duplicated. This should normally always be avoided

Part III. Server Reference

This part gives some insights in the internal structure of GeoNetwork opensource. It describes some basic operations, like compiling and running the software, protocols used, a description of XML services and the Settings structure.

If you are a software developer and want to develop, customize or integrate GeoNetwork services, this part is for you.

16. Software development

16.1 System Requirements

GeoNetwork is a Java application that runs as a servlet so the Java Runtime Environment (JRE) must be installed in order to run it. You can get the JRE from the following address <http://java.sun.com> and you have to download the Java 5 Standard Edition (SE). GeoNetwork won't run with Java 1.4 and Java 6 has some problems with it so we recommend to use Java 5. Being written in Java, GeoNetwork can run on any platform that supports Java, so it can run on Windows, Linux and Mac OSX. For the latter one, make sure to use version 10.4 (Tiger) or newer. Version 10.3 (Panther) has only Java 1.4 so it cannot run GeoNetwork.

Next, you need a servlet container. GeoNetwork comes with an embedded one (Jetty) which is fast and well suited for most applications. If you need a stronger one, you can install Tomcat from the Apache Software Foundation (<http://tomcat.apache.org>). It provides load balance, fault tolerance and other corporate needed stuff. If you work for an organisation, it is probable that you already have it up and running. The tested version is 5.5 but GeoNetwork should work with all other versions.

Regarding storage, you need a Database Management System (DBMS) like Oracle, MySQL, PostgreSQL and so on. GeoNetwork comes with an embedded one (McKoi) which is used by default during installation. This DBMS can be used for small or desktop installations, where the speed is not an issue. You can use this DBMS for several thousands of metadata. If you manage more than 10.000 metadata it is better to use a professional, stand alone DBMS. In this case, using a separate DBMS also frees up some memory for the application.

GeoNetwork does not require a strong machine to run. A good performance can be obtained even with 128 Mb of RAM. The suggested amount is 512 Mb. For the hard disk space, you have to consider the space required for the application itself (about 40 Mb) and the space required for data maps, which can require 50 GB or more. A simple disk of 250 GB should be OK. Maybe you can choose a fast one to reduce backup time but GeoNetwork itself does not speed up on a faster disk. You also need some space for the search index which is located in `web/WEB-INF/lucene`. Even with a lot of metadata the index is small so usually 10-20 Mb of space is enough.

16.2 Running the software with a servlet engine

The software is run in different ways depending on the servlet container you are using:

Tomcat You can use the manager web application to start/stop GeoNetwork. You can also use the `startup.*` and `shutdown.*` scripts located into Tomcat's bin folder (`.*` means `.sh` or `.bat` depending on your OS) but this way you restart all applications you are running, not only GeoNetwork. After installation and before running GeoNetwork you must link it to Tomcat. **Jetty** If you use the provided container you can use the scripts into GeoNetwork's bin folder. The scripts are `start-geonetwork.*` and `stop-geonetwork.*` and you must be inside the bin folder to run them. You can use these scripts just after installation.

16.3 Development

Compiling GeoNetwork

To compile GeoNetwork you first need to install the source code during installation. If you do so, you get a `build.xml` script and a `src` folder with the full source.

You also need the Ant tool to run the build script. You can download Ant from <http://ant.apache.org>. Version 1.6.5 works but any other recent version should be OK. Once installed, you should have the ant command in your path (on Windows systems, you have to open a shell to check).

When all is in place, go inside the GeoNetwork's root folder (the one where the build.xml file is located) and issue the ant command. You should see an output like this one:

```
gemiini:/geonetwork/trunk# ant
Buildfile: build.xml
compile:
[delete] Deleting: /geonetwork/trunk/web/WEB-INF/lib/geonetwork.jar
[delete] Deleting: /geonetwork/trunk/csw/lib/csw-client.jar
[delete] Deleting: /geonetwork/trunk/csw/lib/csw-common.jar
[delete] Deleting: /geonetwork/trunk/gast/gast.jar
[mkdir] Created dir: /geonetwork/trunk/.build
[javac] Compiling 267 source files to /geonetwork/trunk/.build
[javac] Note: Some input files use or override a deprecated API.
[javac] Note: Recompile with -Xlint:deprecation for details.
[javac] Note: Some input files use unchecked or unsafe operations.
[javac] Note: Recompile with -Xlint:unchecked for details.
[copy] Copying 1 file to /geonetwork/trunk/.build
[jar] Building jar: /geonetwork/trunk/web/WEB-INF/lib/geonetwork.jar
[jar] Building jar: /geonetwork/trunk/csw/lib/csw-client.jar
[jar] Building jar: /geonetwork/trunk/csw/lib/csw-common.jar
[jar] Building jar: /geonetwork/trunk/gast/gast.jar
[delete] Deleting directory /geonetwork/trunk/.build
BUILD SUCCESSFUL
Total time: 9 seconds
gemiini:/geonetwork/trunk#
```

The compilation phase, if it has success, puts all jars into the proper place (most of them will be copied into web/geonetwork/WEB-INF/lib and web/intermap/WEB-INF/lib). After this phase, simply restart GeoNetwork to see the effects.

Source code documentation

The GeoNetwork Java source code is based on Javadoc. Javadoc is a tool for generating API documentation in HTML format from doc comments in source code. To see documentation generated by the Javadoc tool, go to:

- [GeoNetwork opensource Javadoc](#)¹
- [InterMap opensource Javadoc](#)²

Creating the installer

You can generate an installer by running the ant command inside the installer directory.

Both platform independent and Windows specific installers are generated by default.

Make sure you update version number and other relevant properties in the installer/build.xml file

You can also create an installer that includes a Java Runtime Environment (JRE) for Windows. This will allow GeoNetwork to run on a compatible, embedded JRE and thus avoid error messages caused by JRE incompatibilities on the PC.

Creating an installer with an embedded JRE requires you to first download and unzip the JRE in a folder jre1.5.0_12 at the project root level. Refer to the installer-config-win-jre.xml file for exact configuration.

17. Harvesting

17.1 Structure

The harvesting capability is built around 3 areas: JavaScript code, Java code and XSL stylesheets (on both the server and client side).

JavaScript code

This refers to the web interface. The code is located in the `web/geonetwork/scripts/harvesting` folder. Here, there is a subfolder for each harvesting type plus some classes for the main page. These are:

1. *harvester.js*: This is an abstract class that must be implemented by harvesting types. It defines some information retrieval methods (`getType`, `getLabel`, etc...) used to handle the harvesting type, plus one `getUpdateRequest` method used to build the XML request to insert or update entries.
2. *harvester-model.js*: Another abstract class that must be implemented by harvesting types. When creating the XML request, the only method `substituteCommon` takes care of adding common information like privileges and categories taken from the user interface.
3. *harvester-view.js*: This is an important abstract class that must be implemented by harvesting types. It takes care of many common aspects of the user interface. It provides methods to add group's privileges, to select categories, to check data for validity and to set and get common data from the user interface.
4. *harvesting.js*: This is the main JavaScript file that takes care of everything. It starts all the submodules, loads XML strings from the server and displays the main page that lists all harvesting nodes.
5. *model.js*: Performs all XML requests to the server, handles errors and decode responses.
6. *view.js*: Handles all updates and changes on the main page.
7. *util.js*: just a couple of utility methods.

Java code

The harvesting package is located in `src/org/fao/geonet/kernel/harvest`. Here too, there is one subfolder for each harvesting type. The most important classes for the implementor are:

1. *AbstractHarvester*: This is the main class that a new harvesting type must extend. It takes care of all aspects like adding, updating, removing, starting, stopping of harvesting nodes. Some abstract methods must be implemented to properly tune the behaviour of a particular harvesting type.
2. *AbstractParams*: All harvesting parameters must be enclosed in a class that extends this abstract one. Doing so, all common parameters can be transparently handled by this abstract class.

All others are small utility classes used by harvesting types.

XSL stylesheets

Stylesheets are spread in some folders and are used by both the JavaScript code and the server. The main folder is located at `web/geonetwork/xsl/harvesting`. Here there are some general stylesheets, plus one subfolder for each harvesting type. The general stylesheets are:

1. *buttons.xsl*: Defines all button present in the main page (*activate, deactivate, run, remove, back, add, refresh*), buttons present in the "add new harvesting" page (*back* and *add*) and at the bottom of the edit page (*back* and *save*).
2. *client-error-tip.xsl*: This stylesheet is used by the browser to build tooltips when an harvesting error occurred. It will show the error class, the message and the stacktrace.
3. *client-node-row.xsl*: This is also used by the browser to add one row to the list of harvesting nodes in the main page.
4. *harvesting.xsl*: This is the main stylesheet. It generates the HTML page of the main page and includes all panels from all the harvesting nodes.

In each subfolder, there are usually 4 files:

1. *xxx.xsl*: This is the server stylesheets who builds all panels for editing the parameters. XXX is the harvesting type. Usually, it has the following panels: site information, search criteria, options, privileges and categories.
2. *client-privil-row.xsl*: This is used by the JavaScript code to add rows in the group's privileges panel.
3. *client-result-tip.xsl*: This is used by the JavaScript code (which inherits from *harvester-view.js*) to show the tool tip when the harvesting has been successful.
4. *client-search-row.xsl*: Used in some harvesting types to generate the HTML for the search criteria panel.

As you may have guessed, all client side stylesheets (those used by JavaScript code) start with the prefix *client-*.

Another set of stylesheets are located in `web/geonetwork/xsl/xml/harvesting` and are used by the `xml.harvesting.get` service. This service is used by the JavaScript code to retrieve all the nodes the system is currently harvesting from. This implies that a stylesheet (one for each harvesting type) must be provided to convert from the internal setting structure to an XML structure suitable to clients.

The last file to take into consideration contains all localised strings and is located at `web/geonetwork/loc/XX/xml/harvesting.xml` (where XX refers to a language code). This file is used by both JavaScript code and the server.

17.2 Data storage

Harvesting nodes are stored inside the Settings table. Further useful information can be found in the chapter Harvesting.

The SourceNames table is used to keep track of the `uuid/name` couple when metadata get migrated to different sites.

17.3 Guidelines

To add a new harvesting type, follow these steps:

1. Add the proper folder in `web/scripts/harvesting`, maybe copying an already existing one.
2. Edit the `harvesting.js` file to include the new type (edit both constructor and init methods).

3. Add the proper folder in web/xsl/harvesting (again, it is easy to copy from an already existing one).
4. Edit the stylesheet web/xsl/harvesting/harvesting.xsl and add the new type
5. Add the transformation stylesheet in web/xsl/xml/harvesting. Its name must match the string used for the harvesting type.
6. Add the Java code in a package inside org.fao.geonet.kernel.harvest.harvester.
7. Add proper strings in web/geonetwork/loc/XX/xml/harvesting.xml.

Here follows a list of general notes to follow when adding a new harvesting type:

1. Every harvesting node (not type) must generate its UUID. This UUID is used to remove metadata when the harvesting node is removed and to check if a metadata (which has another UUID) has been already harvested by another node.
2. If a harvesting type supports multiple searches on a remote site, these must be done sequentially and results merged.
3. Every harvesting type must save in the folder images/logos a GIF image whose name is the node's UUID. This image must be deleted when the harvesting node is removed. This is necessary to propagate harvesting information to other GeoNetwork nodes.
4. When a harvesting node is removed, all collected metadata must be removed too.
5. During harvesting, take in mind that a metadata could have been removed just after being added to the result list. In this case the metadata should be skipped and no exception raised.
6. The only settable privileges are: view, dynamic, featured. It does not make sense to use the others.
7. If a node raises an exception during harvesting, that node will be deactivated.
8. If a metadata already exists (its UUID exists) but belong to another node, it must not be updated even if it has been changed. This way the harvesting will not conflict with the other one. As a side effect, this prevent locally created metadata from being changed.
9. The harvesting engine does not store results on disk so they will get lost when the server will be restarted.
10. When some harvesting parameters are changed, the new harvesting type must use them during the next harvesting without requiring to reboot the server.

18. Metadata Exchange Format v1.1

18.1 Introduction

The metadata exchange format (MEF in short) is a special designed file format whose purpose is to allow metadata exchange between different platforms. A metadata exported into this format can be imported by any platform which is able to understand it. This format has been developed with GeoNetwork in mind so the information it contains is mainly related to it. Nevertheless, it can be used as an interoperability format between any platform.

This format has been designed with these needs in mind:

1. Export a metadata record for backup purposes
2. Import a metadata record from a previous backup
3. Import a metadata record from a different GeoNetwork version to allow a smooth migration from one version to another.

All these operations regard the metadata and its related data as well.

In the paragraphs below, some terms should be intended as follows:

1. the term actor is used to indicate any system (application, service etc...) that operates on metadata.
2. the term reader will be used to indicate any actor that can import metadata from a MEF file.
3. the term writer will be used to indicate any actor that can generate a MEF file.

18.2 File format

A MEF file is simply a ZIP file which contains the following files:

1. *metadata.xml*: this file contains the metadata itself, in XML format. The text encoding of the metadata is that one specified into the XML declaration.
2. *info.xml*: this is a special XML file which contains information related to the metadata but that cannot be stored into it. Examples of such information are the creation date, the last change date, privileges on the metadata and so on. Now this information is related to the GeoNetwork's architecture.
3. *public*: this is a directory used to store the metadata thumbnails and other public files. There are no restrictions on the images' format but it is strongly recommended to use the portable network graphics (PNG), the JPEG or the GIF formats.
4. *private*: this is a directory used to store all data (maps, shape files etc...) associated to the metadata. Files in this directory are *private* in the sense that an authorisation is required to access them. There are no restrictions on the file types that can be stored into this directory.

Any other file or directory present into the MEF file should be ignored by readers that don't recognise them. This allows actors to add custom extensions to the MEF file.

A MEF file can have empty public and private folders depending on the export format, which can be:

1. *simple*: both public and private are omitted.
2. *partial*: only public files are provided.

3. *full*: both public and private files are provided.

It is recommended to use the .mef extension when naming MEF files.

18.3 The info.xml file

This file contains general information about a metadata. It must have an info root element with a mandatory version attribute. This attribute must be in the X.Y form, where X represents the major version and Y the minor one. The purpose of this attribute is to allow future changes of this format maintaining compatibility with older readers. The policy behind the version is this:

1. A change to Y means a minor change. All existing elements in the previous version must be left unchanged: only new elements or attributes may be added. A reader capable of reading version X.Y is also capable of reading version X.Y' with Y'>Y.
2. A change to X means a major change. Usually, a reader of version X.Y is not able to read version X'.Y with X'>X.

The root element must have the following children:

1. *general*: a container for general information. It must have the following children:
 - a. *UUID*: this is the universally unique identifier assigned to the metadata and must be a valid UUID. This element is optional and, when omitted, the reader should generate one. A metadata without a UUID can be imported several times into the same system without breaking uniqueness constraints. When missing, the reader should also generate the siteld value.
 - b. *createDate*: This date indicates when the metadata was created.
 - c. *changeDate*: This date keeps track of the most recent change to the metadata.
 - d. *siteld*: This is an UUID that identifies the actor that created the metadata and must be a valid UUID. When the UUID element is missing, this element should be missing too. If present, it will be ignored.
 - e. *siteName*: This is a human readable name for the actor that created the metadata. It must be present only if the siteld is present.
 - f. *schema*: Indicates the metadata's schema. The value can be assigned as will but if the schema is one of those describe below, that value must be used:
 - i. *dublin-core*: A metadata in the Dublin Core format as described in <http://dublincore.org>
 - ii. *fgdc-std*: A metadata in the Federal Geographic Data Committee.
 - iii. *iso19115*: A metadata in the ISO 19115 format
 - iv. *iso19139*: A metadata in the ISO 19115/2003 format for which the ISO19139 is the XML encoding.
 - g. *format*: Indicates the MEF export format. The element's value must belong to the following set: { *simple*, *partial*, *full* }.
 - h. *localId*: This is an optional element. If present, indicates the id used locally by the sourceId actor to store the metadata. Its purpose is just to allow the reuse of the same local id when reimporting a metadata.

- i. *isTemplate*: A boolean field that indicates if this metadata is a template used to create new ones. There is no real distinction between a real metadata and a template but some actors use it to allow fast metadata creation. The value must be: { *true*, *false* }.
 - j. *rating*: This is an optional element. If present, indicates the users' rating of the metadata ranging from 1 (a bad rating) to 5 (an excellent rating). The special value 0 means that the metadata has not been rated yet. Can be used to sort search results.
 - k. *popularity*: Another optional value. If present, indicates the popularity of the metadata. The value must be positive and high values mean high popularity. The criteria used to set the popularity is left to the writer. Its main purpose is to provide a metadata ordering during a search.
2. *categories*: a container for categories associated to this metadata. A category is just a name, like 'audio-video' that classifies the metadata to allow an easy search. Each category is specified by a category element which must have a name attribute. This attribute is used to store the category's name. If there are no categories, the categories element will be empty.
 3. *privileges*: a container for privileges associated to this metadata. Privileges are operations that a group (which represents a set of users) can do on a metadata and are specified by a set of group elements. Each one of these, has a mandatory name attribute to store the group's name and a set of operation elements used to store the operations allowed on the metadata. Each operation element must have a name attribute which value must belong to the following set: { *view*, *download*, *notify*, *dynamic*, *featured* }. If there are no groups or the actor does not have the concept of group, the privileges element will be empty. A group element without any operation element must be ignored by readers.
 4. *public*: All metadata thumbnails (and any other public file) must be listed here. This container contains a file element for each file. Mandatory attributes of this element are name, which represents the file's name and changeDate, which contains the date of the latest change to the file. The public element is optional but, if present, must contain all the files present in the metadata's public directory and any reader that imports these files must set the latest change date on these using the provided ones. The purpose of this element is to provide more information in the case the MEF format is used for metadata harvesting.
 5. *private*: This element has the same purpose and structure of the public element but is related to maps and all other private files.

Any other element or attribute should be ignored by readers that don't understand them. This allows actors to add custom attributes or subtrees to the XML.

Figure 18.1, "Example of info file" shows an example of info file.

Date format

Unless differently specified, all dates in this file must be in the ISO/8601 format. The pattern must be YYYY-MM-DDTHH:mm:ss and the timezone should be the local one.

```
<info version="1.0">
  <general>
    <UUID>0619abc0-708b-eea-8202-000d98959033</uuid>
    <createDate>2006-12-11T10:33:21</createDate>
    <changeDate>2006-12-14T08:44:43</changeDate>
    <siteId>0619cc50-708b-11da-8202-000d9335906e</siteId>
    <siteName>FAO main site</siteName>
    <schema>iso19139</schema>
    <format>full</format>
    <localId>204</localId>
    <isTemplate>>false</isTemplate>
  </general>
  <categories>
    <category name="maps" />
    <category name="datasets" />
  </categories>
  <privileges>
    <group name="editors">
      <operation name="view" />
      <operation name="download" />
    </group>
  </privileges>
  <public>
    <file name="small.png" changeDate="2006-10-07T13:44:32" />
    <file name="large.png" changeDate="2006-11-11T09:33:21" />
  </public>
  <private>
    <file name="map.zip" changeDate="2006-11-12T13:23:01" />
  </private>
</info>
```

Figure 18.1. Example of info file

19. XML Services

19.1 Calling specifications

Calling XML services

GeoNetwork provides access to several internal structures through the use of XML services. These are much like HTML addresses but return XML instead. As an example, consider the `xml.info` service: you can use this service to get some system's information without fancy styles and graphics. In GeoNetwork, XML services have usually the `xml.` prefix in their address.

Request

Each service accepts a set of parameters, which must be embedded into the request. A service can be called using different HTTP methods, depending on the structure of its request:

GET The parameters are sent using the URL address. On the server side, these parameters are grouped into a flat XML document with one root and several simple children. A service can be called this way only if the parameters it accepts are not structured. Figure 19.1, "A GET request to a XML service and its request encoding" shows an example of such request and the parameters encoded in XML. POST There are 3 variants of this method:

ENCODED The request has one of the following content types: `application/x-www-form-urlencoded` or `multipart/form-data`. The first case is very common when sending web forms while the second one is used to send binary data (usually files) to the server. In these cases, the parameters are not structured so the rules of the GET method applies. Even if the second case could be used to send XML documents, this possibility is not considered on the server side.

XML The content type is `application/xml`. This is the common case when the client is not a browser but a specialised client. The request is a pure XML document in string form, encoded using the encoding specified into the prologue of the XML document. Using this form, any type of request can be made (structured or not) so any service can be called.

SOAP The content type is `application/soap+xml`. SOAP is a simple protocol used to access objects and services using XML. Clients that use this protocol can embed XML requests into a SOAP structure. On the server side, GeoNetwork will remove the SOAP structure and feed the content to the service. Its response will be embedded again into a SOAP structure and sent back to the caller. It makes sense to use this protocol if it is the only protocol understood by the client.

```
<request>
  <hitsPerPage>10</hitsPerPage>
  <any />
</request>
```

Figure 19.1. A GET request to a XML service and its request encoding

Response

The response of an XML service always has a content type of `application/xml` (the only exception are those services which return binary data). The document encoding is the one specified into the document's prologue. Anyway, all GeoNetwork services return documents in the UTF-8 encoding.

On a GET request, the client can force a SOAP response adding the `application/soap+xml` content type to the Accept header parameter.

Exception handling

A response document having an error root element means that the XML service raised an exception. This can happen under several conditions: bad parameters, internal errors et cetera. In this cases the returned XML document has the following structure:

- *error*: This is the root element of the document. It has a mandatory *id* attribute that represents an identifier of the error from a common set. See Table 19.1, “Summary of error ids” for a list of all id values.
 - *message*: A message related to the error. It can be a short description about the error type or it can contain some other information that completes the id code.
 - *class*: The Java class of the raised error (name without package information).
 - *stack*: The server’s stacktrace up to the point that generated the exception. It contains several at children, one for each nested level. Useful for debugging purposes.
 - *at*: Information about a nested level of called code. It has the following mandatory attributes:
class Java class of the called method. *method* Java called method. *line* Line, inside the called method’s source code where there the method call of the next nested level. *file* Source file where the class is defined.
 - *object*: An optional container for parameters or other values that caused the exception. In case a parameter is an XML object, this container will contain that object in XML form.
 - *request*: A container for some useful information that can be needed to debug the service.
 - *language*: Language used when the service was called.
 - *service*: Name of the called service.

Table 19.1. Summary of error ids

<i>id</i>	Meaning of message element	Meaning of object element
<i>error</i>	General message, human readable	
<i>bad-format</i>	Reason	-
<i>bad-parameter</i>	Name of the parameter	Parameter's bad value
<i>file-not-found</i>	-	File's name
<i>file-upload-too-big</i>	-	-
<i>missing-parameter</i>	Name of the parameter	XML container where the parameter should have been present.
<i>object-not-found</i>	-	Object's name
<i>operation-aborted</i>	Reason of abort	If present, the object that caused the abort
<i>operation-not-allowed</i>	-	-
<i>resource-not-found</i>	-	Resource's name
<i>service-not-allowed</i>	-	Service's name
<i>service-not-found</i>	-	Service's name
<i>user-login</i>	User login failed message	User's name
<i>user-not-found</i>	-	User's id or name
<i>metadata-not-found</i>	The requested metadata was not found	Metadata's id

Figure 19.2, "An example of generated exception" shows an example of exception generated by the `mef.export` service. The service complains about a missing parameter, as you can see from the content of the `id` attribute. The object element contains the xml request with an unknown test parameter while the mandatory UUID parameter (as specified by the message element) is missing.

```

<error>
  <message>UUID</message>
  <class>MissingParameterEx</class>
  <stack>
    <at class="jeeves.utils.Util" file="Util.java" line="66"
      method="getParam" />
    <at class="org.fao.geonet.services.mef.Export" file="Export.java"
      line="60" method="exec" />
    <at class="jeeves.server.dispatchers.ServiceInfo" file="ServiceInfo.java"
      line="226" method="execService" />
    <at class="jeeves.server.dispatchers.ServiceInfo" file="ServiceInfo.java"
      line="129" method="execServices" />
    <at class="jeeves.server.dispatchers.ServiceManager" file="ServiceManager.java"
      line="370" method="dispatch" />
  </stack>
  <object>
    <request>
      <asd>ee</asd>
    </request>
  </object>
  <request>
    <language>en</language>
    <service>mef.export</service>
  </request>
</error>

```

Figure 19.2. An example of generated exception

19.2 General services

xml.info

The xml.info service can be used to query the site about its configuration, services, status and so on. For example, it is used by the harvesting web interface to retrieve information about a remote node.

Request

The XML request should contain at least one type element to indicates the kind of information to retrieve. More type elements can be specified to obtain more information at once. The set of allowed values are:

1. *site*: Returns general information about the site like its name, id, etc...
2. *categories*: Returns all site's categories
3. *groups*: Returns all site's groups visible to the requesting user. If the user does not authenticate himself, only the Intranet and the all groups are visible.
4. *operations*: Returns all possible operations on metadata
5. *regions*: Returns all geographical regions usable for queries
6. *sources*: Returns all GeoNetwork sources that the remote site knows.

The result will contain:

- The remote node's name and siteld
- All source UUIDs and names that have been discovered through harvesting.
- All source UUIDs and names of metadata that have been imported into the remote node through the MEF format.

- Administrators can see all users into the system (normal, other administrators, etc...)
- User administrators can see all users they can administrate and all other user administrators in the same group set. The group set is defined by all groups visible to the user administration, beside the All and the Intranet groups.
- An authenticated user can see only himself.
- A guest cannot see any user.

```
<request>
  <type>site</type>
  <type>groups</type>
</request>
```

Figure 19.3. Request example

Response

Each type element produces an XML subtree so the response to the previous request is like this:

```
<info>
  <site>...</site>
  <categories>...</categories>
  <groups>...</groups>
  ...
</info>
```

Figure 19.4. Response example

Here follows the structure of each subtree:

- *site*: This is the container
 - *name*: Human readable site name
 - *siteId*: Universal unique identifier of the site
 - *platform*: This is just a container to hold the site's back end
 - *name*: Platform name. For GeoNetwork installations it must be GeoNetwork.
 - *version*: Platform version, given in the X.Y.Z format
 - *subVersion*: Additional version notes, like 'alpha-1' or 'beta-2'.

Example:

```
<site>
  <name>My site</name>
  <organisation>FAO</organization>
  <siteId>0619cc50-708b-11da-8202-000d9335906e</siteId>
  <platform>
    <name>geonetwork</name>
    <version>2.2.0</version>
  </platform>
</site>
```

Figure 19.5. Example site information

- *categories*: This is the container for categories.

- *category* [0..n]: A single GeoNetwork's category. This element has an id attribute which represents the local identifier for the category. It can be useful to a client to link back to this category.
- *name*: Category's name
- *label*: The localised labels used to show the category on screen. See Figure 19.6, "Example response for categories".

```
<categories>
  <category id="1">
    <name>datasets</name>
    <label>
      <en>Datasets</en>
      <fr>Jeux de données</fr>
    </label>
  </category>
</categories>
```

Figure 19.6. Example response for categories

- *groups*: This is the container for groups
- *group* [2..n]: This is a GeoNetwork group. There are at least the Internet and Intranet groups. This element has an id attribute which represents the local identifier for the group.
 - *name*: Group's name
 - *description*: Group's description
 - *referrer*: The user responsible for this group
 - *email*: The email address to notify when a map is downloaded
 - *label*: The localised labels used to show the group on screen. See Figure 19.7, "Example response for groups".

```
<groups>
  <group id="1">
    <name>editors</name>
    <label>
      <en>Editors</en>
      <fr>Éditeurs</fr>
    </label>
  </group>
</groups>
```

Figure 19.7. Example response for groups

- *operations*: This is the container for the operations
- *operation* [0..n]: This is a possible operation on metadata. This element has an id attribute which represents the local identifier for the operation.
 - *name*: Short name for the operation.
 - *reserved*: Can be y or n and is used to distinguish between system reserved and user defined operations.

- *label*: The localised labels used to show the operation on screen. See Figure 19.8, “Example response for operations”.

```
<operations>
  <operation id="0">
    <name>view</name>
    <label>
      <en>View</en>
      <fr>Voir</fr>
    </label>
  </operation>
</operations>
```

Figure 19.8. Example response for operations

- *regions*: This is the container for geographical regions
- *region [0..n]*: This is a region present into the system. This element has an id attribute which represents the local identifier for the operation.
 - *north*: North coordinate of the bounding box.
 - *south*: South coordinate of the bounding box.
 - *west*: West coordinate of the bounding box.
 - *east*: east coordinate of the bounding box.
 - *label*: The localised labels used to show the region on screen. See Figure 19.9, “Example response for regions”.

```
<regions>
  <region id="303">
    <north>82.99</north>
    <south>26.92</south>
    <west>-37.32</west>
    <east>39.24</east>
    <label>
      <en>Western Europe</en>
      <fr>Western Europe</fr>
    </label>
  </region>
</regions>
```

Figure 19.9. Example response for regions

- *sources*: This is the container.
 - *source [0..n]*: A source known to the remote node.
 - *name*: Source’s name
 - *UUID*: Source’s unique identifier

```

<sources>
  <source>
    <name>My Host</name>
    <UUID>0619cc50-708b-11da-8202-000d9335906e</uuid>
  </source>
</sources>

```

Figure 19.10. Example response for a source

- *users*: This is the container for user information
- *user [0..n]*: A user of the system
 - *id*: The local identifier of the user
 - *username*: The login name
 - *surname*: The user's surname. Used for display purposes.
 - *name*: The user's name. Used for display purposes.
 - *profile*: User's profile, like Administrator, Editor, UserAdmin etc...
 - *address*:
 - *state*:
 - *zip*:
 - *country*:
 - *email*:
 - *organisation*:
 - *kind*:

```

<users>
  <user>
    <id>3</id>
    <username>eddi</username>
    <surname>Smith</surname>
    <name>John</name>
    <profile>Editor</profile>
    <address/>
    <state/>
    <zip/>
    <country/>
    <email/>
    <organisation/>
    <kind>gov</kind>
  </user>
</users>

```

Figure 19.11. Example response for a user

localised entities

localised entities have a general label element which contains the localised strings in all supported languages. This element has as many children as the supported languages. Each child has a name that reflect the language code while its content is the localised text. Here is an example of such elements:

```
<label>
  <en>Editors</en>
  <fr>Éditeurs</fr>
  <es>Editores</es>
</label>
```

xml.forward

This is just a router service. It is used by JavaScript code to connect to a remote host because a JavaScript program cannot access a machine other than its server. For example, it is used by the harvesting web interface to query a remote host and retrieve the list of site ids.

Request

```
<request>
  <site>
    <url>...</url>
    <type>...</type>
    <account>
      <username>...</username>
      <password>...</password>
    </account>
  </site>
  <params>...</params>
</request>
```

Figure 19.12. The service's request

Where:

1. *site*: A container for site information where the request will be forwarded.
2. *url*: Refers to the remote URL to connect to. Usually it points to a GeoNetwork XML service but it can point to any XML service.
3. *type*: Its only purpose is to distinguish GeoNetwork nodes which use a different authentication scheme. The value GeoNetwork refers to these nodes. Any other value, or if the element is missing, refers to a generic node.
4. *account*: This element is optional. If present, the provided credentials will be used to authenticate to the remote site.
5. *params*: This is just a container for the request that must be executed remotely.

```
<request>
  <site>
    <url>http://mynode.org:8080/geonetwork/srv/en/xml.info</url>
  </site>
  <params>
    <request>
      <type>site</type>
    </request>
  </params>
</request>
```

Figure 19.13. Request for info from a remote server

Please note that this service uses the GeoNetwork's proxy configuration.

Response

The response is just the response from the remote service.

19.3 Harvesting services

Introduction

This chapter provides a detailed explanation of the GeoNetwork's harvesting services. These services allow a complete control over the harvesting behaviour. They are used by the web interface and can be used by any other client.

xml.harvesting.get

Retrieves information about one or all configured harvesting nodes.

Request

Called with no parameters returns all nodes. Example:

```
<request/>
```

Otherwise, an id parameter can be specified:

```
<request>
  <id>123</id>
</request>
```

Response

When called with no parameters the service provide its output inside a nodes container. You get as many node elements as are configured. Figure 19.14, "Example of an xml.harvesting.get response for a GeoNetwork node" shows an example of output.

```

<nodes>
  <node id="125" type="geonetwork">
    <site>
      <name>test 1</name>
      <UUID>0619cc50-708b-11da-8202-000d9335aaaa</uuid>
      <host>localhost</host>
      <port>8080</port>
      <servlet>geonetwork</servlet>
      <account>
        <use>false</use>
        <username />
        <password />
      </account>
    </site>
    <searches>
      <search>
        <freeText />
        <title />
        <abstract />
        <keywords />
        <digital>false</digital>
        <hardcopy>false</hardcopy>
        <source>
          <UUID>0619cc50-708b-11da-8202-000d9335906e</uuid>
          <name>Food and Agriculture organisation</name>
        </source>
      </search>
    </searches>
    <options>
      <every>90</every>
      <oneRunOnly>false</oneRunOnly>
      <status>inactive</status>
    </options>
    <info>
      <lastRun />
      <running>false</running>
    </info>
    <groupsCopyPolicy>
      <group name="all" policy="copy"/>
      <group name="mygroup" policy="createAndCopy"/>
    </groupsCopyPolicy>
    <categories>
      <category id="4"/>
    </categories>
  </node>
</nodes>

```

Figure 19.14. Example of an `xml.harvesting.get` response for a GeoNetwork node

If you specify an id, you get a response like that one in Figure 19.15, “Example of an `xml.harvesting.get` response for a WebDAV node” (for a WebDAV node).

```

<node id="165" type="webdav">
  <site>
    <name>test_1</name>
    <UUID>0619cc50-708b-11da-8202-000d9335aaae</uuid>
    <url>http://www.mynode.org/metadata</url>
    <icon>default.gif</icon>
    <account>
      <use>true</use>
      <username>admin</username>
      <password>admin</password>
    </account>
  </site>
  <options>
    <every>90</every>
    <oneRunOnly>false</oneRunOnly>
    <recurse>false</recurse>
    <validate>true</validate>
    <status>inactive</status>
  </options>
  <privileges>
    <group id="0">
      <operation name="view" />
    </group>
    <group id="14">
      <operation name="download" />
    </group>
  </privileges>
  <categories>
    <category id="2"/>
  </categories>
  <info>
    <lastRun />
    <running>false</running>
  </info>
</node>

```

Figure 19.15. Example of an `xml.harvesting.get` response for a WebDAV node

The node's structure has a common XML format, plus some additional information provided by the harvesting types. In the following structure, each element has a cardinality specified using the `[x..y]` notation, where `x` and `y` denote the minimum and the maximum values. The cardinality `[1..1]` is omitted for clarity.

- *node*: The root element. It has a mandatory *id* attribute that represents the internal identifier and a mandatory *type* attribute which indicates the harvesting type.
- *site*: A container for site information.
 - *name (string)*: The node's name used to describe the harvesting.
 - *UUID (string)*: This is a system generated unique identifier associated to the harvesting node. This is used as the source field into the Metadata table to group all metadata from the remote node.
 - *account*: A container for account information.
 - *use (boolean)*: true means that the harvester will use the provided username and password to authenticate itself. The authentication mechanism depends on the harvesting type.
 - *username (string)*: Username on the remote node.
 - *password (string)*: Password on the remote node.

- **options**: A container for generic options.
 - *every (integer)*: Harvesting interval in minutes.
 - *oneRunOnly (boolean)*: After the first run, the entry's status will be set to inactive.
 - *status (string)*: Indicates if the harvesting from this node is stopped (inactive) or if the harvester is waiting for the timeout (active).
- **privileges [0..1]**: A container for privileges that must be associated to the harvested metadata. This optional element is present only if the harvesting type supports it.
 - *group [0..n]*: A container for allowed operations associated to this group. It has the id attribute which value is the identifier of a GeoNetwork group.
 - *operation [0..n]*: Specifies an operation to associate to the containing group. It has a name attribute which value is one of the supported operation names. The only supported operations are: *view*, *dynamic*, *featured*.
- **categories [0..1]**: This is a container for categories to assign to each imported metadata. This optional element is present if the harvesting type supports it.
 - *category (integer) [0..n]*: Represents a local category and the id attribute is its local identifier.
- **info**: A container for general information.
 - *lastRun (string)*: The lastRun element will be filled as soon as the harvester starts harvesting from this entry. The value is the
 - *running (boolean)*: True if the harvester is currently running.
- **error**: This element will be present if the harvester encounters an error during harvesting.
 - *code (string)*: The error code, in string form.
 - *message (string)*: The description of the error.
 - *object (string)*: The object that caused the error (if any). This element can be present or not depending on the case.

Errors

- **ObjectNotFoundEx** If the id parameter is provided but the node cannot be found.

xml.harvesting.add

Create a new harvesting node. The node can be of any type supported by GeoNetwork (GeoNetwork node, web folder etc...). When a new node is created, its status is set to inactive. A call to the `xml.harvesting.start` service is required to start harvesting.

Request

The service requires an XML tree with all information the client wants to add. In the following sections, default values are given in parenthesis (after the parameter's type) and are used when the parameter

is omitted. If no default is provided, the parameter is mandatory. If the type is boolean, only the true and false strings are allowed.

All harvesting nodes share a common XML structure that must be honoured. Please, refer to the previous section for elements explanation. Each node type can add extra information to that structure. The common structure is here described:

- *node*: The root container. The type attribute is mandatory and must be one of the supported harvesting types.
 - *site* [0..1]
 - *name* (*string*, "")
 - *account* [0..1]
 - *use* (*boolean*, 'false')
 - *username* (*string*, "")
 - *password* (*string*, "")
 - *options* [0..1]
 - *every* (*integer*, '90')
 - *oneRunOnly* (*boolean*, 'false')
 - *privileges* [0..1]: Can be omitted but doing so the harvested metadata will not be visible. Please note that privileges are taken into account only if the harvesting type supports them.
 - *group* [0..n]: It must have the *id* attribute which value should be the identifier of a GeoNetwork group. If the id is not a valid group id, all contained operations will be discarded.
 - *operation* [0..n]: It must have a *name* attribute which value must be one of the supported operation names.
 - *categories* [0..1]: Please, note that categories will be assigned to metadata only if the harvesting type supports them.
 - *category* (*integer*) [0..n]: The mandatory id attribute is the category's local identifier.

Please note that even if clients can store empty values ("") for many parameters, before starting the harvesting entry those parameters should be properly set in order to avoid errors.

In the following sections, the XML structures described inherit from this one here so the common elements have been removed for clarity reasons (unless they are containers and contain new children).

Standard GeoNetwork harvesting

To create a node capable of harvesting from another GeoNetwork node, the following XML information should be provided:

- *node*: The type attribute is mandatory and must be GeoNetwork.
 - *site*

- *host* (*string*, "): The GeoNetwork node's host name or IP address.
- *port* (*string*, '80'): The port to connect to.
- *servlet* (*string*, 'geonetwork'): The servlet name chosen in the remote site.
- *searches* [0..1]: A container for search parameters.
 - *search* [0..n]: A container for a single search on a siteID. You can specify 0 or more searches. If no search element is provided, an unconstrained search is performed.
 - *freeText* (*string*, "): Free text to search. This and the following parameters are the same used during normal search using the web interface.
 - *title* (*string*, "): Search the title field.
 - *abstract* (*string*, "): Search the abstract field.
 - *keywords* (*string*, "): Search the keywords fields.
 - *digital* (*boolean*, 'false'): Search for metadata in digital form.
 - *hardcopy* (*boolean*, 'false'): Search for metadata in printed form.
 - *source* (*string*, "): One of the sources present on the remote node.
- *groupsCopyPolicy* [0..1]: Container for copy policies of remote groups. This mechanism is used to retain remote metadata privileges.
 - *group*: There is one copy policy for each remote group. This element must have 2 mandatory attributes: *name* and *policy*. The name attribute is the remote group's name. If the remote group is renamed, it is not found anymore and the copy policy is skipped. The policy attribute represents the policy itself and can be: *copy*, *createAndCopy*, *copyToIntranet*. *copy* means that remote privileges are copied locally if there is locally a group with the same name as the *name* attribute. *createAndCopy* works like *copy* but the group is created locally if it does not exist. *copyToIntranet* works only for the remote group named *all*, which represents the public group. This policy copies privileges of the remote group named *all* to the local Intranet group. This is useful to restrict metadata access.

Figure 19.16, "Example of an xml.harvesting.add request for a GeoNetwork node" shows an example of an XML request to create a GeoNetwork node.

```

<node type="geonetwork">
  <site>
    <name>South Africa</name>
    <host>south.africa.org</host>
    <port>8080</port>
    <servlet>geonetwork</servlet>
    <account>
      <use>true</use>
      <username>admin</username>
      <password>admin</password>
    </account>
  </site>
  <searches>
    <search>
      <freeText />
      <title />
      <abstract />
      <keywords />
      <digital>true</digital>
      <hardcopy>>false</hardcopy>
      <source>0619cc50-708b-11da-8202-000d9335906e</source>
    </search>
  </searches>
  <options>
    <every>90</every>
    <oneRunOnly>>false</oneRunOnly>
  </options>
  <groupsCopyPolicy>
    <group name="all" policy="copy"/>
    <group name="mygroup" policy="createAndCopy"/>
  </groupsCopyPolicy>
  <categories>
    <category id="4"/>
  </categories>
</node>

```

Figure 19.16. Example of an `xml.harvesting.add` request for a GeoNetwork node

WebDAV harvesting

To create a web DAV node, the following XML information should be provided.

- *node*: The type attribute is mandatory and must be WebDAV.
- *site*
 - *url* (*string*, *"*): The URL to harvest from. If provided, must be a valid URL starting with 'HTTP://'.
 - *icon* (*string*, *'default.gif'*): Icon file used to represent this node in the search results. The icon must be present into the images/harvesting folder.
- *options*
 - *recurse* (*boolean*, *'false'*): When true, folders are scanned recursively to find metadata.
 - *validate* (*boolean*, *'false'*): When true, GeoNetwork will validate every metadata against its schema. If the metadata is not valid, it will not be imported.

This type supports both privileges and categories assignment.

Figure 19.17, "Example of an `xml.harvesting.add` request for a WebDAV node" shows an example of an XML request to create a web DAV entry.

```

<node type="webdav">
  <site>
    <name>Asia remote node</name>
    <url>http://www.mynode.org/metadata</url>
    <icon>default.gif</icon>
    <account>
      <use>true</use>
      <username>admin</username>
      <password>admin</password>
    </account>
  </site>
  <options>
    <every>90</every>
    <oneRunOnly>>false</oneRunOnly>
    <recurse>>false</recurse>
    <validate>>true</validate>
  </options>
  <privileges>
    <group id="0">
      <operation name="view" />
    </group>
    <group id="14">
      <operation name="features" />
    </group>
  </privileges>
  <categories>
    <category id="4"/>
  </categories>
</node>

```

Figure 19.17. Example of an `xml.harvesting.add` request for a WebDAV node

CSW harvesting

To create a node to harvest from a CSW capable server, the following XML information should be provided:

- *node*: The type attribute is mandatory and must be `csw`.
- *site*
 - *capabilitiesUrl* (string): URL of the capabilities file that will be used to retrieve the operations address.
 - *icon* (string, 'default.gif'): Icon file used to represent this node in the search results. The icon must be present into the `images/harvesting` folder.
- *searches* [0..1]
 - *search* [0..n]: Contains search parameters. If this element is missing, an unconstrained search will be performed.
 - *freeText* (string, "): Search the entire metadata.
 - *title* (string, "): Search the `dc:title` queryable.
 - *abstract* (string, "): Search the `dc:abstract` queryable.
 - *subject* (string, "): Search the `dc:subject` queryable.

This type supports both privileges and categories assignment.

Figure 19.18, “Example of an `xml.harvesting.add` request for a CSW node” shows an example of an XML request to create a CSW entry.

```
<node type="csw">
  <site>
    <name>Minos CSW server</name>
    <capabilitiesUrl>http://www.minos.org/csw?request=GetCapabilities
      &service=CSW&acceptVersions=2.0.1</capabilitiesUrl>
    <icon>default.gif</icon>
    <account>
      <use>true</use>
      <username>admin</username>
      <password>admin</password>
    </account>
  </site>
  <options>
    <every>90</every>
    <oneRunOnly>>false</oneRunOnly>
    <recurse>>false</recurse>
    <validate>>true</validate>
  </options>
  <privileges>
    <group id="0">
      <operation name="view" />
    </group>
    <group id="14">
      <operation name="features" />
    </group>
  </privileges>
  <categories>
    <category id="4"/>
  </categories>
</node>
```

Figure 19.18. Example of an `xml.harvesting.add` request for a CSW node

Response

The service’s response is the output of the `xml.harvesting.get` service of the newly created node.

Summary

The following table:

Table 19.2. Summary of features of the supported harvesting types

Harvesting type	Authentication	Privileges ?	Categories ?
GeoNetwork	native	through policies	yes
Web DAV	HTTP digest	yes	yes
CSW	HTTP Basic	yes	yes

`xml.harvesting.update`

This service is responsible for changing the node’s parameters. A typical request has a node root element and must include the `id` attribute:

```
<node id="24">
  ...
</node>
```

The body of the node element depends on the node's type. The update policy is this:

- If an element is specified, the associated parameter is updated.
- If an element is not specified, the associated parameter will not be changed.

So, you need to specify only the elements you want to change. However, there are some exceptions:

1. *privileges*: If this element is omitted, privileges will not be changed. If specified, new privileges will replace the old ones.
2. *categories*: Like the previous one.
3. *searches*: Some harvesting types support multiple searches on the same remote note. When supported, the updated behaviour should be like the previous ones.

Note that you cannot change the type of an node once it has been created.

Request

The request is the same as that used to add an entry. Only the id attribute is mandatory.

Response

The response is the same as the `xml.harvesting.get` called on the updated entry.

`xml.harvesting.remove/start/stop/run`

These services are put together because they share a common request interface. Their purpose is obviously to remove, start, stop or run a harvesting node. In detail:

1. *start*: When created, a node is in the inactive state. This operation makes it active, that is the countdown is started and the harvesting will be performed at the timeout.
2. *stop*: Makes a node inactive. Inactive nodes are never harvested.
3. *run*: Just start the harvester now. Used to test the harvesting.

Request

A set of ids to operate on. Example:

```
<request>
  <id>123</id>
  <id>456</id>
  <id>789</id>
</request>
```

If the request is empty, nothing is done.

Response

The same as the request but every id has a status attribute indicating the success or failure of the operation. For example, the response to the previous request could be:

```
<request>
  <id status="ok">123</id>
  <id status="not-found">456</id>
  <id status="inactive">789</id>
</request>
```

Table 19.3, “Summary of status values” summarises, for each service, the possible status values.

Table 19.3. Summary of status values

Status value	remove	start	stop	run
ok	+	+	+	+
not-found	+	+	+	+
inactive	-	-	-	+
already-inactive	-	-	+	-
already-active	-	+	-	-
already-running	-	-	-	+

19.4 System configuration

Introduction

The GeoNetwork’s configuration is made up of a set of parameters that can be changed to accommodate any installation need. These parameters are subdivided into 2 groups:

- parameters that can be easily changed through a web interface.
- parameters not accessible from a web interface and that must be changed when the system is not running.

The first group of parameters can be queried or changed through 2 services: `xml.config.get` and `xml.config.update`. The second group of parameters can be changed using the GAST tool.

`xml.config.get`

This service returns the system configuration’s parameters.

Request

No parameters are needed.

Response

The response is an XML tree similar to the system hierarchy into the settings structure. The response has the following elements:

- *site*: A container for site information.
 - *name*: Site’s name.
 - *organisation*: Site’s organisation name.
- *server*: A container for server information.
 - *host*: Name of the host from which the site is reached.
 - *port*: Port number of the previous host.
- *Intranet*: Information about the Intranet of the organisation.

- *network*: IP address that specifies the network.
- *netmask*: netmask of the network.
- *z3950*: Configuration about Z39.50 protocol.
 - *enable*: true means that the server component is running.
 - *port*: Port number to use to listen for incoming Z39.50 requests.
- *proxy*: Proxy configuration
 - *use*: true means that the proxy is used when connecting to external nodes.
 - *host*: Proxy's server host.
 - *port*: Proxy's server port.
 - *username*: Proxy's credentials.
 - *password*: Proxy's credentials.
- *feedback*: A container for feedback information
 - *email*: Administrator's email address
 - *mailServer*: Email server to use to send feedback
 - *host*: Email's host address
 - *port*: Email's port to use in host address
- *removedMetadata*: A container for removed metadata information
 - *dir*: Folder used to store removed metadata in MEF format
- *ldap*: A container for LDAP parameters
 - *use*:
 - *host*:
 - *port*:
 - *defaultProfile*:
 - *login*:
 - *userDN*:
 - *password*:
 - *distinguishedNames*:
 - *base*:
 - *users*:

- *userAttribs*:
 - *name*:
 - *password*:
 - *profile*:

Figure 19.19, “Example of `xml.config.get` response” shows an example of `xml.config.get` response.

```
<config>
  <site>
    <name>dummy</name>
    <organisation>dummy</organization>
  </site>
  <server>
    <host>localhost</host>
    <port>8080</port>
  </server>
  <Intranet>
    <network>127.0.0.1</network>
    <netmask>255.255.255.0</netmask>
  </intranet>
  <z3950>
    <enable>true</enable>
    <port>2100</port>
  </z3950>
  <proxy>
    <use>false</use>
    <host />
    <port />
    <username>proxyuser</username>
    <password>proxypass</password>
  </proxy>
  <feedback>
    <email />
    <mailServer>
      <host />
      <port>25</port>
    </mailServer>
  </feedback>
  <removedMetadata>
    <dir>WEB-INF/removed</dir>
  </removedMetadata>
  <ldap>
    <use>false</use>
    <host />
    <port />
    <defaultProfile>RegisteredUser</defaultProfile>
    <login>
      <userDN>cn=Manager</userDN>
      <password />
    </login>
    <distinguishedNames>
      <base>dc=fao,dc=org</base>
      <users>ou=people</users>
    </distinguishedNames>
    <userAttribs>
      <name>cn</name>
      <password>userPassword</password>
      <profile>profile</profile>
    </userAttribs>
  </ldap>
</config>
```

Figure 19.19. Example of `xml.config.get` response

xml.config.update

This service is used to update the system's information and so it is restricted to administrators.

Request

The request format must have the same structure returned by the `xml.config.get` service and can contain only elements that the caller wants to be updated. If an element is not included, it will not be updated. However, when included some elements require mandatory information (i.e. the value cannot be empty). Please, refer to Table 19.4, “Mandatory and optional parameters for the `xml.config.update` service”.

Table 19.4. Mandatory and optional parameters for the *xml.config.update* service

Parameter	Type	Mandatory
site/name	string	yes
site/organization	string	-
server/host	string	yes
server/port	integer	-
intranet/network	string	yes
intranet/netmask	string	yes
z3950/enable	bool	yes
z3950/port	integer	-
proxy/use	bool	yes
proxy/host	string	-
proxy/port	integer	-
proxy/username	string	-
proxy/password	string	-
feedback/email	string	-
feedback/mailServer/host	string	-
feedback/mailServer/port	integer	-
removedMetadata/dir	string	yes
ldap/use	bool	yes
ldap/host	string	-
ldap/port	integer	-
ldap/defaultProfile	string	yes
ldap/login/userDN	string	yes
ldap/login/password	string	-
ldap/distinguishedNames/base	string	yes
ldap/distinguishedNames/users	string	yes
ldap/userAttribs/name	string	yes
ldap/userAttribs/password	string	yes
ldap/userAttribs/profile	string	-

Response

On success, the service returns a response element with the OK text. Example:

```
<response>ok</response>
```

Otherwise a proper error element is returned.

19.5 MEF services

Introduction

This chapter describes the services related to the Metadata Exchange Format. These services allow to import/export metadata using the MEF format.

mef.export

As the name suggests, this service exports a GeoNetwork's metadata using the MEF file format.

This service is public but metadata access rules apply. For a partial export, the view privilege is enough but for a full export the download privilege is also required. Without a login step, only partial exports on public metadata are allowed.

This service uses the system's temporary directory to build the MEF file. With full exports of big data maybe it is necessary to change this directory. In this case, use the Java's -D command line option to set the new directory before running GeoNetwork (if you use Jetty, simply change the script into the bin directory).

Request

This service accepts requests in GET/POST and XML form. The input parameters are:

UUID the universal unique identifier of the metadata

format which format to use. Can be one of: *simple*, *partial*, *full*.

skipUuid If provided, tells the exporter to not export the metadata's UUID. Without the UUID (which is a unique key inside the database) the metadata can be imported over and over again. Can be one of: *true*, *false*. The default value is *false*.

Response

The service's response is a MEF file with these characteristics:

- the name of the file is the metadata's UUID
- the extension of the file is *mef*

mef.import

This service is reserved to administrators and is used to import a metadata provided in the MEF format.

Request

The service accepts a *multipart/form-data* POST request with a single *mefFile* parameter that must contain the MEF information.

Response

If all goes well, the service returns an OK element containing the local id of the created metadata. Example:

```
<ok>123</ok>
```

Metadata ownership

Version 1.0 of the MEF format does not take into account the metadata owner (the creator) and the group owner. This implies that this information is not contained into the MEF file. During import, the user that is performing this operation will become the metadata owner and the group owner will be set to null.

19.6 Relations

Introduction

This chapter describes general services used to get and set relations between metadata records inside GeoNetwork. The association is performed by a Relations table which stores a metadata id and a metadata relatedId fields (see Table 19.5, "Structure of table Relations").

Table 19.5. Structure of table Relations

Field	Datatype	Description
id	foreign key to Metadata(id)	Source metadata whose relation is being described.
relatedId	foreign key to Metadata(id)	Metadata related to the source one

xml.relation.get

This service retrieves all relations between metadata.

Request

The request accepts an id and a relation parameters, whose meaning is this:

- *id (integer)*: This is the local GeoNetwork identifier of the metadata whose relations are requested.
- *relation (string, 'normal')*: This optional parameter identifies the kind of relation that the client wants to be returned. It can be one of these values:
 - *normal*: The service performs a query into the id field and returns all relatedId records.
 - *reverse*: The service performs a query into the relatedId field and returns all id records.
 - *full*: Includes both normal and reverse queries (duplicated ids are removed).

Here is an example of POST/XML request:

```
<request>
  <id>10</id>
  <relation>full</relation>
</request>
```

Response

The response has a response root element with several metadata children depending on the relations found. Example:

```
<response>
  <metadata>...</metadata>
  <metadata>...</metadata>
  ...
</response>
```

Each metadata element has the following structure:

- *title*: Metadata title
- *abstract*: A brief explanation of the metadata
- *keyword*: Keywords found inside the metadata
- *image*: Information about thumbnails
- *link*: A link to the source site
- *geoBox*: coordinates of the bounding box
- *geonet:info*: A container for GeoNetwork related information

```
<metadata>
  <title>Globally threatened species of the world</title>
  <abstract> Contains information on animals.</abstract>
  <keyword>biodiversity</keyword>
  <keyword>endangered animal species</keyword>
  <keyword>endangered plant species</keyword>
  <link type="url">http://www.mysite.org</link>
  <geoBox>
    <westBL>-180.0</westBL>
    <eastBL>180.0</eastBL>
    <southBL>-90.0</southBL>
    <northBL>90.0</northBL>
  </geoBox>
  <geonet:info>
    <id>11</id>
    <schema>fgdc-std</schema>
    <createDate>2005-03-31T19:13:31</createDate>
    <changeDate>2007-03-12T14:52:46</changeDate>
    <isTemplate>n</isTemplate>
    <title/>
    <source>38b75c1b-634b-443e-9c36-a12e89b4c866</source>
    <UUID>84b4190b-de43-4bd7-b25f-6ed47eb239ac</uuid>
    <isHarvested>n</isHarvested>
    <view>true</view>
    <admin>>false</admin>
    <edit>>false</edit>
    <notify>>false</notify>
    <download>true</download>
    <dynamic>>false</dynamic>
    <featured>>false</featured>
  </geonet:info>
</metadata>
```

Figure 19.20. Example of a metadata record

19.7 Schema information

Introduction

GeoNetwork is able to handle several metadata schema formats. Up to now, the supported schemas are:

- *ISO-19115 (iso19115)*: GeoNetwork implements an old version of the draft, which uses short names for elements. This is not so standard so this schema is obsolete and will be removed in future releases.
- *ISO-19139 (iso19139)*: This is the XML encoding of the ISO 19115:2007 metadata and ISO 19119 service metadata specifications.
- *Dublin core (dublin-core)*: This is a simple metadata schema based on a set of elements capable of describing any metadata.
- *FGDC (fgdc-std)*: It stands for Federal Geographic Data Committee and it is a metadata schema used in North America.

In parenthesis is indicated the name used by GeoNetwork to refer to that schema. These schemas are handled through their XML schema files (XSD), which GeoNetwork loads and interprets to allow the editor to add and remove elements. Beside its internal use, GeoNetwork provides some useful XML services to find out some element properties, like label, description and so on.

xml.schema.info

This service returns information about a set of schema elements or codelists. The returned information consists of a localised label, a description, conditions that the element must satisfy etc...

Request

Due to its nature, this service accepts only the POST binding with `application/XML` content type. The request can contain several element and codelist elements. Each element indicate the will to retrieve information for that element. Here follows the element descriptions:

- *element*: It must contain a *schema* and a *name* attribute. The first one must be one of the supported schemas (see the section above). The second must be the qualified name of the element which information must be retrieved. The namespace must be declared into this element or into the root element of the request.
- *codelist*: Works like the previous one but returns information about codelists.

```
<request xmlns:gmd="http://www.isotc211.org/2005/gmd">
  <element schema="iso19139" name="gmd:constraintLanguage" />
  <codelist schema="iso19115" name="DateTypCd" />
</request>
```

Note

The returned text is localised depending on the language specified during the service call. A call to `/geonetwork/srv/en/xml.schema.info` will return text in the English language.

Response

The response's root element will be populated with information of the elements/codelists specified into the request. The structure is the following:

- *element*: A container for information about an element. It has a name attribute which contains the qualified name of the element.
 - *label*: The human readable name of the element, localised into the request's language.
 - *description*: A generic description of the element.
 - *condition [0..1]*: This element is optional and indicates if the element must satisfy a condition, like the element is always mandatory or is mandatory if another one is missing.
- *codelist*: A container for information about a codelist. It has a name attribute which contains the qualified name of the codelist.
 - *entry [1..n]*: A container for a codelist entry. There can be many entries.
 - *code*: The entry's code. This is the value that will be present inside the metadata.
 - *label*: This is a human readable name, used to show the entry into the user interface. It is localised.
 - *description*: A generic localised description of the codelist.

```

<response>
  <element name="gmd:constraintLanguage">
    <label>Constraint language</label>
    <description>language used in Application Schema</description>
    <condition>mandatory</condition>
  </element>
  <codelist name="DateTypCd">
    <entry>
      <code>creation</code>
      <label>Creation</label>
      <description>date when the resource was brought into existence</description>
    </entry>
    <entry>
      <code>publication</code>
      <label>Publication</label>
      <description>date when the resource was issued</description>
    </entry>
    <entry>
      <code>revision</code>
      <label>Revision</label>
      <description>date identifies when the resource was examined
        or re-examined and improved or amended</description>
    </entry>
  </codelist>
</response>

```

Error management

Beside the normal exceptions management, the service can encounter some errors trying to retrieve an element/codelist information. In this case, the object is copied verbatim to the response with the addition of an error attribute that describes the encountered error. The returned errors are described in Table 19.6, "Possible errors returned by xml.schema.info service". Here follows an example of such response:

```

<response>
  <element schema="iso19139" name="blablabla" error="not-found"/>
</response>

```

Table 19.6. Possible errors returned by xml.schema.info service

Error code	Description
unknown-schema	The specified schema is not supported
unknown-namespace	The namespace of the specified prefix was not found
not-found	The requested element / codelist was not found

20. Settings hierarchy

20.1 Introduction

GeoNetwork stores many options and information inside the Settings table. Information is grouped into hierarchies where each node has a key/value pair and can have many children. Each key is limited to 32 characters while each value is limited to 250. The 2 top level hierarchies are system and harvesting.

In the following sections, the indentation is used to show hierarchies. Names in bold represent keys with the value's datatype in parenthesis. An *italic* font is used to indicate basic types (string, integer, boolean) while normal font with a | is used to represent a set of allowed values. Regarding the boolean type, value can be only true or false. A missing datatype means that the value of the node is not used. Square brackets indicate cardinality. If they are missing, a cardinality of [1..1] should be considered.

20.2 The system hierarchy

- site: Contains information about the site
 - name (*string*): Name used to present this site to other sites. Used to fill comboboxes or lists.
 - organisation (*string*): Name of the organization/company/institute that is running GeoNetwork
 - siteId (*string*): A UUID that uniquely identifies the site. It is generated by the installer.
- platform: Contains information about the current version
 - version (*string*): GeoNetwork's version in the X.Y.Z format
 - subVersion (*string*): A small description about the version, like 'alpha-1', 'beta' etc...
- server: Used when it is necessary to build absolute URLs to the GeoNetwork server. This is the case, for example, when creating links inside a metadata or when providing CSW capabilities.
 - host (*string*): Main HTTP server's address
 - port (*integer*): Main HTTP server's port (can be empty)
- Intranet: specify the network of the Intranet
 - network (*string*): Network's address
 - netmask (*string*): Network's netmask
- z3950: A container for Z39.50 server parameters
 - enable (*boolean*): If true, GeoNetwork will start the Z30.50 server
 - port (*integer*): The port opened by GeoNetwork to listen to Z39.50 requests. Usually is 2100.
- proxy: This container specify proxy configuration to use
 - use (*boolean*): If true, GeoNetwork will use the given proxy for outgoing connections
 - host (*string*): Proxy's host
 - port (*integer*): Proxy's port

- username (*string*): Proxy's credentials.
- password (*string*): Proxy's credentials.
- feedback: Feedback is sent with proper web form or when downloading a resource.
 - email (*string*): email address of a GeoNetwork administrator or someone else
 - mailServer: This container represents the mail server that will be used to send email
 - host (*string*): Address of the SMTP server to use
 - port (*string*): SMTP port to use
- removedMetadata: This container contains settings about removed metadata.
 - dir: This folder will contain removed metadata in MEF format. It gets populated when the user deletes a metadata using the web interface.
- LDAP: Parameters for LDAP authentication
 - use (*boolean*)
 - host (*string*)
 - port (*integer*)
 - defaultProfile (*string*): Default GeoNetwork's profile to use when the profile user attribute does not exist.
 - login
 - userDN (*string*)
 - password (*string*)
 - distinguishedNames
 - base (*string*)
 - users (*string*)
 - userAttribs: A container for user attributes present into the LDAP directory that must be retrieved and used to create the user in GeoNetwork.
 - name (*string*)
 - password (*string*)
 - profile (*string*)

20.3 Harvesting nodes

The second top level hierarchy is harvesting. All nodes added using the web interface are stored here. Each child has node in its key and its value can be GeoNetwork, WebDAV, CSW or another depending on the node type.

All harvesting nodes share a common setting structure, which is used by the harvesting engine to retrieve these common parameters. This implies that any new harvesting type must honour this structure, which is the following:

- **site**: A container for site information.
 - **name** (*string*): Node name as shown in the harvesting list.
 - **UUID** (*string*): A unique identifier assigned by the system when the harvesting node is created.
 - **useAccount** (*boolean*): Indicates if the harvester has to authenticate to access the data.
 - **username** (*string*):
 - **password** (*string*):
- **options**:
 - **every** (*integer*): Timeout, in minutes, between 2 consecutive harvesting.
 - **oneRunOnly** (*boolean*): If true, the harvester will harvest one time from this node and then it will set the status to inactive.
 - **status** (*active|inactive*): Indicates if the harvesting from this node is stopped (inactive) or if the harvester is waiting until the timeout comes.
- **privileges** [0..1]: This is a container for privileges to assign to each imported metadata
 - **group** (*integer*) [0..n]: Indicate a local group. The node's value is its local identifier. There can be several group nodes each with its set of privileges.
 - **operation** (*integer*) [0..n]: Privilege to assign to the group. The node's value is the numeric id of the operation like 0=view, 1=download, 2=edit etc...
- **categories** [0..1]: This is a container for categories to assign to each imported metadata
 - **category** (*integer*) [0..n]: Indicate a local category and the node's value is its local identifier.
- **info**: Just a container for some information about harvesting from this node.
 - **lastRun** (*string*): If not empty, tells when the harvester harvested from this node. The value is the current time in milliseconds since 1 January, 1970.

Privileges and categories nodes can or cannot be present depending on the harvesting type. In the following structures, this common structure is not shown. Only extra information specific to the harvesting type is described.

Nodes of type GeoNetwork

This is the native harvesting supported by GeoNetwork 2.1 and above.

- **site**: Contains host and account information
 - **host** (*string*)
 - **port** (*integer*)

- servlet (*string*)
- search [0..n]: Contains the search parameters. If this element is missing, an unconstrained search will be performed.
- freeText (*string*)
- title (*string*)
- abstract (*string*)
- keywords (*string*)
- digital (*boolean*)
- hardcopy (*boolean*)
- source (*string*)
- groupsCopyPolicy [0..n]: Represents a copy policy for a remote group. It is used to maintain remote privileges on harvested metadata.
 - name (*string*): Internal name (not localised) of a remote group.
 - policy (*string*): Copy policy. For the group all, policies are: copy, copyToIntranet. For all other groups, policies are: copy, createAndCopy. The Intranet group is not considered.

Nodes of type geonetwork20

This type allows harvesting from older GeoNetwork 2.0.x nodes.

- site: Contains host and account information
 - host (*string*)
 - port (*integer*)
 - servlet (*string*)
- search [0..n]: Contains the search parameters. If this element is missing no harvesting will be performed but the host's parameters will be used to connect to the remote node.
 - freeText (*string*)
 - title (*string*)
 - abstract (*string*)
 - keywords (*string*)
 - digital (*boolean*)
 - hardcopy (*boolean*)
 - siteld (*string*)

Nodes of type WebDAV

This harvesting type is capable of connecting to a web server which is WebDAV enabled.

- Site: Contains the URL to connect to and account information
 - URL (*string*): URL to connect to. Must be well formed, starting with 'http://', 'file://' or a supported protocol.
 - Icon (*string*): This is the icon that will be used as the metadata source's logo. The image is taken from the images/harvesting folder and copied to the images/logos folder.
- options
 - Recurse (*boolean*): Indicates if the remote folder must be recursively scanned for metadata.
 - Validate (*boolean*): If set, the harvester will validate the metadata against its schema and the metadata will be harvested only if it is valid.

Nodes of type CSW

This type of harvesting is capable of querying a Catalogue Services for the Web (CSW) server and retrieving all found metadata.

- site
 - capabUrl (*string*): URL of the capabilities file that will be used to retrieve the operations address.
 - icon (*string*): This is the icon that will be used as the metadata source's logo. The image is taken from the images/harvesting folder and copied to the images/logos folder.
- search [0..n]: Contains search parameters. If this element is missing, an unconstrained search will be performed.
 - freeText (*string*)
 - title (*string*)
 - abstract (*string*)
 - subject (*string*)

Appendix A. Frequently Asked Questions

A.1. Users FAQ

A.1.1.Where do I learn more about the use and functionality of the GeoNetwork opensource catalog?

The Quick Start Guide will provide you with an excellent first introduction. The Guide can be downloaded from the GeoNetwork Community website¹

A.2. Administrators FAQ

A.2.1I want to run GeoNetwork on another port than port 8080. What should I do?

To change the default port 8080, take the following steps:

1. Change the servlet engine configuration. When running on the embedded Jetty, the configuration file to change is bin/jetty.xml
2. Search intermap/WEB-INF/config.xml for 8080 and change it to the new port number
3. Change the map service URLs for the local GeoServer layers that GeoNetwork connects to by default. The file to change is web/intermap/WEB-INF/mapServers.xml
4. On Windows systems, change the startup script bin/win/start-stdin.html
5. Start the GeoNetwork application and login as Administrator. In the System configuration panel you will need to change the Server port number
6. If you changed the port number after uploading data in the system while using another port, you may need to start GAST and run the Metadata Sync in order to update the download links in the metadata.

Other affected, but non-critical files:

- geonetwork/xml/csw/test/csw-Harvest.xml
- geonetwork/xml/schemas/iso19139/process/thumbnails-host-url-relocator.xsl
- geonetwork/xml/validation/csw202_apiso100/csw/2.0.2/examples/wSDL/2.0.2/service.wSDL
- intermap/static/iframepoc.html

A.2.2I am having difficulty installing multiple instances of GeoNetwork on the same server

To run multiple installation you have to change the ports that GeoNetwork uses in order to avoid conflicts. The ports are:

- **Z39.50 listening port.** This is the most probable source of conflicts. You can change the port in the System preferences panel under Administrative tools. Use for example 2101 but keep in mind that remote nodes usually use 2100 so your second node will not be reachable. You cannot use the system configuration web form the first time because if the port conflicts, the server won't start.

- **If you are using Jetty.**
 - **Jetty's listening port.** This can be modified in the `bin/jetty.xml` configuration file. The default value is 8080. To run a second installation use a different value, like 8081.
 - **Jetty's stop port.** This is defined into the scripts `bin/start-geonetwork.*` and `bin/stop-geonetwork.*` (for both Windows and Linux). The provided value is 8079 as the value of the `STOP.PORT` parameter. Use another value for the second installation, like 8078. If you don't change this value, the stop script will stop all instances.
- **If you are using the embedded McKoi DBMS.**
 - **McKoi listening port.** This can be easily modified using GAST. The default value is 9157. For other installations you can use 9158, 9159 and so on. The affected files are `web/WEB-INF/config.xml` and `web/WEB-INF/db/db.conf`.

A.2.3. What is normally logged when running GeoNetwork opensource?

GeoNetwork has its own internal logging based on log4j Logging services² (written by default to the file `jetty/logs/geonetwork.log` and `jetty/logs/intermap.log` for Intermap application). Additionally there are log files generated by the web server (Jetty³, Apache Tomcat⁴ etc..) and by the DBMS used (for example by the internal McKoi SQL⁵).

A.2.4. How do I control what is written to the GeoNetwork internal log file?

The logging is configured in the files `web/geonetwork/WEB-INF/log4j.cfg` and `web/intermap/WEB-INF/log4j.cfg`. You can change the settings by editing the file in a text editor.

The default configuration uses a `DailyRollingFileAppender` that writes the daily log to the file `logs/geonetwork.log` and at midnight is copied to the file `logs/geonetwork.log.yyyy-MM-dd` starting in the file `logs/geonetwork.log` the logging for the new day.

```
log4j.appender.jeeves = org.apache.log4j.DailyRollingFileAppender
log4j.appender.jeeves.DatePattern = 'yyyy-MM-dd
log4j.appender.jeeves.file = logs/geonetwork.log
```

The console logging is also configured to show the datetime, the log intensity of the message (see below), the category and the logging message.

```
log4j.appender.console = org.apache.log4j.ConsoleAppender
log4j.appender.console.layout = org.apache.log4j.PatternLayout
log4j.appender.console.layout.ConversionPattern=%d{ISO8601} %-5p [%c] -
%m%n
```

For operational systems it is suggested to put all log options to OFF or FATAL. The log options are, with increasing log intensity:

- *OFF* - The OFF Level has the highest possible rank and is intended to turn off logging.
- *FATAL* - The FATAL level designates very severe error events that will presumably lead the application to abort.
- *ERROR* - The ERROR level designates error events that might still allow the application to continue running.

- *WARN* - The *WARN* level designates potentially harmful situations.
- *INFO* - The *INFO* level designates informational messages that highlight the progress of the application at coarse-grained level.
- *DEBUG* - The *DEBUG* Level designates fine-grained informational events that are most useful to debug an application.
- *ALL* - The *ALL* Level has the lowest possible rank and is intended to turn on all logging.

A.3. Developers FAQ

A.3.1. What is Free and Open Source Software (FOSS) and how can I use, participate and contribute to the GeoNetwork opensource project?

The book "Producing Open Source Software" (shown in Figure A.1, "Producing Open Source Software") is a highly recommended book for anyone working on open source software projects. It provides insight in all aspects of FOSS development and on how to make a project successful. If you are interested in participating in the GeoNetwork opensource project, please spend some time reading through this book. It's definitely worth the time and money (so buy the hardcopy if you can afford it!).

Producing Open Source Software is a book about the human side of open source development. It describes how successful projects operate, the expectations of users and developers, and the culture of free software.

The book is available in bookstores and from the publisher (O'Reilly Media⁶), or you can browse or download it from <http://producingoss.com/>. Producing Open Source Software is released under an open copyright that allows everyone to share and modify the book freely. The latest version is always available on the website. The online version is the same as the commercially available print version ? in other words, you can buy a printed copy and know that it's up-to-date.



Figure A.1. Producing Open Source Software

Appendix B. Glossary of Metadata Fields Description

This glossary provides you with brief descriptions of the minimum set of metadata fields required to properly describe a geographical data as well as some optional elements highly suggested for a more extensive standard description.

Access constraints. Access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the resource

Abstract. Brief narrative summary of the content of the resource(s)

Administrative area. State, province of the location

* **Temporal extent - Begin date.** Formatted as 2007-09-12T15:00:00 (YYYY-MM-DDTHH:mm:ss)

Character set. Full name of the character coding standard used for the metadata set

Grid spatial representation - Cell geometry. Identification of grid data as point or cell

City. City of the location

Reference System Info - Code. Alphanumeric value identifying an instance in the namespace

Country. Country of the physical address

Data quality info. Provides overall assessment of quality of a resource(s)

Date. Reference date and event used to describe it (YYYY-MM-DD)

Date stamp. Date that the metadata was created (YYYY-MM-DDThh:mm:ss)

Date type. Event used for reference date

Delivery point. Address line for the location (as described in ISO 11180, annex A)

Equivalent scale - Denominator. The number below the line in a vulgar fraction

Data Quality - Description. Description of the event, including related parameters or tolerances

OnLine resource - Description. Detailed text description of what the online resource is/does

Descriptive keywords. Provides category keywords, their type, and reference source

Grid spatial representation - Dimension name. Name of the axis i.e. row, column

Grid spatial representation - Dimension size. Number of elements along the axis

Dimension size Resolution. Number of elements along the axis

Distribution info. Provides information about the distributor of and options for obtaining the resource(s)

Geographic bounding box - East bound longitude. Eastern-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)

Edition. Version of the cited resource

Electronic mail address. Address of the electronic mailbox of the responsible organisation or individual

Temporal extent - End date. Formatted as 2007-09-12T15:00:00 (YYYY-MM-DDTHH:mm:ss)

Equivalent scale. Level of detail expressed as the scale of a comparable hardcopy map or chart

Extent. Information about spatial, vertical, and temporal extent

Facsimile. Telephone number of a facsimile machine for the responsible organisation or individual

File identifier. Unique identifier for this metadata file

Vector spatial representation - Geometric object type. Name of point and vector spatial objects used to locate zero-, one-and two-dimensional spatial locations in the dataset

Vector spatial representation - Geometric object count. Total number of the point or vector object type occurring in the dataset

Geographic bounding box. Geographic position of the dataset

Grid spatial representation. Information about grid spatial objects in the dataset

Grid spatial representation - Resolution value. Degree of detail in the grid dataset

Grid spatial representation - Transformation parameter availability. Indication of whether or not parameters for transformation exists

Data Quality - Hierarchy level. Hierarchical level of the data specified by the scope

Identification info. Basic information about the resource(s) to which the metadata applies

Point of Contact - Individual name. Name of the responsible person- surname, given name, title separated by a delimiter

Keyword. Commonly used word(s) or formalised word(s) or phrase(s) used to describe the subject

Data Language. Language used for documenting data

Metadata - Language. Language used for documenting metadata

Data Quality - Lineage. Non-quantitative quality information about the lineage of the data specified by the scope. Mandatory if report not provided

OnLine resource - Linkage. Location (address) for on-line access using a Uniform Resource Locator address or similar addressing scheme such as <http://www.statkart.no/isotc211>

Maintenance and update frequency. Frequency with which changes and additions are made to the resource after the initial resource is completed

Metadata author. Party responsible for the metadata information

Metadata standard name. Name of the metadata standard (including profile name) used

Metadata standard version. Version (profile) of the metadata standard used

OnLine resource - Name. Name of the online resource

Geographic bounding box - North bound latitude. Northern-most, coordinate of the limit of the dataset extent expressed in latitude in decimal degrees (positive north)

Grid spatial representation - Number of dimensions. Number of independent spatial-temporal axes

Distribution Info - OnLine resource. Information about online sources from which the resource can be obtained

Point of Contact - Organisation name. Name of the responsible organisation

Other constraints. Other restrictions and legal prerequisites for accessing and using the resource

Point of contact. Identification of, and means of communication with, person(s) and organisations(s) associated with the resource(s)

Point of contact - Position name. Role or position of the responsible person

Postal code. ZIP or other postal code

Presentation form. Mode in which the resource is represented

OnLine resource - Protocol. Connection protocol to be used

Purpose. Summary of the intentions with which the resource(s) was developed

Reference system info. Description of the spatial and temporal reference systems used in the datasetData

Data Quality - Report. Quantitative quality information for the data specified by the scope. Mandatory if lineage not provided

Grid spatial representation - Resolution value. Degree of detail in the grid dataset

Point of contact - Role. Function performed by the responsible party

Geographic bounding box - South bound latitude. Southern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north)

Spatial representation info. Digital representation of spatial information in the dataset

Spatial representation type. Method used to spatially represent geographic information

Data Quality - Statement. General explanation of the data producer's knowledge about the lineage of a dataset

Status. Status of the resource(s)

Supplemental Information. Any other descriptive information about the dataset

Temporal Extent. Time period covered by the content of the dataset

Title. Name by which the cited resource is known

Topic category code. High-level geographic data thematic classification to assist in the grouping and search of available geographic data sets. Can be used to group keywords as well. Listed examples

are not exhaustive. NOTE It is understood there are overlaps between general categories and the user is encouraged to select the one most appropriate.

Grid spatial representation - Transformation parameter availability. Indication of whether or not parameters for transformation exists

Vector spatial representation - Topology level. Code which identifies the degree of complexity of the spatial relationships

Type. Subject matter used to group similar keywords

URL. Unified Resource Locator

Use constraints. Constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations or warnings on using the resource

Vector spatial representation. Information about the vector spatial objects in the dataset

Voice. Telephone number by which individuals can speak to the responsible organisation or individual

Geographic bounding box - West bound longitude. Western-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)

Appendix C. ISO Topic Categories

ISO Topic Categories and Keywords

Isotopic Category	Main Topic	Examples	Keywords
<i>Base Maps</i>	Base Maps, Scanned Maps, and Charts		Base Map
<i>Biota</i>	Biologic and Ecologic Information Flora and/or fauna in natural environment	wildlife, vegetation, biological sciences, ecology, wilderness, sea life, wetlands, habitat	
<i>Boundaries</i>	Administrative Legal land descriptions and Political Boundaries	political and administrative boundaries	Administrative boundaries,
<i>Climatology Meteorology Atmosphere</i>	cloud cover, weather, climate, atmospheric conditions, climate change, precipitation		NDWI, Drought, Floods
<i>Earth Cover</i>	Earth Surface Characteristics and Land Cover		Land Cover
<i>Economy</i>	Business and Economic Information, Economic activities, conditions and employment	production, labour, revenue, commerce, industry, tourism and ecotourism, forestry, fisheries, commercial or subsistence hunting, exploration and exploitation of resources such as minerals, oil and gas	
<i>Elevation</i>	Elevation and Derived Products, Height above or below sea level	altitude, bathymetry, digital elevation models, slope, derived products	Digital Elevation Model
<i>Environment</i>	Environmental Monitoring and Modelling, Environmental resources, protection and conservation	environmental pollution, waste storage and treatment, environmental impact assessment, monitoring environmental risk, nature reserves, landscape	
<i>Farming</i>	Agriculture and Farming Rearing of animals and/or cultivation of plants	agriculture, irrigation, aquaculture, plantations, herding, pests and diseases affecting crops and livestock	Agriculture, Crop Production, Livestock
<i>Geoscientific Information</i>	Geologic and Geophysical Information, Information pertaining to earth sciences	geophysical features and processes, geology, minerals, sciences dealing with the composition, structure and origin of the earth's rocks, risks of earthquakes, volcanic activity, landslides, gravity	

		information, soils, permafrost, hydrogeology, erosion	
Health	Human Health and Health, health services, human ecology, and Disease safety	disease and illness, factors affecting health, hygiene, substance abuse, mental and physical health, health services	Malnutrition, Wasting, Stunting, Underweight, Food Deficit, Crop Disease, Livestock Disease,
Imagery	Images and Photographs		
Imagery Base Maps Earth Cover	Base maps	land cover, topographic maps, imagery, unclassified images, annotations	
Inland Waters	Inland Water Resources and Characteristics, Inland water features, drainage systems and their characteristics	rivers and glaciers, salt lakes, water utilisation plans, dams, currents, floods, water quality, <u>hydrographic charts</u>	Rivers,
Intelligence Military	Military bases, structures, activities	barracks, training grounds, military transportation, information collection	
Location	Geodetic Networks Positional information and services and Control Points	addresses, geodetic networks, control points, postal zones and services, place names	Cartography
Oceans	Ocean and Estuarine Resources and Characteristics(excluding inland waters), Features and characteristics of salt water bodies	tides, tidal waves, coastal information, reefs	
Planning Cadastre	Cadastral and Legal Land Descriptions. Information used for appropriate actions for future use of the land	land use maps, zoning maps, cadastral surveys, land ownership	
Recreation	Tourism and Recreation		
Society	Society and Cultural and Demographic Information, Characteristics of society and cultures	Settlements, anthropology, archaeology, education, traditional beliefs, manners and customs, demographic data, recreational areas and activities, social impact assessments, crime and justice, census information	Vulnerability, Early Warning, Emergency, <u>IDPS</u> , Refugees, Population, Poverty, Food Security, Regional <u>Bureaux</u> , <u>WFP</u> Facilities, School feeding
Structure	Facilities, Man-made construction, Buildings and Structures	buildings, museums, churches, factories, housing, monuments, shops, towers	
Transportation	Transportation Means and aids for conveying persons and/or goods, Networks and Models	roads, airports/airstrips, shipping routes, tunnels, nautical charts, vehicle or vessel location, aeronautical	Infrastructure, <u>COMPAS</u> , Food Aid, Food Beneficiaries, Railways, Roads, Shipments

		charts, railways	
<i>Utilities Communication</i>	Utility Distribution Networks, Energy, water and waste systems and communications infrastructure and services	hydroelectricity, geothermal, solar and nuclear sources of energy, water purification and distribution, sewage collection and disposal, electricity and gas distribution, data communication, telecommunication, radio, communication networks	

Appendix D. Free and Open Source Software for Geospatial Information Systems

A range of related software packages can be used in addition to GeoNetwork opensource to deploy a full Spatial Data Infrastructure. These include Web Map Server software, GIS desktop applications and Web Map Viewers.

Below you will find some examples of open source software available for each categories.

D.1 Web Map Server software

- GeoServer (All)¹ <http://www.geoserver.org>
- MapServer (All) <http://www.mapserver.org/>
- MapGuide Open Source (Windows & Linux) <http://www.osgeo.org/mapguide>
- Deegree (All) <http://www.osgeo.org/deegree>

D.2 GIS Desktop software

- GRASS (All) <http://www.osgeo.org/grass>
- gvSIG (All) <http://www.gvsig.gva.es/>
- uDig (All) <http://udig.refractor.net>
- Quantum GIS (All) <http://www.osgeo.org/qgis>
- OSSIM (Windows & OSX) <http://www.osgeo.org/ossim>

D.3 Web Map Viewer and Map Server Management

- OpenLayers (All) <http://www.osgeo.org/openlayers>
- MapBender (All) <http://www.osgeo.org/mapbender>

Glossary

CSW	Catalog Service for the Web. The OGC Catalog Service defines common interfaces to discover, browse, and query metadata about data, services, and other potential resources. See Also OGC.
ISO	International Standards Organisation is an international-standard-setting body composed of representatives from various national standards organizations. http://www.iso.org . See Also ISO TC211.
ISO TC211	ISO/TC 211 is a standard technical committee formed within ISO, tasked with covering the areas of digital geographic information (such as used by geographic information systems) and geomatics. It is responsible for preparation of a series of International Standards and Technical Specifications numbered in the range starting at 19101.
GeoNetwork	GeoNetwork opensource is a standards based, Free and Open Source catalog application to manage spatially referenced resources through the web. http://geonetwork-opensource.org
GN	See GeoNetwork.
XML	Extensible Markup Language is a general-purpose specification for creating custom markup languages.
XSD	XML Schema, published as a W3C recommendation in May 2001, is one of several XML schema languages. http://en.wikipedia.org/wiki/XSD
DB (or DBMS)	A database management system (DBMS) is computer software that manages databases. DBMSes may use any of a variety of database models, such as the network model or relational model. In large systems, a DBMS allows users and other software to store and retrieve data in a structured way.
SOA	Service Oriented Architecture provides methods for systems development and integration where systems package functionality as interoperable services. A SOA infrastructure allows different applications to exchange data with one another.
FGDC	The Federal Geographic Data Committee (FGDC) is an interagency committee that promotes the coordinated development, use, sharing, and dissemination of geospatial data on a national basis in the USA. See http://www.fgdc.gov
SOAP	Simple Object Access Protocol is a protocol specification for exchanging structured information in the implementation of Web Services in computer networks.
OGC	Open Geospatial Consortium. A standards organization for geospatial information systems http://www.opengeospatial.org

OSGeo	The Open Source Geospatial Foundation (OSGeo), is a non-profit non-governmental organization whose mission is to support and promote the collaborative development of open geospatial technologies and data. http://www.osgeo.org
FAO	Food and Agriculture Organisation of the United Nations is a specialised agency of the United Nations that leads international efforts to defeat hunger. http://www.fao.org
WFP	World Food Programme of the United Nations is the food aid branch of the United Nations, and the world's largest humanitarian organization. http://www.wfp.org
UNEP	The UN Environment Programme (UNEP) coordinates United Nations environmental activities, assisting developing countries in implementing environmentally sound policies and encourages sustainable development through sound environmental practices. http://www.unep.org
OCHA	United Nations Office for the Coordination of Humanitarian Affairs is designed to strengthen the UN's response to complex emergencies and natural disasters. http://ochaonline.un.org/
URL	A Uniform Resource Locator specifies where an identified resource is available and the mechanism for retrieving it.
GAST	GeoNetwork Administrator Survival Tool. A desktop application that allows administrators of a GeoNetwork catalog to perform a range of admin operations.
WebDAV	Web-based Distributed Authoring and Versioning. WebDAV is a set of extensions to the Hypertext Transfer Protocol (HTTP) that allows users to edit and manage files collaboratively on remote World Wide Web servers.
OAI-PMH	Open Archive Initiative Protocol for Metadata Harvesting. It is a protocol developed by the Open Archives Initiative. It is used to harvest (or collect) the metadata descriptions of the records in an archive so that services can be built using metadata from many archives.
WMS	Web Map Service is a standard protocol for serving georeferenced map images over the Internet that are generated by a map server using data from a GIS database. The specification was developed and first published by the Open Geospatial Consortium in 1999. See Also OGC.
WFS	Web Feature Service provides an interface allowing requests for geographical features across the web using platform-independent calls. One can think of geographical features as the "source code" behind a map. See Also OGC.

WCS	Web Coverage Service provides an interface allowing requests for geographical coverages across the web using platform-independent calls. The coverages are objects (or images) in a geographical area See Also OGC.
WPS	Web Processing Service is designed to standardize the way that GIS calculations are made available to the Internet. WPS can describe any calculation (i.e. process) including all of its inputs and outputs, and trigger its execution as a Web Service. See Also OGC.
UUID	A Universally Unique Identifier (UUID) is an identifier standard used in software construction, standardized by the Open Software Foundation (OSF) as part of the Distributed Computing Environment (DCE).
MAC address	Media Access Control address (MAC address) is a unique identifier assigned to most network adapters or network interface cards (NICs) by the manufacturer for identification, and used in the Media Access Control protocol sublayer. See also MAC address² on Wikipedia
MEF	Metadata Exchange Format. An export format developed by the GeoNetwork community. More details can be found in this manual in Chapter Metadata Exchange Format.
SKOS	The Simple Knowledge Organisation Systems (SKOS) http://www.w3.org/2004/02/skos/ is an area of work developing specifications and standards to support the use of knowledge organisation systems (KOS) such as thesauri, classification schemes.
Z39.50 protocol	Z39.50 is a client-server protocol for searching and retrieving information from remote computer databases. It is covered by ANSI/NISO standard Z39.50, and ISO standard 23950. The standard's maintenance agency is the Library of Congress.
SMTP	Simple Mail Transfer Protocol is an Internet standard for electronic mail (e-mail) transmission across Internet Protocol (IP) networks.
LDAP	Lightweight Directory Access Protocol is an application protocol for querying and modifying directory services running over TCP/IP.
Shibboleth	The Shibboleth System is a standards based, open source software package for web single sign-on across or within organisational boundaries. It allows sites to make informed authorisation decisions for individual access of protected online resources in a privacy-preserving manner.
DC	The Dublin Core metadata element set is a standard for cross-domain information resource description. It provides a simple and standardised set of conventions for describing things online in ways that make them easier to find.
ESA	European Space Agency is an intergovernmental organisation dedicated to the exploration of space. http://www.esa.int

FOSS Free and Open Source Software, also F/OSS, FOSS, or FLOSS (free/libre/open source software) is software which is liberally licensed to grant the right of users to study, change, and improve its design through the availability of its source code. <http://en.wikipedia.org/wiki/FOSS>

JDBC The Java Database Connectivity (JDBC) API is the industry standard for database-independent connectivity between the Java programming language and a wide range of databases – SQL databases and other tabular data sources, such as spreadsheets or flat files. The JDBC API provides a call-level API for SQL-based database access.

JDBC technology allows you to use the Java programming language to exploit "Write Once, Run Anywhere" capabilities for applications that require access to enterprise data. With a JDBC technology-enabled driver, you can connect all corporate data even in a heterogeneous environment.

Index

A

abstract, 23
access constraints, 24

D

data quality, 24
date, 23
descriptive keywords, 24
distribution info, 24

G

geographic bounding box, 24

L

language, 24
login,

M

maintenance and update frequency, 24
metadata,
Metadata
 Dublin Core, 32
 FGDC, 32
 ISO19115, 32
 ISO19139, 32
 migrating, 33
 profiles, 33
 standards, 32
 transforming, 33
metadata author, 25

O

online resource, 24
other constraints, 24

P

point of contact, 23
presentation form, 23
purpose, 23

R

reference system info, 24

S

scale denominator, 24

security

 access restrictions, 11
 privileges, 11
 roles, 11
 user groups, 11
spatial representation type, 24
status, 23
supplemental information, 24

T

temporal extent, 24
title, 23
topic category, 24

U

user constraints, 24