



Deliverable 7.4

Final version of system management tools

Authors and affiliation:

*Martin Hansen GEUS
Pavla Kramolišová CGS*

E-mail of lead author:

mh@geus.dk

Version: 19-10-2021

This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation programme under grant agreement number 731166.





Deliverable Data		
Deliverable number	D7.4	
Dissemination level	Public	
Deliverable name	Final version of system management tools	
Work package	WP7, Developments (central)	
Lead WP/Deliverable beneficiary	GeoZS	
Deliverable status		
Submitted (Author(s))	13/10/2021	Martin Hansen
Verified (WP leader)	15/10/2021	Andrej Vihtelič
Approved (Coordinator)	28/10/2021	Jørgen Tulstrup



GENERAL INTRODUCTION

This document gives an overview of the EGDI management tools and the procedures that must be followed to deliver data for which EGDI by now doesn't have management tool. As these management tools has extensive online documentations including recorded webinars and training videos, this report for a large part contains links to these.

TABLE OF CONTENTS

1. DEFINITIONS	4
2. INTRODUCTION	6
3. THEADMINISTRATION MODULE.....	7
4. METADATA CATALOGUE	9
5. DATA TYPES THAT CANNOT BY LOADED THROUGH THE WEB INTERFACES	11
5.1 Grids with multiple values per cell.....	11
5.2 Spreadsheets with formulas	11
5.3 3D geological models	11
5.4 Project vocabularies.....	12
6. MONITORING TOOL.....	13



1. DEFINITIONS

Below there are several definitions, conventions and abbreviations that are used throughout the document. Readers may read this section initially or refer to it when there is a term whose meaning they do not understand, or they want to review.

Definitions are sorted alphabetically to make it easier for users to find a particular definition.

Database: an organized collection of data, generally stored and accessed electronically from a computer system.

Docsify: a tool for generation online documentation. It generates the documentation website on the fly. It is based on Markdown files and loads and parses these files and displays them as a website. <https://docsify.js.org/#/>

Functionality: the range of operations that can be run on a computer or other electronic system.

GeoPackage: Open file format based on SQLite made to store spatial data. A GeoPackage can store multiple spatial tables as well as regular data tables.

JSON: a simple text format for data exchange. It is a subset of the literal object notation of JavaScript, although, because of its wide adoption as an alternative to XML, it is considered a language-independent format.

GeoERA: Establishing the European Geological Surveys Research Area to deliver a Geological Service for Europe.

GeoTIFF: is a public domain metadata standard which allows georeferencing information to be embedded within a TIFF file.

GIP-P: GeoERA Information Platform Project.

GitLab: A opensource DevOps system that provides tools for storage source code with version control, issue tracking and documentations.

GSP: GeoERA Scientific Project. The 14 scientific projects of the GeoERA programme.

Markdown: is a lightweight markup language for creating formatted text using a plain-text editor.

Metadata: data that provides information about spatial and non-spatial data (e.g., purpose of the data, time of creation, authors, etc.).

NetCDF: Open file format made to store grid data with multiple values pr. cell.

Project vocabulary: collections of terms with short descriptions, bibliographic citations and links to unstructured web contents used to define scientific parameters and concepts.

SQLite: An open database format where all tables and functions are stored in a single file.



Tiff: Tag Image File Format a computer file format for storing raster graphics images. Tiff files are compressed but without loss of information's.

URI: A Uniform Resource Identifier (URI) is a unique sequence of characters that identifies a logical or physical resource used by web technologies.

Web application (or web app): application software that runs on a web server, unlike computer-based software programs that are run locally on the operating system (OS) of the device. Web applications are accessed by the user through a web browser with an active network connection.



2. INTRODUCTION

The GeoERA Information Platform Project (GIP-P) is establishing a new common platform for organising, disseminating and sustaining digital harmonised data from the geoscientific GeoERA projects on subsurface energy, water and raw material resources from all over Europe. Great efforts are being put into making the data Findable, Accessible, Interoperable and Reusable (FAIR) in one place and thereby be as valuable as possible for the stakeholders.

The Platform consists of web applications, databases, a digital repository, opensource tools and services that connect all this together.

The new GeoERA Information Platform will become a part of the EuroGeoSurveys (<http://www.eurogeosurveys.org/>) European Geological Data Infrastructure (EGDI) which will care for its future sustainability.

The platform is made to store and display data for different EU projects. To get a user account to deliver data for a specific project, the project lead must make an agreement with the EGDI team.

This document gives an overview of the management tools used in the system. These tools are divided into two groups. The first and most extensive group consists of the tool the data providers use to deliver their data to the EGDI platform. The second group consists of tools used to manage the platform.

The first group of tools (the ones described in this report) are public and intended for the project participants delivering data to the system. This group contains user guides / cookbooks, presentations and recorded webinars and training videos. For some of the data the user cannot deliver the data directly to the platform but has to send the data on predefined formats to EGDI staff who then takes care of loading the data into the system. This applies for grid data with multiple cell values on NetCDF format, spreadsheets with formulas, 3D model and project vocabularies.

The second group is internal and much more technical, only intended for the system developers and system maintainers and as such consists of the systems documented source codes, code snippets and descriptions. This documentation is stored in the EGDI GitLab system. These are only available for people that have been granted access to the system (as some of this information can help hackers to gain access to the system).



3. THE ADMINISTRATION MODULE

The “Administration Module” is the most important part of the process of setting up data sets, delivering data to EGDI and to configure maps and layers.

The Administration module can be used by registered users to upload spatial and unstructured data to the EGDI portal. The module is developed to make it possible for the participants of the different GSPs to upload their spatial data and unstructured data and to set up maps and layers for the Web GIS.

In the user guide the user can find:

- what is needed in order to deliver data to the platform,
- recommendations on how to structure and name the data,
- which format can be uploaded directly to the platform and which must be handled by the EGDI team,
- a section about how to most effectively store and structure the data in GeoPackages and
- how to deliver the data to the platform and set up layers and maps.

The user guide also has a section where we have tried to answer the frequently asked questions.

The user guide is online and can be found [here](#).

The user guide is divided into the following 7 parts:

Introduction	A short introduction to the user guide including links to the webinar to the documentation of the EGDI Metadata Catalogue
Delivering data to EGDI	Description of the considerations that are needed to take before delivering data to EGDI and which effects the different choices has on the functionality and availability of the data
Creating spatial data sets	General description of the way the user should organise the data spatial data, which projection to use and how special fields in your data set should be named
Working with GeoPackages	To deliver more complex spatial data that can fit into a Shapefile GeoPackages are recommended. This section describes through examples how data can be organised in GeoPackages



Uploading data to EGDI	Description of how to upload data and documents to EGDI. It also describes how to thematize the data and organise the different data on maps
Used terms and	Description of the terms used in the user guide
Frequent asked questions	Some frequently asked questions along with their answers

The guide is written using docsify and based on Markdown files stored in the EGDI GitLab.

A training video has also been produced on how to upload unstructured data and add metadata. This video can be found [here](#).



4. METADATA CATALOGUE

A very important part of delivering data is the metadata. When uploading spatial data to EGDI the user is required to add a link to the metadata describing the dataset in the EGDI Metadata catalogue.

The documentation of the EGDI Metadata Catalogue, is, like the one for the Administration, an online user guide. The guide is divided into four parts.

Cookbook for creating metadata records using the EGDI Metadata Catalogue (MIcKA, version 6.0)	User Guide of the EGDI Metadata Catalogue application - general navigation, metadata input, harvesting, minimum required for establishing metadata record, detailed instructions for filling the EGDI metadata profile elements in the EGDI-Lite editing form for a spatial dataset (including extensions for 3D models), service and application
EGDI Metadata profile	Methodology for the unified metadata description of the results of GeoERA projects within the EGDI with the extension to describe 3D geological models
FAQ, frequently asked questions for metadata of GeoERA projects in the EGDI Metadata Catalogue	11 frequently asked questions on these topics: 1. Metadata URL for uploading data, 2. Inability to save metadata record, 3. Metadata record not visible to all users, 4. Search metadata by Project Name, 5. Upload metadata from XML file, 6. Use of example records, 7. Unique resource identifier as URI, 8. Adding Web Map URL to the dataset metadata as an on-line Resource Locator, 9. Linking of dataset metadata - Source Citation and Parent Identifier elements, 10. Metadata of a set of spatial products for one area, 11. Validation of 3D models extended by vertical items compared to INSPIRE validation

The guide is written using docsify and based on Markdown files stored in the [Czech Geological Surveys GitHub site](#).

Along with the written documentations a set of four videos has been produced describing the systems. These videos describe:

Introduction to the EGDI Metadata Catalogue	General navigation in the EGDI Metadata Catalogue, description of main menu, tools, login, and example of a creation of basic metadata with 5+1 mandatory
---	---



	minimum elements for uploading data to the EGDI platform
EGDI Metadata Cookbook Lite	Step by step through the Cookbook for creating metadata records using the EGDI Lite editing form in the EGDI Metadata Catalogue
Training video 1. Metadata of structured data – basic	How to use the EGDI Metadata Catalogue and how to create metadata records compliant with INSPIRE via the EGDI-Lite editing form, following INSPIRE validation panel step by step
Training video 2. Metadata of structured data – detailed	How to use the EGDI Metadata Catalogue. The best practise of creation and editing metadata records via the EGDI-Lite editing form follow the INSPIRE directive validation. Example of uploading spatial data to the EGDI platform. (Following Training video 1. Metadata of structured data – basic)



5. DATA TYPES THAT CANNOT BY LOADED THROUGH THE WEB INTERFACES

When the GIP Project was planned, we did a survey among the GSPs about which kind of data they would collect. The Administration module was developed to be able to handle the mostly used formats. At the same time, it was decided to **not** develop functionality to load the less used formats through the Administration module but load this manually by GIP-P staff. The following is an explanation of these formats and how they will be treated.

5.1 Grids with multiple values per cell

A few projects have gridded data with multiple values per cell. It was decided that these should be delivered on the NetCDF format and loaded manually into the EGDI system by GIP-P staff.

5.2 Spreadsheets with formulas

EGDI has a document repository where users can deliver documents, pictures and data. From the beginning it was decided only to accept a set of open formats as the goal of the repository is to store the documents on open formats that will be readable in the future. During upload these documents will receive a static URL (a URI) that will remain constant in time and can be used to refer to the document.

A few of the project have developed spreadsheet with built in formulas that cannot be stored on an open format. These files cannot be uploaded through the Administration module and must be uploaded by hand by GIP staff.

5.3 3D geological models

With the lack of a common format for 3D models and the range of different types of models a common tool to upload all kinds of models was not possible within the project. Instead, a list of formats capable of handling the different types of models (layer-, voxel- and truly 3D models) was established.

- Raster file formats
 - *.asc
 - *.grd
- Tin file formats
 - *.obj
 - *.ts
- Voxel file formats
 - *.xyz

The models must then be exported on one of the formats and sent to a GIP-P person for upload by hand.



5.4 Project vocabularies

Data for project vocabularies are handled by spreadsheets delivered to a GIP-P person for loading into the triple store database. A zip-file with an empty spreadsheet and instructions on how to fill out the spreadsheet can be downloaded from this [GitHub page](#).



6. MONITORING TOOL

If some of the layers don't work, we established a web page from where the status of the different services delivering data to EGDl can be seen. The different services are displayed together with one or more smileys going from green and happy to red and sad depending on the presence of metadata and how the service replies.

If a user clicks on the name on service or the status smiley, he/she is led to a page showing the history of the services behaviour.

If the user clicks on the metadata smiley, he/she is forwarded to the metadata in the EGDl Metadata Catalogue.



Service	Metadata	Status	Open in map
Onshore Geology			
Lithology			
Friuli area (eENVplus) 1:100,000	😊	😊	<input type="checkbox"/>
West Slovenia (eENVplus) 1:100,000	😊	😊	<input type="checkbox"/>
Pan-European (EGDI) 1:1,000,000	😊	😊	<input type="checkbox"/>
Pan-European (IGME5000) 1:5,000,000	😊	😊	<input type="checkbox"/>
Pan-European faults (IGME5000) 1:5,000,000	😊	😊	<input type="checkbox"/>
Age			
Friuli area (eENVplus) 1:100,000	😊	😞	<input type="checkbox"/>
West Slovenia (eENVplus) 1:100,000	😊	😊	<input type="checkbox"/>
Pan-European (EGDI) 1:1,000,000	😊	😊	<input type="checkbox"/>
Marine Geology			
Marine Lithology			
Pan-European (IGME5000) 1:5,000,000	😊	😊	<input type="checkbox"/>
Pan-European faults (IGME5000) 1:5,000,000	😊	😊	<input type="checkbox"/>
Sea-floor Quaternary geology (EMODnet)	😊	😊	<input type="checkbox"/>
Sea-floor Pre-Quaternary bedrock geology (EMODnet)	😊	😊	<input type="checkbox"/>
Seabed substrates (EMODnet) 1:100 000	😊	😊	<input type="checkbox"/>
Seabed substrates (EMODnet) 1:250 000	😊	😊	<input type="checkbox"/>
Seabed substrates (EMODnet) 1:1,000,000	😊	😊	<input type="checkbox"/>
Marine Age			
Sea-floor Quaternary geology (EMODnet)	😊	😊	<input type="checkbox"/>
Sea-floor Pre-Quaternary bedrock geology (EMODnet)	😊	😊	<input type="checkbox"/>
Other topics			
Coastal migration (EMODnet)	😊	😊	<input type="checkbox"/>
Sea-floor geomorphology (EMODnet)	😊	😊	<input type="checkbox"/>

This monitoring / smiley page can be found [here](#).