

---

# **Data Discoverability Guidance**

*Release 0.1*

**Astun Technology**

**Mar 29, 2021**



# CONTENTS:

- 1 Data sharing and metadata creation made easy with GeoNetwork Open Source** **1**
- 1.1 Introduction . . . . . 1
- 1.2 Requirements . . . . . 1
- 1.3 Configuration . . . . . 2
- 1.4 Workflow . . . . . 3
- 1.5 Classification Systems . . . . . 6
- 1.6 Adding Snippets . . . . . 6
- 1.7 Structured Data . . . . . 6
- 1.8 Publishing and sharing your data . . . . . 6
- 1.9 Harvesting . . . . . 6
- 1.10 Search-Engine Optimisation . . . . . 6
  
- 2 Indices and tables** **7**



# DATA SHARING AND METADATA CREATION MADE EASY WITH GEONETWORK OPEN SOURCE

## 1.1 Introduction

This is a guide to extending the GeoNetwork Metadata Catalog to meet best practice guidance on storing and sharing metadata for both spatial and non-spatial datasets, and for improving data discoverability following [best practice guidance from the Geospatial Commission](#).

The objective is to provide guidance on the configuration changes and additional schema plugins that you'll need for full spatial and non-spatial data discoverability, along with a suggested workflow for metadata creators.



It's an ongoing project by Astun Technology, and was funded in part by a grant from the Open Data Institute.



The raw files for this documentation are hosted on GitHub at <https://github.com/AstunTechnology/datadiscoverabilityguidance>, so if you spot a mistake then head over there and let us know. You can also suggest changes, ask questions, or even submit fixes!

## 1.2 Requirements

---

**Important:** To follow this guidance you will need to have an installation of GeoNetwork Open Source, with version 3.10.x. It will not work with earlier versions, and has not been extensively tested yet with GeoNetwork version 4.0.x.

---

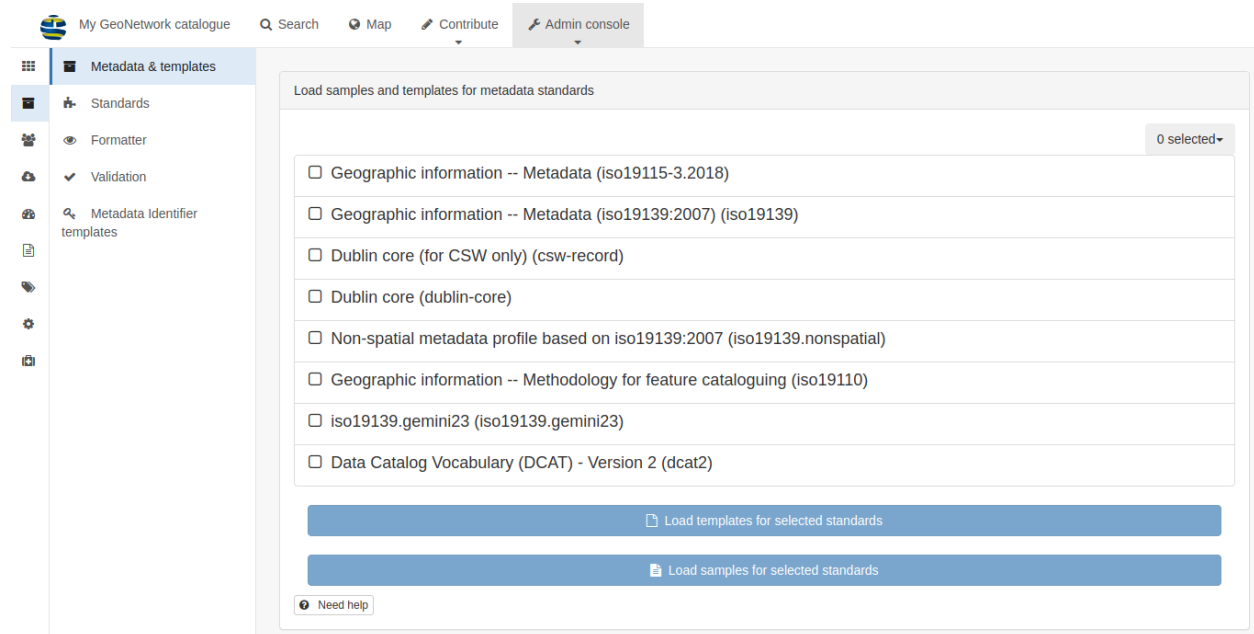
For instructions on installing GeoNetwork, and guidance on getting started, please see the official documentation at <https://www.geonetwork-opensource.org/manuals/trunk/en/install-guide/index.html>.

You will also need to install the plugins for [Gemini 2.3](#) (for spatial data), and [iso19139.nonspatial](#) (for non-spatial data). Additionally, you can install the plugin for metadata in [DCAT-AP v2](#) format. Follow the links to each repository for instructions on how to install. Note that you will need access to the file system on the host computer to do this.

Once you've restarted GeoNetwork, you can check that the metadata profiles have loaded correctly by logging in as an Administrator and going to the admin console -> metadata and templates page.

The list should include the following additional entries (alongside the pre-loaded ones):

- Non-spatial metadata profile based on iso19139:2007
- iso19139.gemini23
- Data Catalog Vocabulary (DCAT) - Version 2



Click the grey dropdown box marked **0 selected** and choose **All**, then click the blue buttons **Load templates for selected standards** and **Load samples for selected standards**.

## 1.3 Configuration

You will need to change a number of settings in the administrator panel to get best use out of GeoNetwork. Login in as an administrator, and visit Admin Console -> Settings.

In the main **System Settings** tab, we recommend making changes to the following sections. Note that there are many other options that you can also change, see [the official documentation](#) for more information:

### Catalog Description

- Fill in a Catalog Name
- Fill in the Organization

### Catalog Server

- Change the **Host**, **Preferred Protocol**, **Port** and **Secure Port** to match your install. For example if you access the catalog at the URL <https://mygeonetwork.com/geonetwork> then you'd set the following:
  - Host: mygeonetwork.com

- Preferred Protocol: https
- Port: blank
- Secure Port: blank

- For the **Timezone**, set the most appropriate one for you. In the UK this is probably *Europe/London*

#### Feedback

- Change the **Email** to the address you want catalog emails to come **from**
- Fill in the address of your mailserver (the **SMTP host**, and set the rest of the options in this section as appropriate.

To test, save your changes and then click the **Test mail configuration** button. This will send an email to the specified address, so make sure it's one you have access to.

**Warning:** If the catalog server is not part of the same domain as the email address, then messages from GeoNetwork may be classified as spam.

#### User feedback

- Click the **Enable feedback** option to allow people to leave comments on records

#### Search statistics

- Click the **Enable** option to store search statistics

#### INSPIRE Directive configuration

- Click the **INSPIRE** option to enable the ability to display records by INSPIRE theme on the home page. Ensure you have the INSPIRE thesaurus installed (see [the page on classification systems](#) for details)
- Click the **INSPIRE search panel** option.

#### Metadata configuration

- Click the **Remove schema location for validation** option. This prevents validation errors from records where the schemalocation is incorrect or cannot be reached. In this case the schema files loaded on your local server are used instead.

---

**Important:** Be sure to click the blue **Save Settings** button to save your changes!

---

## 1.4 Workflow

This is a workflow for using GeoNetwork to meet the Government's [guidance on sharing tabular data](#).

### 1.4.1 Choosing a data format

We recommend using CSV for your non-spatial tabular datasets to meet [Government data sharing guidance](#) but json may be a more suitable format if the data is more complex. See [Government API guidelines](#) for information on good practice for json.

If you are more accustomed to sharing data as an Excel spreadsheet, we would definitely recommend that you convert to CSV as above for data sharing to avoid security risks from macros, or problems arising from Excel's auto-formatting functionality.

### 1.4.2 Formatting your data as a CSV

The Government's guidance on a [tabular data standard](#) recommends that you share non-spatial metadata in CSV format, meeting the following specifications:

- 0 or 1 header rows (preferably 1)
- After the header row, each row should represent a record (eg no blank lines, totals or so on)
- Fields are separated by commas, with text optionally delimited with double quotes
- All rows have the same number of fields
- Line-breaks use windows style “\r\n”
- Use UTF8 for encoding
- No Byte Order Mark (see the link above for more information)

### 1.4.3 Creating a metadata record

Log into GeoNetwork as a user with at least Editor privileges or higher, and go to the Contribute Tab. Choose **Add a new record** and then select **nonGeographicDataset** from the list on the left. Assuming you have followed the [configuration instructions](#) you should be offered the template **Template for metadata in ISO19139 non-spatial format**. Select that by clicking it, then choose the group you wish to create the record in, and finally click the green **+Create** button.



Search Map Contribute Admin console

## Add a new record

- Fill in all the fields shown in the default non-spatial view

### 1.4.4 Uploading your dataset

In your non-spatial record, use the **Associated resources** wizard in the top right and click **+Add**. From the list of options, choose **Link an online resource**.

In the **metadata file store** section to the right, click the green **+Choose or drop resource here** button to navigate to your CSV file. Once it is uploaded select it from the list so that some of the options in the boxes on the left are auto-completed for you.

Fill in a description, and choose **Download** from the list of functions. You can leave the **Application profile** section blank. Finally click the green **Add online resource** button.

**Important:** GeoNetwork will check that the URL to the CSV file is reachable, and will show you an error message at the bottom if it is not. In that case, check the URL is correct.

### 1.4.5 Creating a Feature Catalog record from your dataset

## 1.5 Classification Systems

TODO: This section will outline how to add new thesauri to GeoNetwork to help with UK-specific data sharing etc

## 1.6 Adding Snippets

TODO: This section will explain how to add snippets from <https://github.com/AstunTechnology/geonetwork-snippets>

## 1.7 Structured Data

TODO: This section will explain about structured data embedded in both iso19139.nonspatial and iso19139.gemini23 for data-sharing and SEO

## 1.8 Publishing and sharing your data

TODO: This section will outline the various output formats (for download, and machine-readable)

## 1.9 Harvesting

TODO: This section will outline the options available for harvesting spatial and non-spatial metadata from common ckan endpoints

## 1.10 Search-Engine Optimisation

TODO: This section will outline the SEO functionality built into GeoNetwork and the metadata profiles

## INDICES AND TABLES

- genindex
- modindex
- search