Guidance for the management and use of geospatial data and technologies in health

Part 2 - Implementing the geospatial data management cycle: 2.5 Cleaning, validating, and documenting the data - 2.5.1 Documenting the data using a metadata profile

Version 1.3 (last update: 19.02.2024)

Revision history

Version	Release date	Comment	By
1.0	20 June 2018	Document created	Steeve Ebener, Izay Pantanilla, Chris Erwin G. Mercado, Richard J. Maude
1.1	04 March 2020	Adjustment of the terminology associated to the concepts of geographical features, geographic objects and master lists to align with other volumes of the series	Steeve Ebener
1.2	10 January 2022	Inclusion of the reference to the new HGL guidance document	Izay Pantanilla
1.3	19 February 2024	Update the layout and the URLs across the document	Steeve Ebener

Authors

Steeve Ebener[1]
Izay Pantanilla[1]
Chris Erwin G. Mercado[2]
Richard J. Maude[2,3,4]

1. Health GeoLab, Manila, Philippines
2. Mahidol-Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand
3. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK
4. Harvard TH Chan School of Public Health, Harvard University, Boston, USA

Acknowledgements

Table of Contents

Purpose and audience

The purpose of the Health GeoLab series of guidance is to inform concerned practitioners about the key elements they need to be aware of when it comes to managing and using geospatial data and technologies in public health and guide them through the processes to be followed in that regard.

The audience for this guidance includes geospatial data managers, technical advisors, and any other practitioners that are directly or indirectly involved in the collection and use of geospatial data and technologies in public health.

Please note that some of the sections in the present guidance require a basic understanding of concepts pertaining to the management and use of geospatial data and technologies.

Abbreviations

ADB	Asian Development Bank
AeHIN	Asia eHealth Information Network
FGDC	Federal Geographic Data Committee
GIS	Geographic Information System
HGL	Health GeoLab
HIS	Health Information System
ISO	International Organization for Standardization
MORU	Mahidol-Oxford Tropical Medicine Research Unit
SDG	Sustainable Development Goal
WHO	World Health Organization

1. Background

The Health GeoLab (HGL) is a regional resource supporting low- and middle-income countries in Asia and the Pacific for them to fully benefit from the power of geography, geospatial data, and technologies to reach the health-related Sustainable Development Goal of healthy lives and well-being for all (SDG 3)[1].

The HGL uses the HIS geo-enabling framework to strengthen in-country capacity. The present document has been developed as part of this approach and with the objective of being used by the largest number of users possible.

This volume is part of a series of guidance started under the umbrella of the AeHIN GIS Lab and now continued by the HGL. The complete series is organized as follows:
- Part 1 - Introduction to the data-information-knowledge-decision continuum and the geospatial data management cycle [1]
- Part 2 - Implementing the geospatial data management cycle:
  - 2.1 Documenting the process and defining the data needs [2]
  - 2.2 Defining the terminology, data specifications, and the ground reference [3]
  - 2.3 Compiling existing data and identifying gaps [4]
  - 2.4 Creating geospatial data
    - 2.4.1 Extracting vector format geospatial data from basemaps [5]
    - 2.4.2 Collecting data in the field [6]
  - 2.5 Cleaning, validating, and documenting the data
    - 2.5.1 Documenting the data using a metadata profile (the present document)
    - 2.5.2 Using advanced Microsoft Excel functions [7]
  - 2.6 Distributing, using, and updating the data
    - 2.6.1 Creating good thematic maps using desktop GIS software [8]
    - 2.6.2 Using thematic maps for decision making [9]
    - 2.6.3 Developing and implementing the appropriate data policy [10]

This guidance is a living document made to evolve based on the inputs received from the users. Please don't hesitate to contact us if you have any suggestions for improvement.

The terms used in the present guidance are defined in the following glossary of terms maintained by the Health GeoLab: https://bit.ly/3ctoHiS

Please also contact us using the same email address should you use this document as part of your activities and would like to have your institution recognized as one of the document's users.

---

[1] https://www.un.org/sustainabledevelopment/health/

2. Introduction

Documenting the data is an important and integral part of the data management cycle [1] and should be applied to the two elements that compose geographic data[2]: geospatial and statistical data.

Such documentation should at least allow answering the following questions about the data in order for the user to gauge if it is appropriate in relation to its anticipated use:

1. What is the data about?
2. Who created the data?
3. When was the data created/collected/last updated?
4. How was the data created?
5. What are the data specifications (geographic coordinate system/projection system, scale, accuracy, language..)?
6. Are there any use or redistribution restrictions attached to the data?
7. Who can I contact if I have questions about the data?

This information, or metadata (information that describes the content, quality, condition, origin, and other characteristics of data or other pieces of information[3]) is being captured in what we call a metadata record which is itself based on a metadata profile (modification of an existing metadata standard to adapt to data issues, cultural issues, or both[4]).

A metadata standard is a requirement which is intended to establish a common understanding of the meaning or semantics of the data, and ensure correct and proper use and interpretation of the data by its owners and users.[5]

In other words, a metadata record should accompany any geospatial or statistical dataset being shared to ensure it is used appropriately.

To facilitate the creation of the metadata record, it is recommended to capture the necessary information included in the metadata profile from the beginning of the data production/collection process and to address issues pertaining to use and redistribution rights long before its dissemination.

The following sections describe the content and structure of the minimum metadata profile recommended by the HGL for geospatial data, georeferenced master lists and statistical data and to describe the process to follow to use this profile to create a metadata record in a GIS software (ArcGIS or QGIS) for the former and in Microsoft Excel for the other two.

---

[2] Information describing the location and attributes of things, including their shapes and representation. Geographic data is the composite of spatial data and attribute data.
[3] http://wiki.gis.com/wiki/index.php/GIS_Glossary/M (Metadata profile)
[4] http://wiki.gis.com/wiki/index.php/GIS_Glossary/M (Metadata)
[5] https://en.wikipedia.org/wiki/Metadata_standard

3. Metadata standards and recommended profile for geospatial data

The following sections present the metadata standards being used most widely nowadays to document geospatial data before describing the minimum metadata profile recommended by the HGL. The process to follow in order to fill such a profile in ArcGIS 10.5 or QGIS 3.0 is then detailed.

3.1 Existing metadata standards for geospatial data

There are four (4) main groups of metadata standards existing to document geospatial data, namely the[6]:
- International Organization for Standardization (ISO)
- Federal Geographic Data Committee (FGDC)
- Dublin Core Metadata Initiative (DCMI)
- Infrastructure for Spatial Information in Europe (INSPIRE)

The first two in this list are the most widely used and the ISO is actually the originator of four (4) standards related to geospatial metadata:
- ISO 19115: Metadata standard defining the geographic information of the data and including its identification, extent, quality, spatio-temporal resolution, spatial reference, and distribution information. There have been a few revisions to this standard and the latest edition is the ISO 19115-1.
- ISO 19110: defines the methodology for cataloguing feature types by classifying, organizing, and creating catalogues of the feature type information (roads, rivers, measurements, etc.)
- ISO 19119: metadata standard for geospatial services, such as web map application services, geoportal architectures, data models, and online data processing services
- ISO 19139: defines the XML schema generated using the format and content information relating to ISO 19115

The FGDC is the originator of the Content Standard for Digital Geospatial Metadata (CSDGM), more commonly known as the FGDC metadata standard.

There is not much difference between the ISO and FGDC standards but because it has been created through an international consensus (FGDC's consensus is limited to U.S. Federal agencies[7]), ISO's standard better supports data sharing across national and cultural boundaries. In addition, the FGDC now endorses the transition of metadata standards to ISO metadata standards[8].

3.2 Recommended minimum metadata profile for geospatial data

In view of the previous section, the minimum metadata profile recommended by the HGL is based on the ISO 19115 metadata standard.

The profile in question is presented in Annex 1 and allows covering the key questions listed in the introduction of the present guidance.

---

[6] http://learningzone.rspsoc.org.uk/index.php/Learning-Materials/Introduction-to-OGC-Standards/4.2-Geospatial-metadata-standards
[7] https://www.fgdc.gov/resources/factsheets/documents/GeospatialMetadata-July2011.pdf
[8] http://learningzone.rspsoc.org.uk/index.php/Learning-Materials/Introduction-to-OGC-Standards/4.2-Geospatial-metadata-standards

3.3 Filling the recommended minimum metadata profile in GIS software (ArcGIS 10.5 and QGIS 3.0)

In ArcGIS, the recommended minimum metadata profile can be created using ArcCatalog. ArcCatalog provides several metadata style options to support different metadata standards – one of which is ISO 19139, the XML version of the ISO 19115 standard. Annex 2 shows the steps to follow to create the recommended minimum metadata profile in ArcCatalog for ArcGIS 10.5 to obtain a metadata record based on the ISO 19139 style sheet. Annex 3 presents the resulting metadata record based on the example included in Annex 1.

In QGIS, it has been possible to create and save a metadata record since version 3.0. The metadata profile provided is nevertheless not based on the ISO 191115 standard and does not cover all the fields from the recommended minimum metadata profile (Annex 1).

This being said, it is possible to store most of the information for the fields that are missing in the abstract section of the metadata profile.

Annex 4 provides the steps to follow to complete the profile according to the example reported in Annex 1 while Annex 5 provides the resulting metadata record for this example.

4. Metadata standards and recommended profile for georeferenced master list and statistical data

4.1 Existing metadata standards for statistical data

As mentioned above, georeferenced master lists and statistical data should also be documented properly. Statistical data are widely used and it is easy to get confused when handling multiple data coming from different sources. It is therefore important to create metadata for them. The same applies to georeferenced master lists when sharing them.

Different metadata standards and models have been developed to facilitate data communication between organizations and information systems to improve the quality of statistical documentation provided to users[9].

Among these, we can mention those generated through the Statistical Data and Metadata Exchange (SDMX) initiative[10] which has been established to set standards that can facilitate the exchange of statistical data and metadata using modern information technologies.

The SDMX initiative is the originator of several standards, some of which have ended as ISO standards like ISO 17369:2013 which provides an integrated approach to facilitating Statistical Data and Metadata Exchange (SDMX) enabling interoperable implementations within and between systems concerned with the exchange, reporting and dissemination of statistical data and related metadata.

These standards tend to nevertheless be difficult to digest by non-IT professionals and focus more on how the data and metadata are being exchanged, using the eXtensible Markup Language (XML) for example, than on the fields that the metadata record should contain.

---

[9] http://www.ihsn.org/documentation-standards
[10] http://www.sdmx.org

4.2 Recommended minimum data catalogue and metadata profile for georeferenced master list and statistical data

The purpose of the present guidance is to provide users with an easy and effective way to document the georeferenced master lists and statistical data being generated, managed, shared and used as part of the geospatial data management cycle.

Two components are key in this regard:
- A data catalogue (simplified version of the data dictionary) describing the content of each of the data element contained in the georeferenced master list and statistical dataset.
- A metadata record providing the essential information allowing the user to answer the questions reported in the introduction of the present guidance.

Tables 1 and 2 provide an example of a data catalogue and metadata record for a dataset containing health facility level statistical data.

Data element code	Data element description
HF_ID	Unique identifier of the health facility
HF_NAME	Official name of the health facility (English)
HF_TYPE	Health facility type as per the official classification
NBR_DOC_15	Number of doctors working in the health facility in 2015
NBR_NUR_15	Number of nurses working in the health facility in 2015
NBR_BED_15	Number of operational beds in the health facility in 2015

Table 1. Example of data catalogue for a health facility level statistical dataset

Title:	Hospitals capacity for the Tolkien Region (2015)
Originator:	Department of Health (DOH)
Collection date:	06.06.2015
Abstract:	This database contains the 2015 health capacity related statistics for the hospital of the Tolkien Region
Process:	This database has been created by extracting data from the DOH Health Information System
Progress:	Complete
Access constraints:	The access to this data is limited to your institution and should in no way been shared with other parties
Use constraints:	The use of this data is limited to the purpose granted to your institution. Please contact the DOH again for any other use.
Disclaimer:	This dataset is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the data lies with the user. In no event shall the DOH be liable for damages arising from its use.
Primary Contact	
Contact Name	Ariela Antonio
Organization	Department of Health (DOH)
Contact Telephone number:	877-0000
Contact Fax Number:	NA
Contact Email Address:	a.antonio@doh.gov

Table 2. Example of metadata record for the health facility statistical dataset reported in Table 1

The metadata record for a georeferenced master list would itself contain one more field to capture the temporal validity of the master list. Table 3 provide an example of metadata for a master list.

Title:	Health Facility Master List for Tolkien Province
Originator:	Department of Health (DOH)
Publication date:	January 2025
Temporal validity:	January 2025
Abstract:	This dataset is the official master list of health facilities for Tolkien Province
Process:	This dataset has been generated by combining different sources of information that are validated by the DOH. The geographic coordinates present different level of accuracy depending on the data collection method used.
Progress:	Ongoing (updated regularly)
Access constraints:	The access to this data is limited to the participants attending the Introduction to geospatial data management and technologies for Malaria Programs training workshop.
Use constraints:	The use of this data is limited to the participants attending the Introduction to geospatial data management and technologies for Malaria Programs training workshop.
Disclaimer:	This dataset is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the data lies with the user. In no event shall the Department of Health be liable for damages arising from its use.
Primary Contact	
Contact Name	Ariela Antonio
Organization	DOH
Contact Telephone number:	877-0000
Contact Fax Number:	NA
Contact Email Address:	a.antonio@doh.gov

Table 3. Example of metadata record for a georeferenced health facility master list

The data catalogue contains only two columns, one with the labels of the fields included in the master list or statistical data (Field column) and one containing a description of the content for each of these fields (Description column). This could nevertheless be expanded to include the description in local language.

The metadata record contains the following fields:
- Title: the name by which the georeferenced master list or statistical dataset is known.
- Originator: complete name of the institution which generated the georeferenced master list or statistical dataset.
- Collection date: date when the data was collected (statistical data)
- Publication date: date when the master list has been published
- Temporal validity: Temporal validity of the master list (date or period)
- Abstract: short description of the content of the georeferenced master list or statistical dataset.
- Process: general description of the steps that were followed to create the georeferenced master list or statistical dataset.
- Progress: development status of the georeferenced master list or statistical dataset.
- Access constraints: access constraints attached to the georeferenced master list or statistical dataset.
- Use constraints: use constraints attached to the georeferenced master list or statistical dataset.
- Disclaimer: liability statement for the georeferenced master list or statistical dataset.
- Primary contact: full name, organization, position and role, phone number, and email address of the person to contact regarding the georeferenced master list or statistical dataset.

4.3 Filling the recommended minimum data dictionary and metadata profile for georeferenced master lists statistical data in Microsoft Excel

Microsoft Excel file is the recommended format here because: 1) it is widely used in countries and 2) it allows storage of data and information on multiple spreadsheets/'Sheets' in the same file which is not the case with other file formats such as .csv files. This allows the dataset, the data dictionary, and the metadata record to be saved together in a single file therefore ensuring that none of this information is lost when shared.

Including the data dictionary and the metadata record with the dataset in an Excel file only requires the addition of two extra worksheets.

The content of the data dictionary and metadata worksheet is then generated based on the fields reported in Tables 1, 2 or 3.

References

[1]   Ebener S. (2016): Guidance for the management and use of geospatial data and technologies in health. Part 1 - Introduction to the data-information-knowledge-decision continuum and the geospatial data management chain. Health GeoLab document: https://www.healthgeolab.net/DOCUMENTS/Guide_HGL_Part1.pdf [Accessed 19 February 2024]

[2]   Ebener S. (2016): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.1 Documenting the process and defining the data needs. Health GeoLab document: https://www.healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_1.pdf [Accessed 19 February 2024]

[3]   Ebener S. (2016): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.2 Defining the terminology, data specifications, and the ground reference. Health GeoLab document: https://www.healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_2.pdf [Accessed 19 February 2024]

[4]   Pantanilla I., Ebener S., Mercado C.E., Maude R. (2018): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.3 Compiling existing data and identifying gaps. Health GeoLab document: https://www.healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_3.pdf [Accessed 19 February 2024]

[5]   Ebener S. (2021): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.4 Creating geospatial data - 2.4.1 Extracting vector format geospatial data from basemaps. Health GeoLab document: https://healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_4_1.pdf [Accessed 19 February 2024]

[6]   Ebener S., Maude R.J., Gault P. (2018): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.4 Creating geospatial data - 2.4.2 Collecting data in the field. Health GeoLab document: https://www.healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_4_2.pdf [Accessed 19 February 2024]

[7]   Ebener S., Pantanilla I. (2019): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.5 Cleaning, validating, and documenting the data - 2.5.2 Using advanced Microsoft Excel functions. Health GeoLab document: https://healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_5_2.pdf [Accessed 19 February 2024]

[8]   Pantanilla I., Ebener S., Maude R. (2018): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.6 Distributing, using, and updating the data. 2.6.1 Creating good thematic maps using desktop GIS software. Health GeoLab document:

https://www.healthgeolab.net/DOCUMENTS/Guide_HGL_Part2_6_1.pdf [Accessed 19 February 2024]

[9]   Ebener S. (under preparation): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.6 Distributing, using, and updating the data - 2.6.2 Using thematic maps for decision making. Health GeoLab document.

[10]   Ebener S. (under preparation): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.6 Distributing, using, and updating the data - 2.6.3 Developing and implementing the appropriate data policy. Health GeoLab document.

Annex 1 – Recommended minimum metadata profile for geospatial data

Question	Expected information	Example	Correspond metadata field in ArcGIS 10.5		Correspond metadata field in QGIS 3.0	
			Section	**Field**	**Tab**	**Field**
What is the data about?	The name by which the geospatial data layer is known	Government health facilities of Region VIII , Philippines (2015)	Overview>Item Description	Title	Identification	Title
	Short abstract describing, the content of the geospatial data layer, its date of creation, and the source	This geospatial data layer has been developed by the Department of Health of the Philippines (DOH) and contains the location of all the government health facilities in Region VIII as observed in 2015	Overview>Item Description	Description (Abstract)	Identification	Abstract
	Reason for which the geospatial data layer has been created	This geospatial data layer has been developed to serve as the official reference for governmental health facilities in Region VIII	Overview>Item Description	Summary (Purpose)	Identification	Abstract
	Generic keyword associated to the geospatial data layer	Health	Overview>Topics&Keywords	Topic Categories	Categories	ISO Categories
	Keywords describing the area covered by the geospatial data layer	Region VIII, Philippines	Overview>Topics&Keywords	Place keywords	Keywords	Concept
Who created the data?	Complete name of the Institution who generated the geospatial data layer	Department of Health of the Philippines (DOH)	Overview>Item Description	Credit	Identification	Parent Identifier
When was the data created/collected/last updated?	Date at which the geospatial data layer has been created/collected	01-05-15	Overview>Citation	Dates (created)	Extent	Temporal extent for this dataset
	Development status of the geospatial data layer	Completed	Resource>Details	Status	Identification	Abstract
	Frequency at which the geospatial data layer is being updated	Continual	Resource>maintenance	Update frequency	Identification	Abstract
How was the data created?	General description of the steps that have been followed to create the geospatial data layer	This geospatial data layer has been generated by compiling different source of geographic coordinates from within and outside the DOH. These coordinates have then been checked before linking them to the information stored in the National health Facility Registry	Resource>Lineage	Lineage	History	History

| Question	Expected information	Example	Correspond metadata field in ArcGIS 10.5		Correspond metadata field in QGIS 3.0	
			Section	Field	Tab	Field
What are the data specifications?	Extent of the area covered by the geospatial data layer expressed according to its unit (decimal degrees for example)	West: 123.8; East: 125.3; South: 5.75; North: 7.5	Overview>Item description	Bounding box	Extent	Extent
	Language(s) in which the information is stored in the attribute table of the geospatial data layer	English	Resource>Details	Languages	Identification	Language
	Denominator of the scale at which the geospatial data layer has been generated	100,000 (meaning 1:100,000 scale)	Resource>Details	Scale Resolution	Identification	Abstract
	Format of the geospatial data layer	Vector	Resource>Details	Spatial representation type	Identification	Type
	Code, code space and version of the spatial reference system (geographic and projected coordinate system) of the geospatial data layer as per http://spatialreference.org/	EPSG 4326 (WGS 84)	Resource>Spatial Reference	Reference system	Extent	Coordinate Reference System
	Scope that will allow for the system to define the quality information that should be captured in the metadata	Dataset	Resource>Quality	Scope level	Identification	Type
	Type of quality report that is being provided about the geospatial data layer	Absolute External Positional Accuracy	Resource>Quality	Report >Report type	Identification	Abstract
	Identifies the axe (horizontal or vertical) to which the spatial quality information applies	horizontal	Resource>Quality	Report>Dimension	Identification	Abstract
	A description of the data quality measure reported for the geospatial data layer	The maximum horizontal error accepted at the scale of work is of 50 meters	Resource>Quality	Report>Measure	Identification	Abstract
	Title and date of publication of the document containing the description of the method used to perform the quality check	DOH guidelines on geospatial data; January 2014	Resource>Quality	Specification (titles, date)	Identification	Abstract
	A description of the evaluation method that has been used to assess the data quality	The maximum horizontal error has been measured for 75% of the points in the dataset using satellite images as ground reference	Resource>Quality	Report>Evaluation method	Identification	Abstract

Question	Expected information	Example	Correspond metadata field in ArcGIS 10.5		Correspond metadata field in QGIS 3.0	
			Section	Field	Tab	Field
What are the data specifications?	Result of the quality assessment that has been conducted	Passed. 82% of the points in the datasets are finding themselves within less than 50 meter from the corresponding location in the satellite image	Resource>Quality	Report>Conformance result	Identification	Abstract
	Name and version of the format in which the geospatial data layer in being distributed	Shape file, ArcGIS 10.5	Resource>Distribution	Distribution format (name and version)	Identification	Type
	Description of each field contained in the attribute table of the geospatial data layer when in vector format	Field: HF_ID; Definition: DOH code of the health facility; Definition source: DOH; Range domain: 1-9999999	Resource>Fields	Attributes (Definition, definition source and domain)	Identification	Abstract
Are there any use or redistribution restrictions attached to the data?	Description of any limitation in the use of the geospatial data layer by third parties	The use of this geospatial data layer is limited to non-commercial use. It is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the data lies with the user. In no event shall the DOH be liable for damages arising from its use.	Overview>Item Description	Use limitation	Access	Licenses
	Access or use constraints attached to the geospatial data layer	Access Constraints: Copyright; Use constraints: Restricted	Resource>Constraints	Legal Constrains	Access	Constraints
	Security level assigned to the geospatial data layer based on the sensitivity or secrecy of the information to be selected from the available options	Unclassified	Resource>Constraints	Security Constraints	Access	Constraints
Who can I contact if I have questions about the data?	Full name, organization, position and role, phone number and email address of the person to contact regarding the geospatial data layer	Antoine Marello; Department of Health of the Philippines (DOH); Data manager; Point of contact; +63301274438; AntoineMarello@gmail.com	Resource>Points of Contact	Contacts	Contact	Name, Organization, Position, Role, Email, Voice
	Full name, organization, position and role, phone number and email address of the person who created the metadata	Antoine Marello; Department of Health of the Philippines (DOH); Data manager; Point of contact; +63301274438; AntoineMarello@gmail.com	Metadata>Contacts	Contact	Identification	Abstract
	Language in which the metadata is filled	English	Metadata>Details	Language	Identification	Abstract

Annex 2 – Filling the recommended minimum metadata profile for geospatial data in ArcCatalog for ArcGIS 10.5

After opening ArcCatalog, the first thing you need to do is to activate the ISO 19139 Metadata Style (XML version of the ISO 19115 standard**).**

This is done by doing the following:

1. Click on *Customize > ArcCatalog Options…* item from the main menu bar
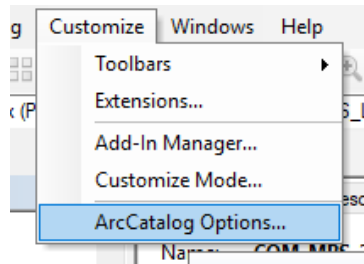
2. In the window that opens, select the "Metadata" tab
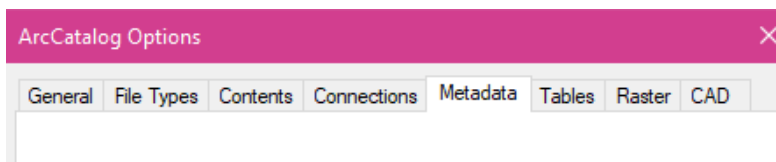
3. Open the scroll-down menu under "Metadata Style" and select *"ISO 19139 Metadata Implementation Specification"* and keep the "Automatically update when metadata is viewed" option checked
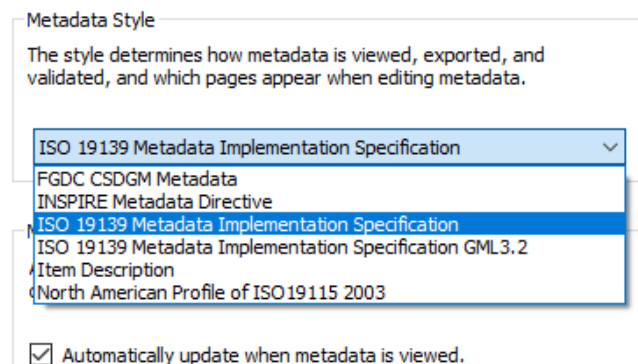
You can now start attaching a metadata record to any dataset you have on your computer.

The following steps are to be followed in order for this metadata record to comply with the recommended minimum metadata profile reported in Annex 1:

1. Launch ArcCatalog and browse to the folder where the file you want to document is located in the "Catalog Tree" (METADATA folder in the present example):
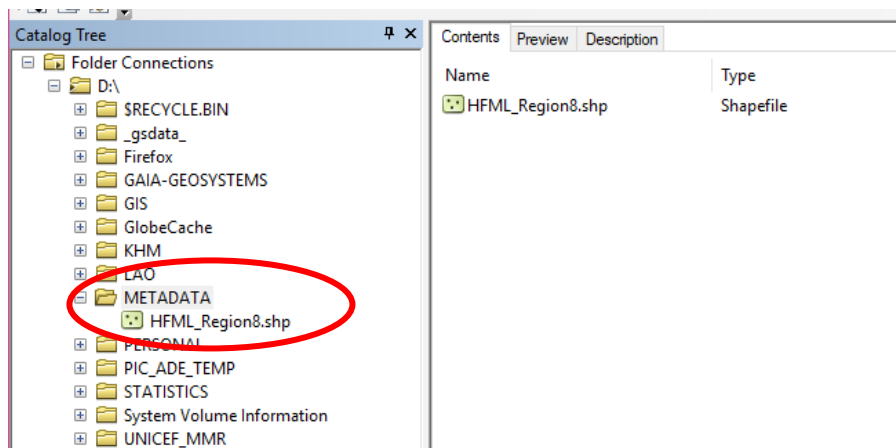
2. Click on the layer you want to document in the Catalog Tree ("HMFL_Region8.shp" in the example here) and then on the "Description" tab appearing on the top of the right side window [Description]

3. The right side window will now display the content of the metadata record attached to this file according to the ISO 19115 standard (mostly empty at this stage):
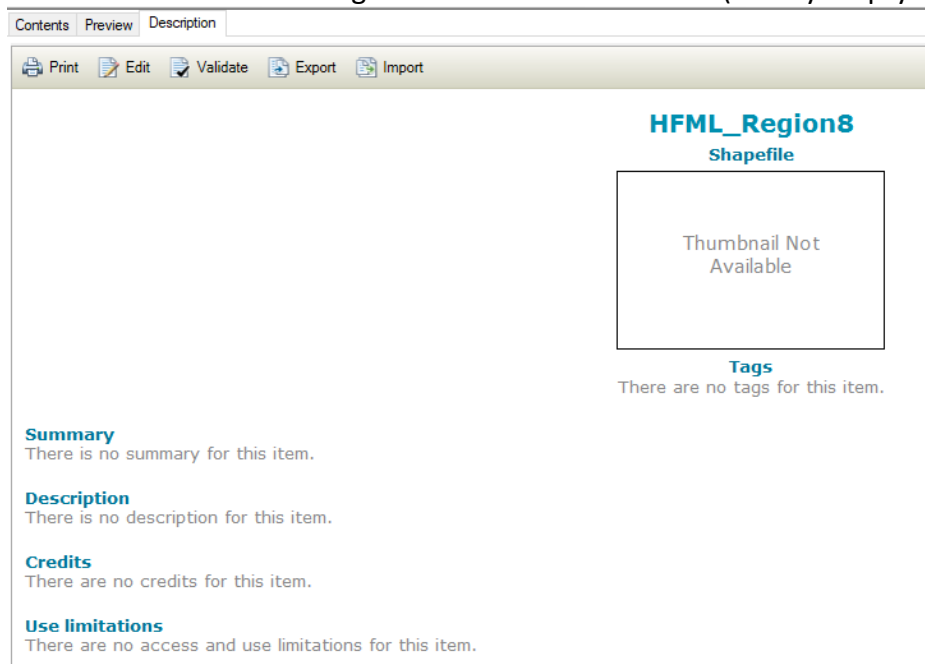
4. Click on the "Edit" [Edit] button on top of the right window to start entering the information in the metadata record.

5. The window contains now the list of sections contained in the metadata record on the list and the editable fields on the right .

6. Go through the different sections of the metadata record listed on the left side of the window to fill all the fields reported in Annex 1 and therefore comply with the minimum metadata profile recommended by the HGL. **Don't forget to save your work from time to time by clicking on the save button on top of the window**
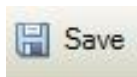
. You need to re-enter the edit mode after that.

7. Once done, check that none of the items in the metadata profile has a red cross next to its name like this . Such a red cross indicates that some fields are still incomplete in that item. Once you click on the item in question, the list of issues to be solved appears on the top of the window like shown here:

8. Your work is done once all the items in the record shows a green check like this
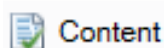
9. Save one last time the metadata record and go through the final result to make sure that all looks good!

Please note that the metadata record you have created is stored in a .shp.xml file placed together with the other files composing the shape files. As per our example, the file in question is therefore named:

Annex 3 – Example of filled metadata record in ArcGIS 10.5 using the example reported in Annex 1

Government health facilities of Region VIII , Philippines (2015)

Shapefile

Thumbnail Not Available
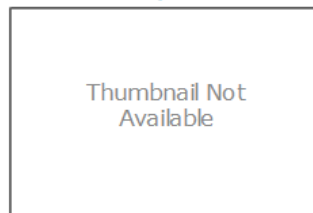
Tags
Region VIII, Philippines

Summary
This geospatial data layer has been developed to serve as the official reference for governmental health facilities in Region VIII

Description
This geospatial data layer has been developed by the Department of Health of the Philippines (DOH) and contains the location of all the government health facilities in Region VIII as observed in 2015

Credits
Department of Health of the Philippines (DOH)

Use limitations
The use of this geospatial data layer is limited to non-commercial use

Extent
West	124.057038	**East**	125.3
North	7.5	**South**	5.75

Scale Range
Maximum (zoomed in) 1:500,000
Minimum (zoomed out) 1:150,000,000

ArcGIS Metadata ▶

Topics and Keywords ▶

THEMES OR CATEGORIES OF THE RESOURCE health

* CONTENT TYPE Downloadable Data
  EXPORT TO FGDC CSDGM XML FORMAT AS RESOURCE DESCRIPTION No

PLACE KEYWORDS Region VIII, Philippines

Hide Topics and Keywords ▲

Citation ▶

TITLE Government health facilities of Region VIII , Philippines (2015)
CREATION DATE 2015-05-01 00:00:00

PRESENTATION FORMATS * digital map

Hide Citation ▲

Resource Details ▶

DATASET LANGUAGES * English (UNITED STATES)
DATASET CHARACTER SET utf8 - 8 bit UCS Transfer Format

STATUS completed
SPATIAL REPRESENTATION TYPE * vector

SPATIAL RESOLUTION
    DATASET'S SCALE
        SCALE DENOMINATOR   100000

  * PROCESSING ENVIRONMENT   Version 6.2 (Build 9200) ; Esri ArcGIS 10.5.0.6491

    CREDITS
        Department of Health of the Philippines (DOH)

Extents ▶

EXTENT
    GEOGRAPHIC EXTENT
        BOUNDING RECTANGLE
            EXTENT TYPE   Extent used for searching
            * WEST LONGITUDE   124.057038
            EAST LONGITUDE   125.3
            NORTH LATITUDE   7.5
            SOUTH LATITUDE   5.75
            * EXTENT CONTAINS THE RESOURCE   Yes

    EXTENT IN THE ITEM'S COORDINATE SYSTEM
        * WEST LONGITUDE   124.057038
        * EAST LONGITUDE   125.293230
        * SOUTH LATITUDE   5.821850
        * NORTH LATITUDE   7.514422
        * EXTENT CONTAINS THE RESOURCE   Yes

    Hide Extents ▲

Resource Points of Contact ▶

POINT OF CONTACT
    INDIVIDUAL'S NAME   Antoine Marello
    ORGANIZATION'S NAME   Department of Health of the Philippines (DOH)
    CONTACT'S POSITION   Data manager
    CONTACT'S ROLE   point of contact

        CONTACT INFORMATION   ▶
            PHONE
                VOICE   +63301274438

            ADDRESS
                TYPE   postal
                COUNTRY   PH
                E-MAIL ADDRESS   AntoineMarello@gmail.com

            Hide Contact information ▲

    Hide Resource Points of Contact ▲

Resource Maintenance ▶

RESOURCE MAINTENANCE
    UPDATE FREQUENCY   continual

    Hide Resource Maintenance ▲

Resource Constraints ▶

CONSTRAINTS
    LIMITATIONS OF USE
        The use of this geospatial data layer is limited to non-commercial use. It is being distributed without warranty
        of any kind, either expressed or implied. The responsibility for the interpretation and use of the data lies with
        the user. In no event shall the DOH be liable for damages arising from its use.

LEGAL CONSTRAINTS
    ACCESS CONSTRAINTS   copyright
    USE CONSTRAINTS   restricted

SECURITY CONSTRAINTS
    CLASSIFICATION   unclassified

    Hide Resource Constraints ▲

Spatial Reference ▶

ARCGIS COORDINATE SYSTEM
    * TYPE   Geographic

* GEOGRAPHIC COORDINATE REFERENCE   GCS_WGS_1984
* COORDINATE REFERENCE DETAILS
  GEOGRAPHIC COORDINATE SYSTEM
    WELL-KNOWN IDENTIFIER   4326
    X ORIGIN   -400
    Y ORIGIN   -400
    XY SCALE   11258999068426.238
    Z ORIGIN   -100000
    Z SCALE   10000
    M ORIGIN   -100000
    M SCALE   10000
    XY TOLERANCE   8.983152841195215e-009
    Z TOLERANCE   0.001
    M TOLERANCE   0.001
    HIGH PRECISION   true
    LEFT LONGITUDE   -180
    LATEST WELL-KNOWN IDENTIFIER   4326
    WELL-KNOWN TEXT   GEOGCS["GCS_WGS_1984",DATUM["D_WGS_1984",SPHEROID
    ["WGS_1984",6378137.0,298.257223563]],PRIMEM["Greenwich",0.0],UNIT
    ["Degree",0.0174532925199433],AUTHORITY["EPSG",4326]]

REFERENCE SYSTEM IDENTIFIER
  * VALUE   4326
  * CODESPACE   EPSG
  * VERSION   6.14(3.0.1)

Hide Spatial Reference ▲

Spatial Data Properties ▶

VECTOR   ▶
  * LEVEL OF TOPOLOGY FOR THIS DATASET   geometry only

  GEOMETRIC OBJECTS
    FEATURE CLASS NAME   HFML_Region8
    * OBJECT TYPE   point
    * OBJECT COUNT   77

  Hide Vector ▲

ARCGIS FEATURE CLASS PROPERTIES   ▶
  FEATURE CLASS NAME   HFML_Region8
    * FEATURE TYPE   Simple
    * GEOMETRY TYPE   Point
    * HAS TOPOLOGY   FALSE
    * FEATURE COUNT   77
    * SPATIAL INDEX   TRUE
    * LINEAR REFERENCING   FALSE

  Hide ArcGIS Feature Class Properties ▲

Hide Spatial Data Properties ▲

Data Quality ▶

SCOPE OF QUALITY INFORMATION   ▶
  RESOURCE LEVEL   dataset

  Hide Scope of quality information ▲

DATA QUALITY REPORT - ABSOLUTE EXTERNAL POSITIONAL ACCURACY   ▶
  DIMENSION   horizontal

  MEASURE DESCRIPTION
    The maximum horizontal error accepted at the scale of of work is of 50 meters

  EVALUATION METHOD
    The maximum horizontal error has been measured for 75% of the points in the dataset using satellite
    images as ground reference

  CONFORMANCE TEST RESULTS
    TEST PASSED   Yes
    RESULT EXPLANATION
      82% of the points in the datasets are finding themselves within less than 50 meter from the
      corresponding location in the satellite image

Lineage ▶

LINEAGE STATEMENT
This geospatial data layer has been generated by compiling different source of geographic coordinates from within and outside the DOH. These coordinates have then been checked before linking them to the information stored in the National health Facility Registry

Hide Lineage ▲

Distribution ▶

DISTRIBUTION FORMAT
    * NAME   Shapefile
    VERSION   ArcGIS 10.5

TRANSFER OPTIONS
    * TRANSFER SIZE   0.002

Hide Distribution ▲

Fields ▶

DETAILS FOR OBJECT HFML_Region8 ▶
    * TYPE   Feature Class
    * ROW COUNT   77
    DEFINITION
    Government health facilities of Region VIII

    DEFINITION SOURCE
    DOH

    FIELD FID ▶
        * ALIAS   FID
        * DATA TYPE   OID
        * WIDTH   4
        * PRECISION   0
        * SCALE   0
        * FIELD DESCRIPTION
        Internal feature number.

        * DESCRIPTION SOURCE
        Esri

        * DESCRIPTION OF VALUES
        Sequential unique whole numbers that are automatically generated.

        Hide Field FID ▲

    FIELD Shape ▶
        * ALIAS   Shape
        * DATA TYPE   Geometry
        * WIDTH   0
        * PRECISION   0
        * SCALE   0
        * FIELD DESCRIPTION
        Feature geometry.

        * DESCRIPTION SOURCE
        Esri

        * DESCRIPTION OF VALUES
        Coordinates defining the features.

23

Metadata Details ▶

* METADATA LANGUAGE   English (UNITED STATES)
* METADATA CHARACTER SET   utf8 - 8 bit UCS Transfer Format

SCOPE OF THE DATA DESCRIBED BY THE METADATA   * dataset
SCOPE NAME   * dataset

* LAST UPDATE   2018-06-11

ARCGIS METADATA PROPERTIES
  METADATA FORMAT   ArcGIS 1.0
  METADATA STYLE   ISO 19139 Metadata Implementation Specification
  STANDARD OR PROFILE USED TO EDIT METADATA   ISO19139

  CREATED IN ARCGIS FOR THE ITEM   2018-06-11   08:46:54
  LAST MODIFIED IN ARCGIS FOR THE ITEM   2018-06-11   16:21:06

  AUTOMATIC UPDATES
    HAVE BEEN PERFORMED   Yes
    LAST UPDATE   2018-06-11   16:21:06

Hide Metadata Details ▲

Metadata Contacts ▶

METADATA CONTACT
  INDIVIDUAL'S NAME   Antoine Marello
  ORGANIZATION'S NAME   Department of Health of the Philippines (DOH)
  CONTACT'S POSITION   Data manager
  CONTACT'S ROLE   point of contact

    CONTACT INFORMATION   ▶
     PHONE
      VOICE   +63301274438

     ADDRESS
      TYPE   postal
      COUNTRY   PH
      E-MAIL ADDRESS   AntoineMarello@gmail.com

Hide Contact information ▲

Metadata Maintenance ▶

MAINTENANCE
  UPDATE FREQUENCY   unknown

Hide Metadata Maintenance ▲

Annex 4 – Filling the minimum recommended metadata profile for geospatial data in QGIS 3.0

The filling of the metadata record for a given geospatial data layer happens directly in QGIS 3.0 based on the following steps:.

1. Launch QGIS 3.0 and add the geospatial data layer you want to document in the view.
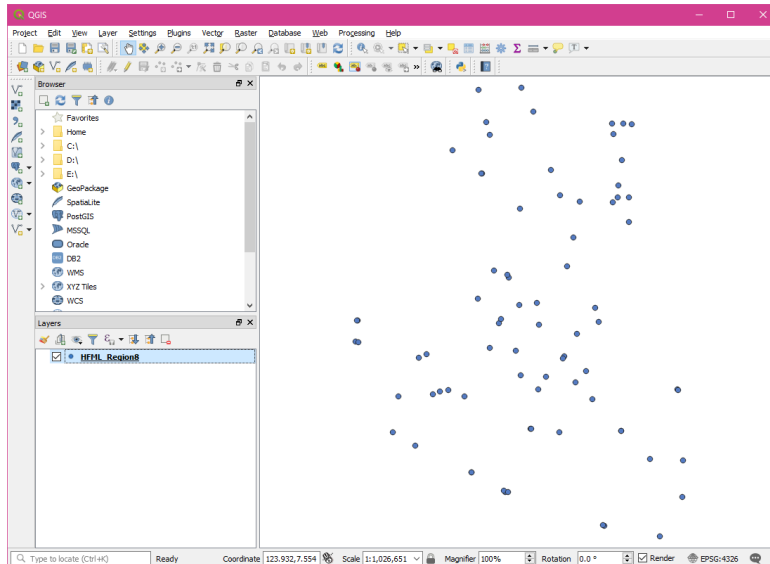
2. Right click on the layer name in the "Layers" panel and select properties.
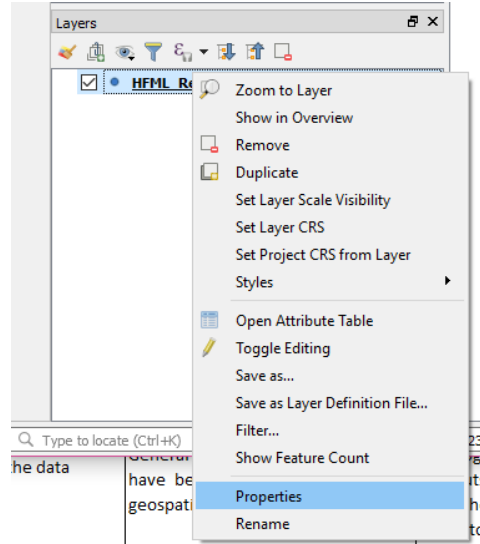
3. Select the "Metadata" item on the left in the window that opens.
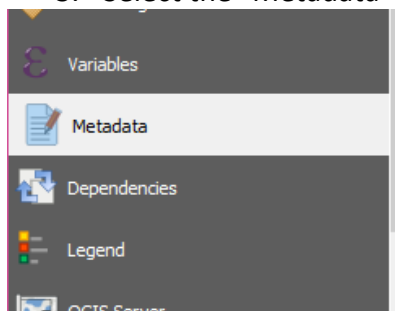
4. This will make all the fields from the QGIS metadata profile appear on the right side of the layer properties window
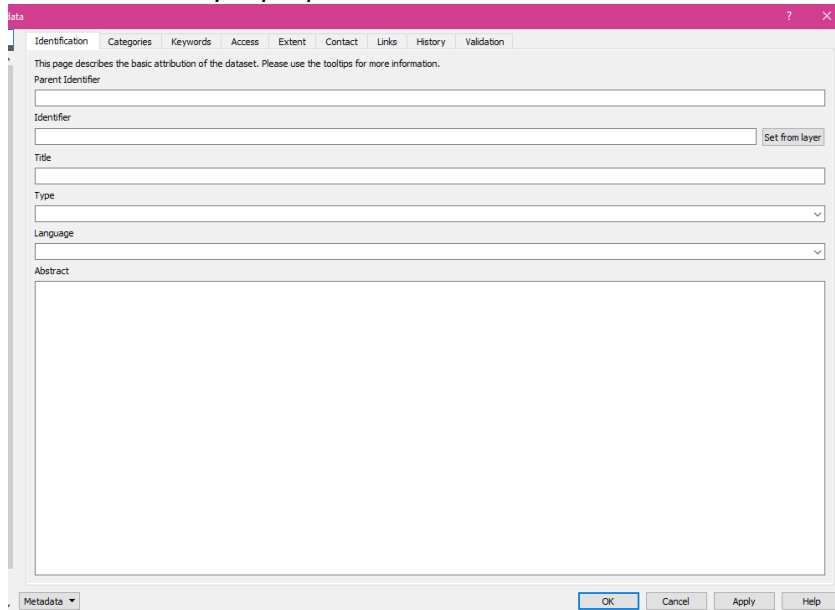
5. Save your metadata record by clicking on the Metadata dropdown button at the bottom left of the window ![Metadata ▼] and then on the "Save metadata" option
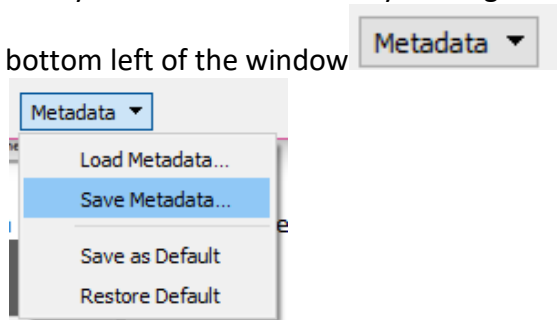
6. This will give you access to a window where you will be able to specify the path and name for the metadata file (.qmd extension). The recommendation is to save it in the same folder and with the same name as the geospatial data layer.

7. Go through the different tabs of the metadata record listed at the top of the window to fill all the fields reported in Annex 1 and therefore comply with the minimum metadata profile recommended by the HGL.

Please note that:

- You might have to use the green ![green plus] or red ![red minus] buttons to respectively store or remove the information in some of the fields.
- Some information like the geographic coordinate system or extent can be directly extracted from the geospatial data layer
- **You should save your work from time to time by following the steps listed under step 5 and overwriting the same metadata record file. If you don't, your changes will only be saved in the QGIS project file.**

8. Verify the content of your metadata record and save it one last time!

Annex 5 – Example of filled metadata record in QGIS 3.0 using the example reported in Annex 1

Identification

Categories

Keywords

Access

Extent

Contact

History