



Report:

Climate Technology Centre & Network (CTCN): National Water and Sewerage Authority, Grenada (NAWASA) – Improvement of water supply management through a GIS based monitoring and control system for water loss reduction in Grenada

Project Output 2: Review of GIS systems and recommendations for future development

Deliverable 2C: Data model for the collection and management of spatial data needed to enhance the resilience of the water supply

1. Introduction

The following report dated the 26th July 2019 was developed by the Wood/GISCAD team as a core deliverable of CTCN project entitled “Improvement of water supply management through a GIS based monitoring and control system for water loss reduction in Grenada”. This project was delivered in collaboration with technical staff within the National Water and Sewerage Authority of Grenada (NAWASA).

The scope of the report relates to the following objectives of Output 2 of the project:

- Defining the core data model for the collection and management of spatial data needed to enhance the resilience of the water supply;

This document outlines the setup of the centralised PostgreSQL/PostGIS databases created for the project and the proposed protocols to follow to assist with their management. The remainder of the document covers the following detailed elements of the data model:

- PostgreSQL database overview – detailing the software products and tools which have been used to define the data model;
- Database schemas - overview of the core data schemas (data groups) defined in the model data model;
- Database roles - overview of the individual (login) and group role defined in the data model;
- Database relationships – outline of the data model relationships between the core clean water datasets held by NAWASA;

- Projections - section providing full details of the two projection systems, which can be defined in the new centralised database. These include the Grenada 1953 / British West Indies Grid projection system or the worldwide WGS84 projection; and
- Metadata management

Details of how the NAWASA GIS users can use (interface and exchange) is covered in accompanying project 2D.

2. PostgreSQL database overview

The centralised databases created for NAWASA was developed using PostgreSQL version 11.4 and the add-in extension PostGIS version 2.5. These products were installed and configured on a new dedicated NAWASA server in June and July 2019.

Access to the databases are available to administrator level users through the PostgreSQL pgAdmin (version 4.8) tool and the front-end GIS package QGIS (version 3.6.3). At the time of writing this document, four databases have been established as shown in Plate 1.

Plate 1 pgAdmin 4 interface used to manage the GIS databases



The scope of the core databases is:

- **nawasa_master**
Master GIS database containing the latest versions of data. Access to this database will be read-only for most GIS users in NAWASA.
- **nawasa_wip**
Work in progress (WIP) database containing temporary / development versions of data. No data held in this database must be used for official analysis of mapping products.
- **nawasa_archive**
Archive database to hold "point in time" versions of specific data. No data held in this database must be used for official analysis of mapping products.
- **nawasa_training**

Database designed specifically for training and testing purposes. No data held in this database must be used for official analysis of mapping products.

In developing this structure, Wood also established a comprehensive database design, with a controlled approach to the assignment of GIS users, roles and data schemas. The different tiers of this model work together to ensure effective management and control of the data held in the databases.

It also important to highlight that only approved and trained administrators with adequate levels of understanding should have access to pgAdmin to directly interface with the PostgreSQL database. All other users will use QGIS to view and/or edit data.

At this stage of the project, several test user accounts have been set up to trial permissions and data interactions within the training database to inform the remainder of this report. This structure will be discussed and agreed with NAWASA prior to the finalisation of this report and accompanying report 2D.

The remainder of this report provides further details of how database schemas, roles and users profiles are used to manage and control data within the database. The report also details the projection requirements for spatial data and how metadata should be managed within the new system.

3. Database schemas

3.1 Overview

PostgreSQL stores individual datasets as tables within separate schemas. Schemas help to group datasets into themes, topics and/or subjects, and are a decisive step in implementing appropriate data management practices. The use of schemas can assist with enabling read/write/admin permissions to user roles by considering what datasets a specific groups/individual requires access to. Further information on how roles are used in the database are provided in section 4.

A series of preliminary schemas were developed for the initial training databases. These are based on key themes identified in the early stages of the project, but the database provides flexibility to be adapted or added to both during the project and if required by NAWASA in the future. Further details of the current schemas are provided below including examples of anticipated datasets which will be stored:

- **schema_basemaps** – this schema will store any base map layers required by NAWASA;
- **schema_customers** – collection of GIS datasets relating to NAWASA’s Customers – including customer meter locations (points); customer meter route areas and summarised water usage data (monthly or annual) linked to individual customer meter locations;
- **schema_dma** – collection of GIS datasets relating to NAWASA’s District Meter Areas – including DMA boundaries and DMA meters;
- **schema_environment** – collection of environmental related GIS datasets. Examples include hydrological data (precipitation and stream flow data), water quality, geology, terrain watersheds, land use and soils;
- **schema_infrasewer** – collection of GIS datasets relating to NAWASA’s core sewerage infrastructure – including sewerage boundaries and sewer zones;
- **schema_infrawater** – collection of GIS datasets relating to NAWASA’s core clean water infrastructure – including clean water pipelines, hydrants and valves, sources (tanks and boreholes); and

- **schema_planning** – collection of GIS datasets relating to planning issues. This schema is currently empty but could include datasets such as locations of affected communities during the dry season (water scarcity) and GPS points of water supply interference (including broken water pipes or scheduled maintenance etc).

3.2 Naming conventions

Naming conventions are a simple and effective method to assist in data management as they provide consistency and clarity on dataset contents, and help reduce naming conflicts. A naming convention should be followed for both schema, and dataset creation. The following rules should be applied in all cases when creating schemas or datasets:

- Titles should be entirely lowercase;
- Titles should not begin with numbers; and
- Titles should contain no spaces or special characters - underscores should be used in place of spaces.

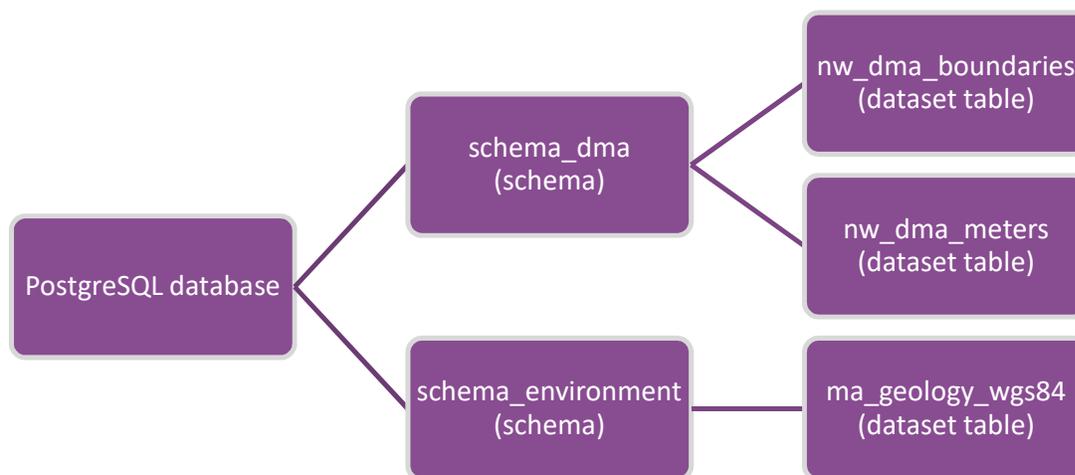
Naming convention specific to datasets:

Alongside the above naming rules, it is advised that a structured approach to the naming of datasets is followed. This is to ensure that NAWASA GIS users and potentially other stakeholders can easily and readily identify the required dataset within the schema. Wood recommends adopting the following naming convention for master versions of all datasets:

`source/dataprovider(2characters)_datasetname_projection(optional)`

The relationship of the individual dataset names within the schema structure is shown in Plate 2.

Plate 2 Database schema naming convention



As shown in Plate 2, the datasets are prefixed with the source provider: as examples **nw** identifies the dataset as coming from NAWASA; **ma** identifies the dataset as coming from the Ministry of Agriculture and Lands. This is limited to 2 characters to avoid overly lengthy dataset names.

The middle section of the dataset name consists of the text description of the dataset contents. This is to enable the GIS user to quickly and efficiently identify what the dataset contains. It can contain additional underscores to help create clarity in the naming, or to further group datasets within the schema, such as seen in the above example separating **dma_boundaries** and **dma_meters**.

The final section of the dataset name includes an optional code to represent if the dataset uses a projection which is not the standard - Grenada 1953 / British West Indies Grid projection system. This include the use of **wgs84** to define any datasets held in the worldwide WGS84 projection. It is recommended that all datasets are held and managed in the Grenada 1953 / British West Indies Grid projection system. This will be the default. Further details of these projection systems are provided in Section 6.

It is important to note that once a dataset has been uploaded to the database and is being used by NAWASA and/or key stakeholders it is recommended that the dataset is not renamed, unless for a specific reason. This will ensure users who have adopted the layers into their GIS projects do not lose connectivity to these datasets. Additionally, other critical systems may require hardcoded names and if a dataset name change is not managed through a structured change request which considers these connections, there may be resultant issues of connectivity failures.

If required, a point in time version of a master dataset can be moved or saved into the "nawasa_archive" database. This might be required to preserve a specific "point in time version" of a dataset for auditing or analysis purposes.

In this case, the dataset naming convention will need to end with a date or version suffix to provide a clear distinction between the latest and previous versions of the dataset.

source/dataprovider(2characters)_datasetname_projection(optional)_date(optional)

The date format is in a reverse order such that 1st July 2019 is written as 190701. This helps the identification of individual layers, especially if multiple, similarly named datasets are created.

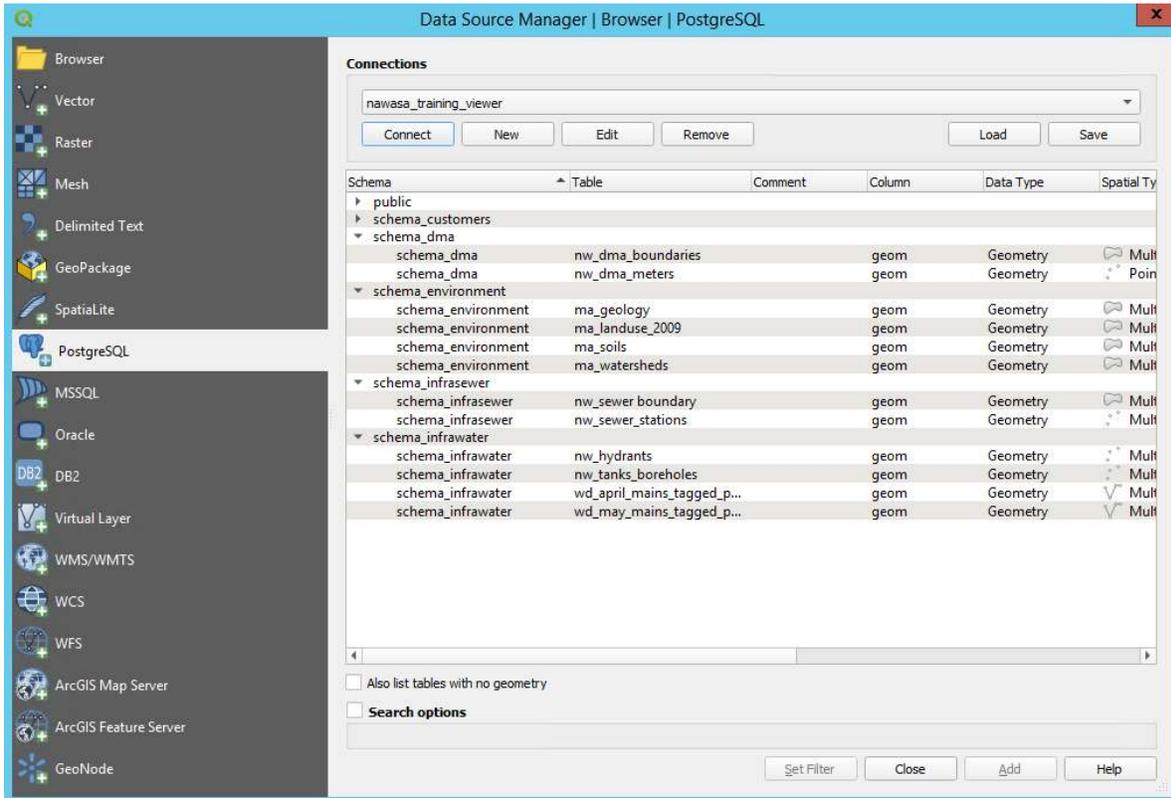
It is recommended that the copying of datasets is avoided as much as possible to limit the duplication of data and increase in the storage volumes held by the database.

4. Database roles

Access to the PostgreSQL database and its subsequent schemas is controlled through roles. Roles can be thought of as either an individual database user (login role), or a group of database users (group role). Roles are granted specific privileges to each schema, allowing for the control of what users can do with the datasets within those schemas.

For example, a role which is limited to only viewing data would have use/select privileges, compared to a role with the ability to edit data having use/select, update, insert and delete privileges. The roles also control how individual users connect to the database through QGIS. This is shown in Plate 3.

Plate 3 Accessing data using the QGIS Data Source Manager



The preferred role set up includes Admin users, Editor users and Viewer users, where the relevant privileges are granted to a group role of the same name i.e. nawasa_admin; nawasa_editor; nawasa_viewer.

On creation of a new login role for an individual user, they can be provided membership to the relevant group role depending on their requirements. This is illustrated in the below in Plate 4.

Plate 4 Relationship between login roles, group roles and database schemas

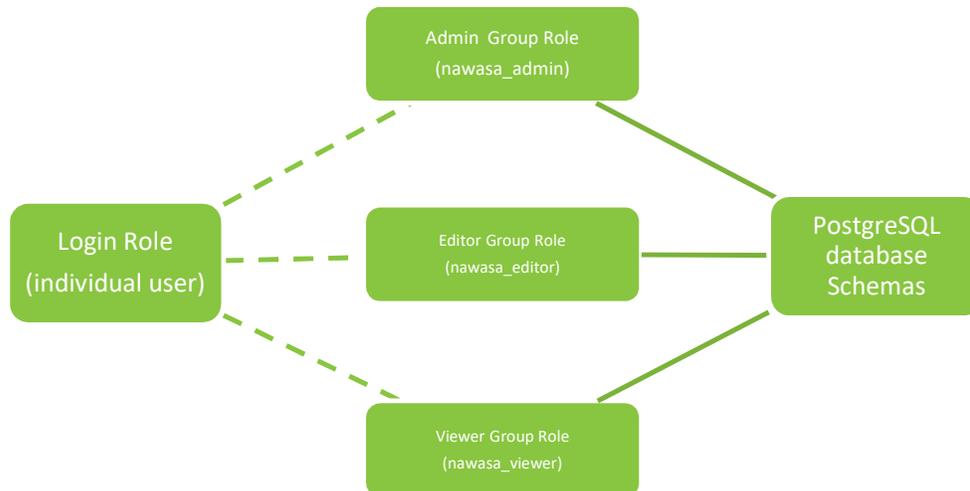


Table 1 outlines the required privileges to be granted to group roles. This is done within the schema options.

Table 1 Security and access privileges for different group roles

	Admin	Editor	Viewer	
Security privileges on schemas	Use	x	X	x
	Create	x		
Default privileges on schemas	All	x		
	Insert		X	
	Select		X	x
	Update		X	
	Delete		X	

An alternative approach, if deemed necessary by NAWASA would be to create admin, editor and viewer group roles for each individual schema. This would enable administrators to control access of individuals by granting different levels of permission to each schema. For example, a user could be granted editor privileges to schema_dma and viewer privileges to all other schemas meaning they were restricted to view only access on all schemas other than schema_dma where they could edit datasets.

This approach is a more complicated data model and is useful where there is requirement useful to restrict access to specific data tables/schemas by different user roles. This might be useful in the future but has not been applied to the new databases at this stage.

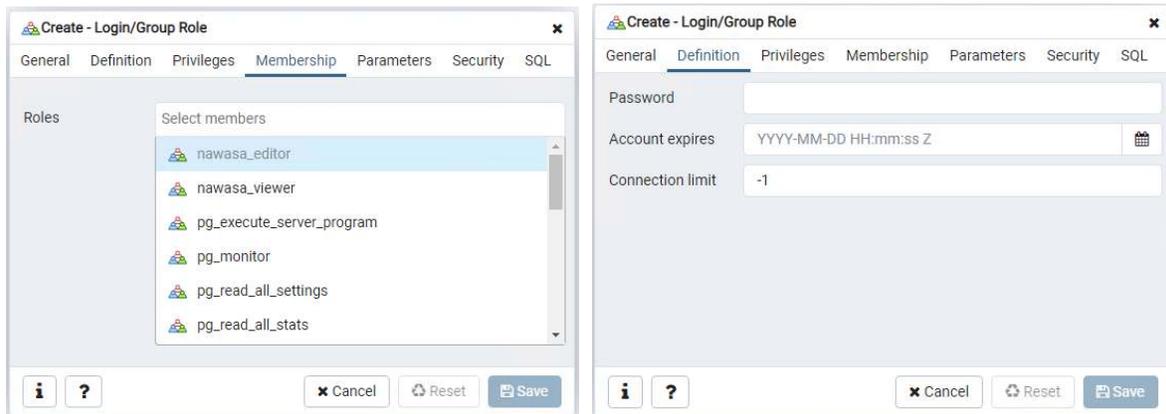
The creation of new roles can be achieved using pgAdmin through the options under the PostgreSQL database tree. This is shown in Plate 5.

Plate 5 Creating new user roles



After creation of the new user, the administrator will be able to assign them to the relevant group role and assign the new user access password. This process is shown below in Plate 6.

Plate 6 Set group role (top) and passwords (bottom) for new users



Administrator level users can also add new users individually or in groups by adding and running SQL code in the pgAdmin Query window. More details can be found on the PostgreSQL official website¹ and will be provided to NAWASA administrators in training sessions.

5. Database relationship model

The development of the new data model has also included the creation of a database relationship model to define the key relationships between NAWASA's core clean water datasets. This structure will be refined during the subsequent stages of the project and ultimately provide a detailed structure for future management of infrastructure data within the organisation. This includes potential extension to the management of sewer infrastructure data which sits outside the scope of the immediate project.

The working data model is presented in Plate 7.

¹ <https://www.postgresql.org/docs/current/sql-createrole.html>

6. Projections

All datasets created and published to the PostgreSQL database must be referenced, with an appropriate projection/coordinate system. The preferred coordinate system for Grenada is the Grenada 1953 / British West Indies Grid projected coordinate system, which is the accepted national grid system. The key parameters of the Grenada 1953 / British West Indies Grid system are provided below in Table 3.

Table 3 Grenada 1953 projection details

Parameter	Value
S_NAME	Grenada 1953 / British West Indies Grid
SRID	2003
AUTH_NAME	EPSG. See 3D CRS for original information source.
Coordinate System Type	Projected
Datum	Grenada 1953
Spheroid	Clarke 1880 (RGS)
Semi-Major Axis	6378249.145
Semi-Minor Axis	6356514.869549776
Prime Meridian	Greenwich
PM Units	Decimal Degree, 0.0174532925199433
Projection	British West Indies Grid (EPSG ID 19942)
Latitude of Origin	0.0
Central Meridian	-62.0
Scale Factor	0.9995
False Easting	400000.0
False Northing	0.0
E/N Units	Meter, 1.0

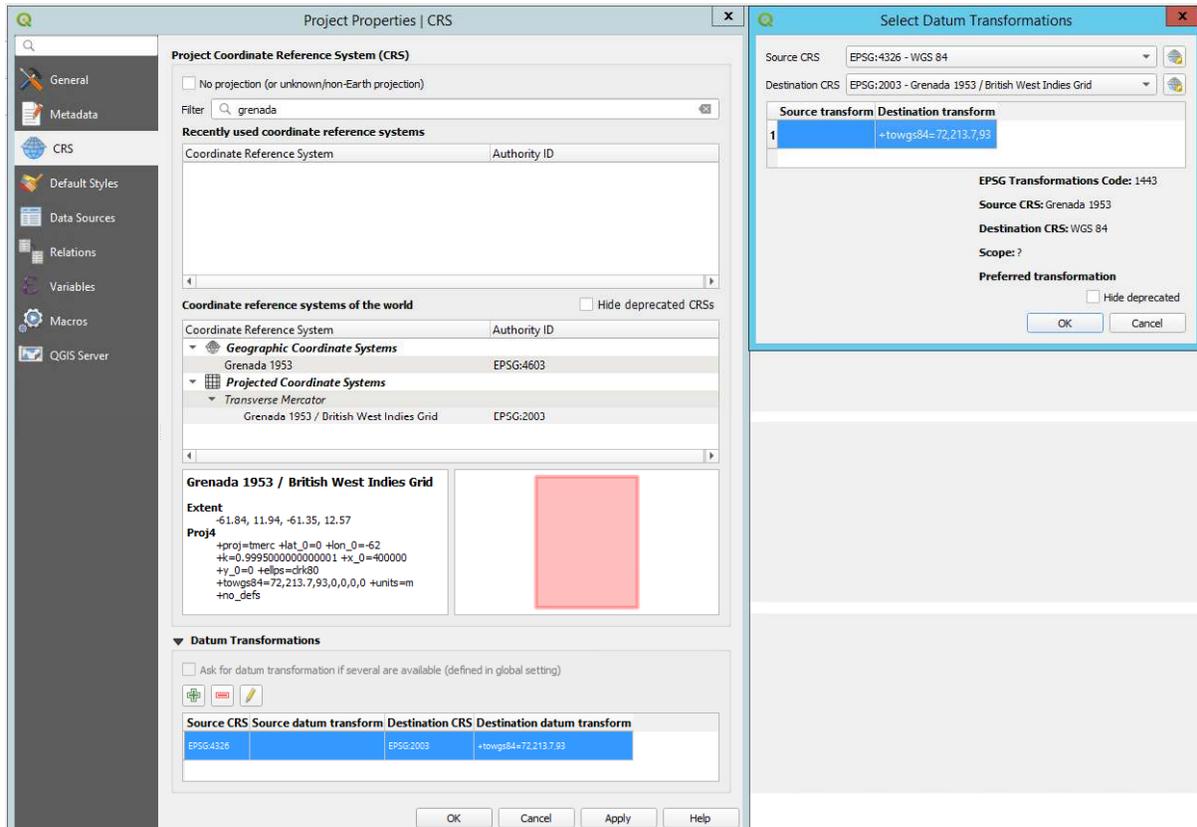
WGS 84 is an acceptable alternative coordinate system for data being provided outside the scope of the Grenada 1953 / British West Indies Grid. WGS 84 is a world coordinate system, the reference coordinate system for Global Positioning System (GPS), with estimated uncertainty of less than 2cm. Table 4 outlines the key parameters of the WGS 84 projection.

Table 4 WGS 84 projection details

Parameter	Value
S_NAME	WGS 84
SRID	4326
AUTH_NAME	EPSG. See 3D CRS for original information source.
Coordinate System Type	Geographic
Datum	World Geodetic System 1984 (EPSG ID 6326)
Spheroid	WGS 84 (EPSG ID 7030)
Semi-Major Axis	6378137
Prime Meridian	Greenwich
PM Units	Decimal Degree, 0.01745329251994328
Semi-Minor Axis	6356752.314245179

Conversion between Grenada 1953 / British West Indies Grid and WGS 84 can be achieved using the transformation "Grenada_1953_To_WGS_84", EPSG transformation code 1443. This can be done within QGIS project properties if a project contains datasets in both coordinate systems – as seen in Plate 8.

Plate 8 Transformation from WGS 84 to Grenada 1953 / British West Indies Grid in QGIS



7. Metadata management

Metadata is an important aspect of data management as it details and records the information on the origin and contents of a dataset. Managing metadata for tables within the PostgreSQL databases can be done through the QGIS plugin "Layer Metadata Search"².

This plugin supports the highly regarded GEMINI 2 metadata standard and is installed on the new NAWASA centralised cloud server. The minimum requirements for metadata for all datasets are:

- **Dataset type** – the primary format of this will be a PostGIS table. Raster files and Web Mapping Services (WMS) can also be included;
- **Title** – the title of the dataset, as defined in Section 3.2;
- **Abstract** – an abstract is the detailed description of what the dataset contains;
- **Layer schema** – the schema the PostGIS table sits within;
- **Layer table** – the name of the PostGIS table itself; and
- **Raster path** – for raster files and WMSs, the path or URL of this data.

Wood propose additional categories be added as custom fields to ensure a robust metadata record, which are:

² <https://github.com/lutraconsulting/qgis-geocat-plugin>

- **Date of last update** – the date that the dataset was last updated on;
- **Custodian organisation** – the source organisation who owns/is responsible for the dataset;
- **Custodian name** – the name of the key contact who owns/is responsible for the dataset; and
- **Custodian email** – the email address of the key contact.

For the Layer Metadata Search tool to work a standalone “metadata” schema is required, this is shown in Plate 1 earlier in this report. This schema houses a single table, “all_metadata”, which contains metadata information for all tables, in all schemas within a database. If a new database has been set up, a metadata schema and metadata table can be added through pgAdmin or SQL code.

The metadata table will need updating as new datasets are added to the database. This can be done directly in pgAdmin or, if the user has the required permissions, through QGIS. As new datasets are added, they will not be dynamically added as a row in the metadata table and instead this will have to be done manually. Alternatively, new datasets can be appended to the metadata table in a batch process through SQL code.

Once the metadata table has been updated with detailed information for all datasets, the Layer Metadata Search plugin can be set up to be used as a data directory. Search terms can be used to filter datasets down and be subsequently added into a QGIS project. To do this, the plugin first needs to be configured as shown in Plate 9.

Plate 9 Layer Metadata Search plugin location



The configuration requires alignment with the required, and additional fields as outlined above and is done within QGIS. Plate 10 illustrates this configuration based on a connection created within our training environment. The custom fields are required to be added to the configuration so that the search functionality can work. It's important to note here that the date_last_update field is widget type “QDateEdit” and the custodian fields are all “QLineEdit (single line)”.

Plate 10 Layer Metadata Search plugin configuration

Layer Metadata Search Configuration

Metadata table

PostGIS connection: nawasa_training_editor

Schema: metadata

Table name: all_metadata

Required metadata table columns

Dataset type column: dataset_type

Vector identifier: PostGIS Table

Raster identifier: Raster File

WMS identifier: WMS Layer

Layer schema column: schema_name

Layer table column: table_name

Raster path column: path

Title column: gem_title

Abstract column: gem_abstract

Custom / additional metadata columns

Description	Column	Widget type
Date of late update	date_last_update	QDateEdit
Owner organisation	custodian_org	QLineEdit (single line)
Owner contact name	custodian_name	QLineEdit (single line)
Owner contact email	custodian_email	QLineEdit (single line)

Advanced

Ignore column: --DISABLED--

Restricted column: --DISABLED--

QGIS PostGIS connection column: --DISABLED--

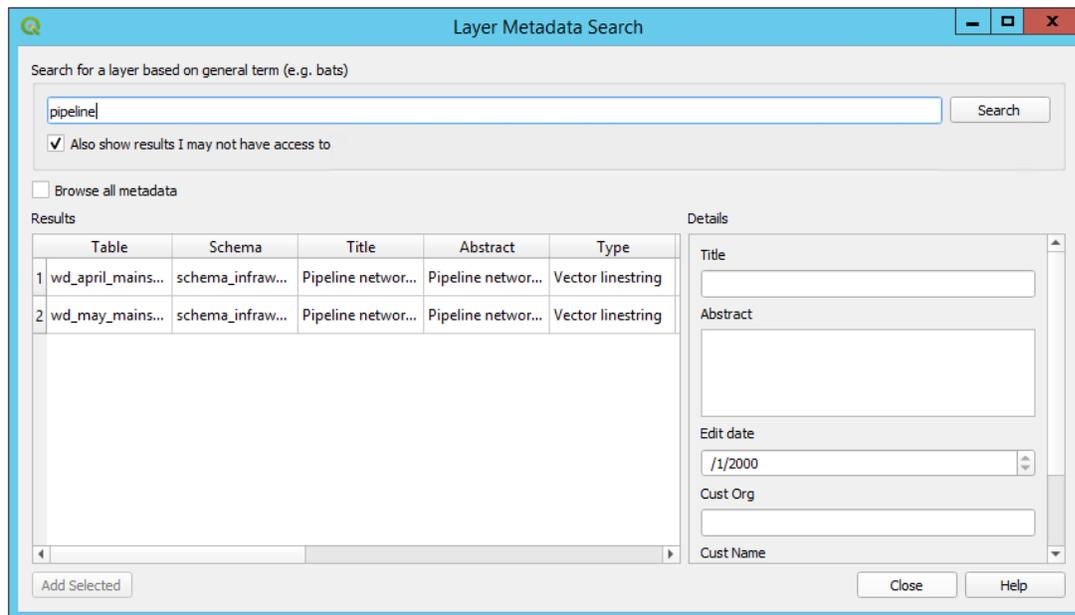
Custom Primary Keys for views:

OK Cancel

Once configuration has been completed, users can use the “Search For Tables Using Metadata” tool within the “Layer Metadata Search” sub-menu to search for datasets using any key words that exist within the metadata table. This enables front end QGIS users to find relevant datasets with ease and add them into their projects without requiring a detailed understanding of the database and schema set up.

Plate 11 illustrates the search functionality of the Layer Metadata Search tool. In this example ‘pipeline’ has been searched for and the results show any dataset which contains this anywhere within its metadata table entry.

Plate 11 Layer Metadata Search, search for table using metadata functionality.



It is recommended that use of the plugin Layer Metadata Search is implemented as the preferred method of searching and adding data into GIS workspaces for all NAWASA GIS users.

Full details for the plugin can be found on the internet website³ and further advice will be provided to NAWASA administrators during future training sessions.

Issued by

Neil Thurston

.....

Approved by

Neil Thurston

.....

Copyright and non-disclosure notice

The contents and layout of this report are subject to copyright owned by Wood (© Wood Environment & Infrastructure Solutions UK Limited 2019) save to the extent that copyright has been legally assigned by us to another party or is used by Wood under licence. To the extent that we own the copyright in this report, it may not be copied or used without our prior written agreement for any purpose other than the purpose indicated in this report. The methodology (if any) contained in this report is provided to you in confidence and must not be disclosed or copied to third parties without the prior written agreement of Wood. Disclosure of that information may constitute an actionable breach of confidence or may otherwise prejudice our commercial interests. Any third party who obtains access to this report by any means will, in any event, be subject to the Third Party Disclaimer set out below.

Third party disclaimer

Any disclosure of this report to a third party is subject to this disclaimer. The report was prepared by Wood at the instruction of, and for use by, our client named on the front of the report. It does not in any way constitute advice to any third party who is able to access it by any means. Wood excludes to the fullest extent lawfully permitted all liability whatsoever for any loss or damage howsoever arising from

³ <https://github.com/lutraconsulting/qgis-geocat-plugin>

reliance on the contents of this report. We do not however exclude our liability (if any) for personal injury or death resulting from our negligence, for fraud or any other matter in relation to which we cannot legally exclude liability.

Management systems

This document has been produced by Wood Environment & Infrastructure Solutions UK Limited in full compliance with our management systems, which have been certified to ISO 9001, ISO 14001 and OHSAS 18001 by LRQA.