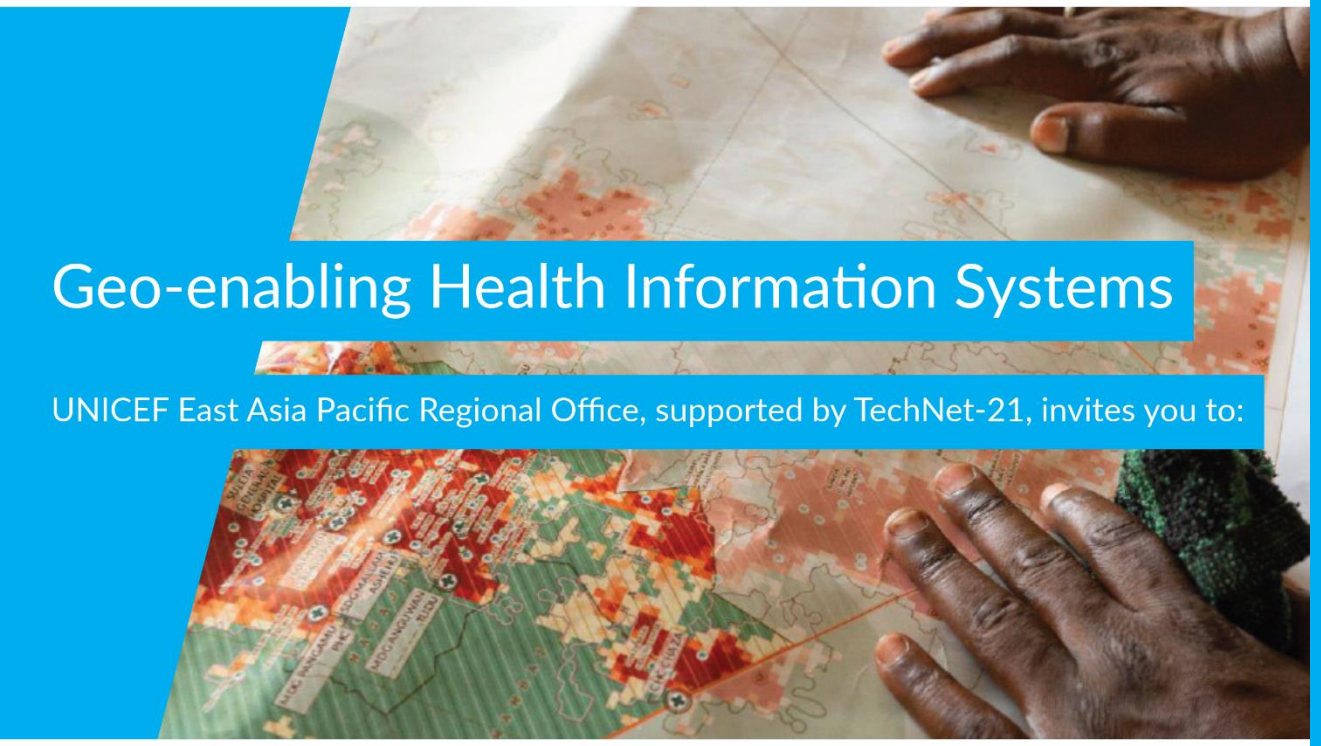
unicef for every child

MORU Tropical Health Network | HEALTH GEOLAB Hub

TechNet-21 The Technical Network for Strengthening Immunization Services



Geo-enabling Health Information Systems

UNICEF East Asia Pacific Regional Office, supported by TechNet-21, invites you to:

Learn how to geo-enable health information systems and programmes

Join us for a bi-weekly web-series starting **19 June 2024**

- Demonstrate the potential of geospatial data and technologies in public health
- Introduce HIS geo-enabling framework and its implementation in countries
- Provide knowledge and resources to implement the HIS geo-enabling framework

Go to https://tn21.org/UNICEF-EAPRO
or Scan QR Code to Register

Ongoing registration

Joining any one session also permitted

6 Modules of around 2 hours each…except today which will reach around 3 hours

Certificates provided on completion by UNICEF & MORU

Workshop Objectives

Disseminate operational guidance materials that can assist countries in implementing the geo-enablement process for health programs in general and the development and implementation of micro plans in particular

More specifically:

- Demonstrate the potential of geospatial data and technologies in public health
- Introduce the HIS geo-enabling framework and its implementation in countries
- Transfer knowledge, expertise and resources that will allow participants to implement the HIS geo-enabling framework in their respective country

➡ At the end of this workshop, it is expected that the participants will have a better understanding of what geospatial data and technologies can bring to public health programs and how to geo-enable their health information system in a sustainable way to benefit from this type of data and technologies

➡ This is not a GIS training

Workshop material

https://bit.ly/4d2nfTS

REFERENCE_MATERIAL

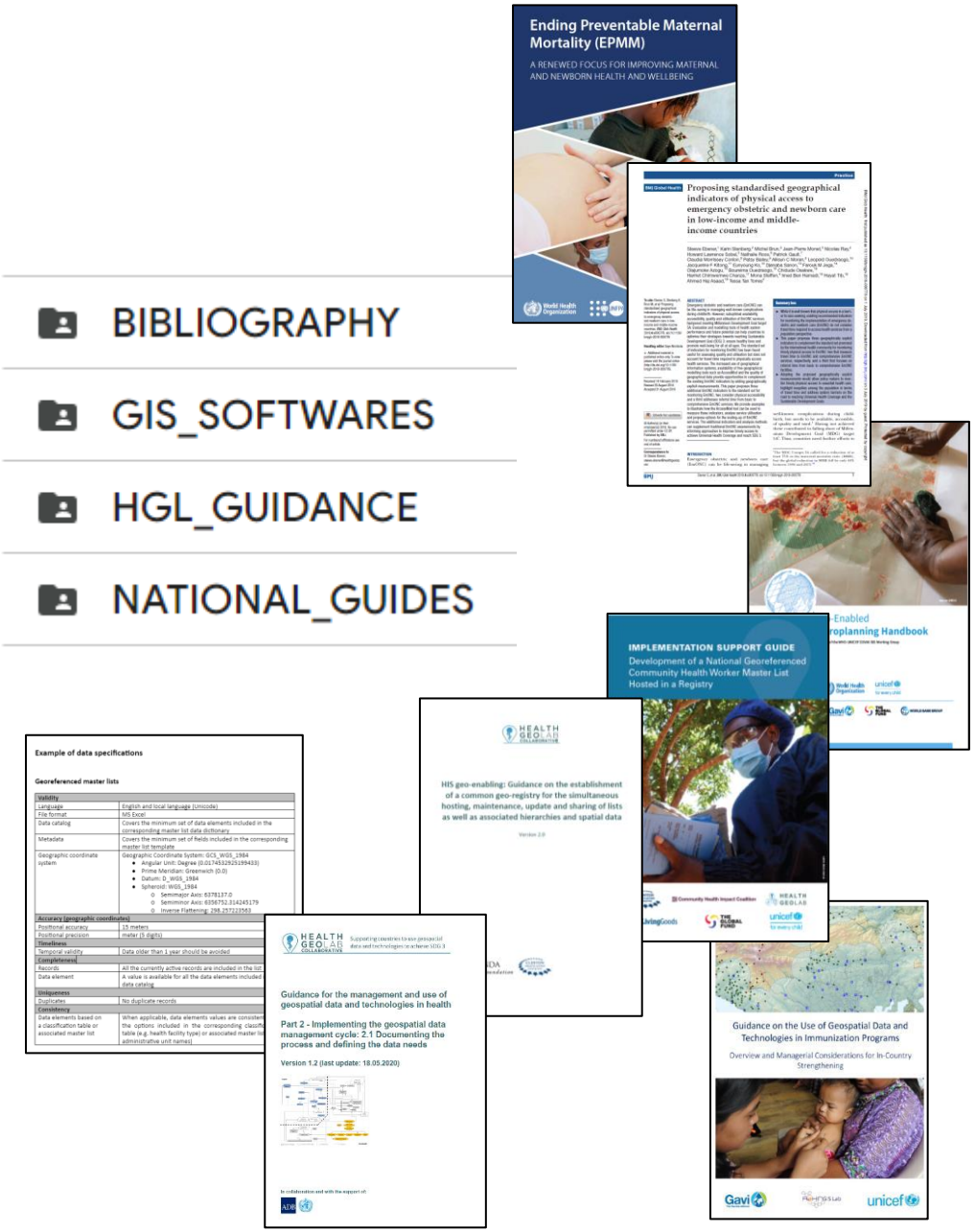
PRESENTATIONS

BIBLIOGRAPHY

GIS_SOFTWARES

HGL_GUIDANCE

NATIONAL_GUIDES



Glossary of terms: https://bit.ly/37Wje0v

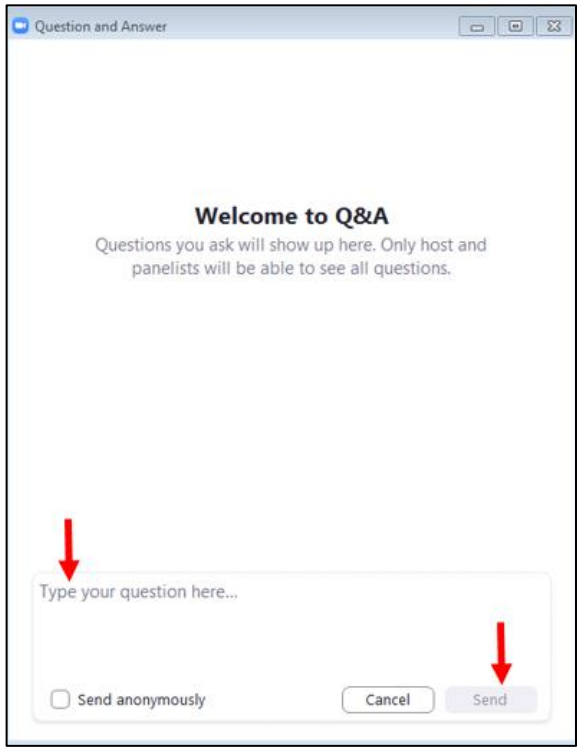
Recording module 1: https://youtu.be/kyLvtGKA27Q

Recording module 2: https://youtu.be/oqOC8064buE

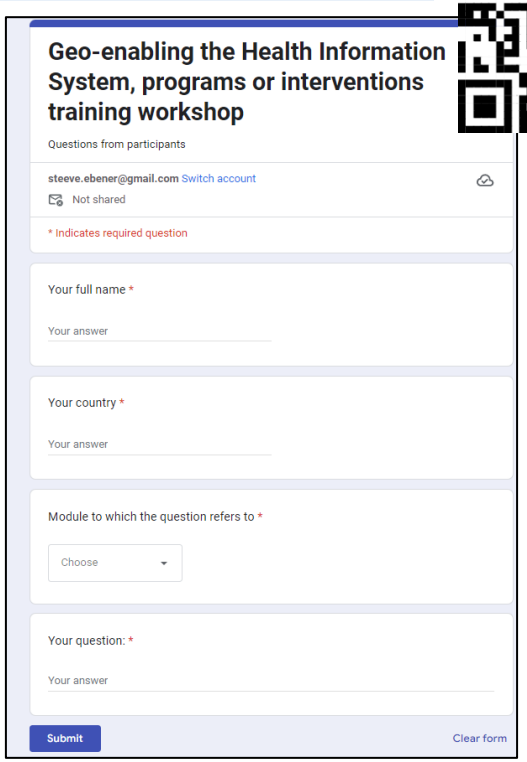
Recording module 3: https://youtu.be/xD4cFEs9Afs

Questions and knowledge sharing during the modules?

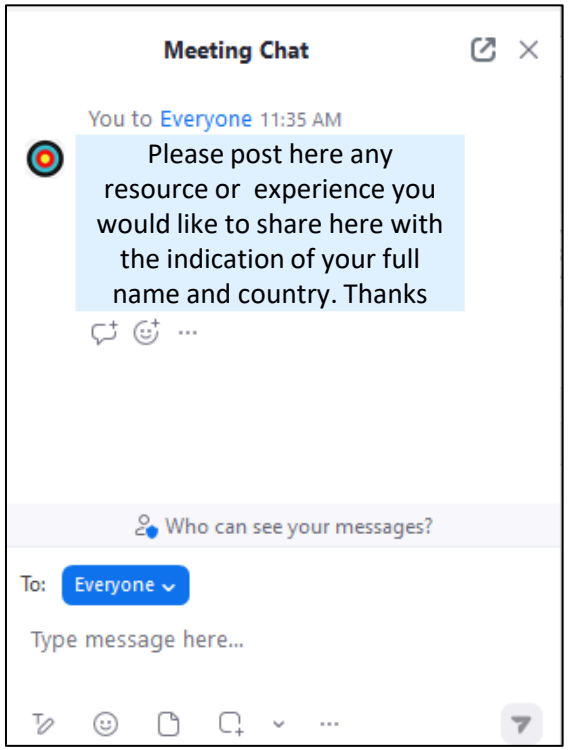
https://tinyurl.com/3999y744



Please post your questions in the Zoom Q&A (not the chat)



You can also ask questions using this short Google form (between modules for example)

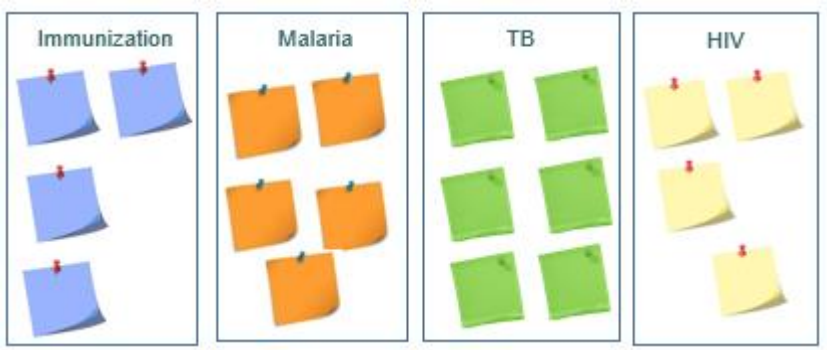


You can share any resource or experience you see relevant to the participants in the chat

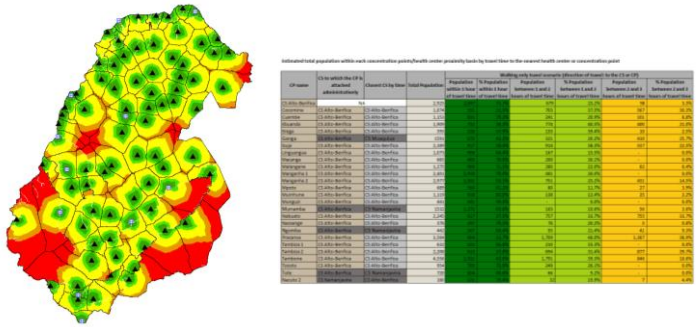
We will also be using the chat to share information

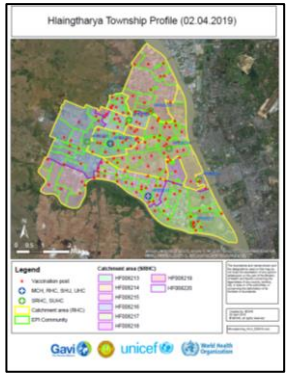
We will answer them as much as possible during the modules

Recap of Module 3

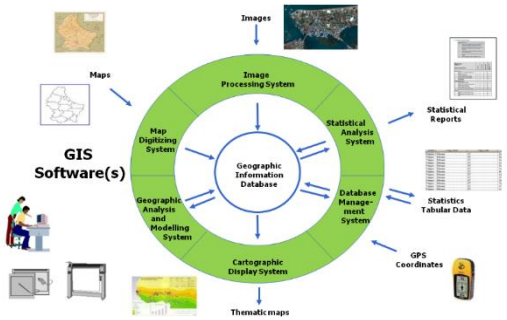


Result of the geographic features identification performed by the participants

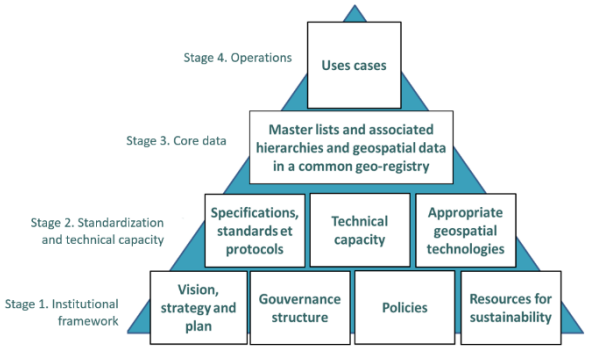




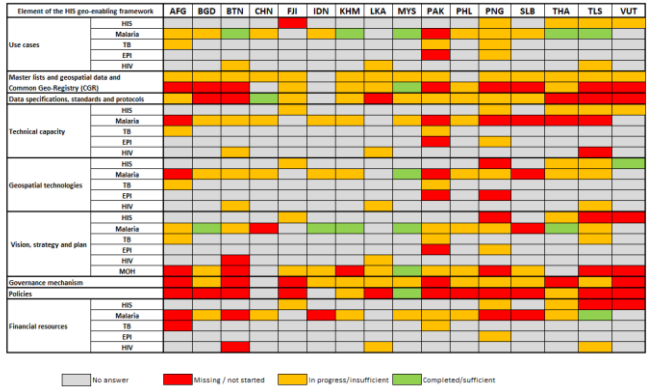
Define the purpose, audience, content and format of the final GIS-based products



Identify needed hardware, software and technical expertise



Assess the current geo-enablement level of the HIS, program or intervention



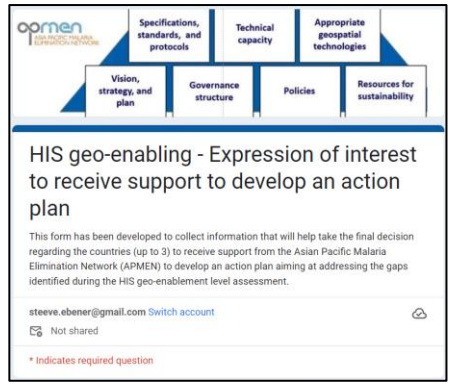
Result of the HIS geo-enablement level assessment conducted for Asia and Pacific

Countries to receive APMEN support

Shortlisted countries (17 July 2024)

Country name	Nbr of MOH programs having completed the questionnaire	Nbr of MOH participants to the first 2 modules	Local partners participating to the first 2 modules	Total
Papua New Guinea	5	8	1	14
Pakistan	4	2	1	7
Timor-Leste	4	1	1	6
Indonesia	2	3	1	6
Afghanistan	2	2	1	5
Bangladesh	1	3	1	5

Expression of interest



Final decision (14 August 2024)

Country	Nbr of MOH individuals contacted	Nbr of partner individuals contacted	Nbr positive MOH answers	Nbr positive partners answers
Papua New Guinea	10	8	3	2
Bangladesh	3	11	1	1
Pakistan	6	3	1	0
Timor-Leste	4	2	1	0
Indonesia	7	10	0	2

Only country supported at this stage to develop an action plan aiming at filling the gaps identified during the assessment

Geo-enabling the Health Information System, programs or interventions training workshop for Asia Pacific

...and beyond

Module 4

Agenda Module 4

15 min - Recap of Module 3 and agenda of Module 4

30 min – **Session 12:** Assess the availability, quality and accessibility of data and information: Introduction to the geospatial data management cycle

60 min - **Session 13:** Implement the geospatial data management cycle (define the terminology, data specifications and the ground reference)

15 min - **Session 14**: Implement the geospatial data management cycle (document the data)

45 min - **Session 15** Implement the geospatial data management cycle (compile existing data, identify and fill data gaps

➡ Geospatial data management cycle

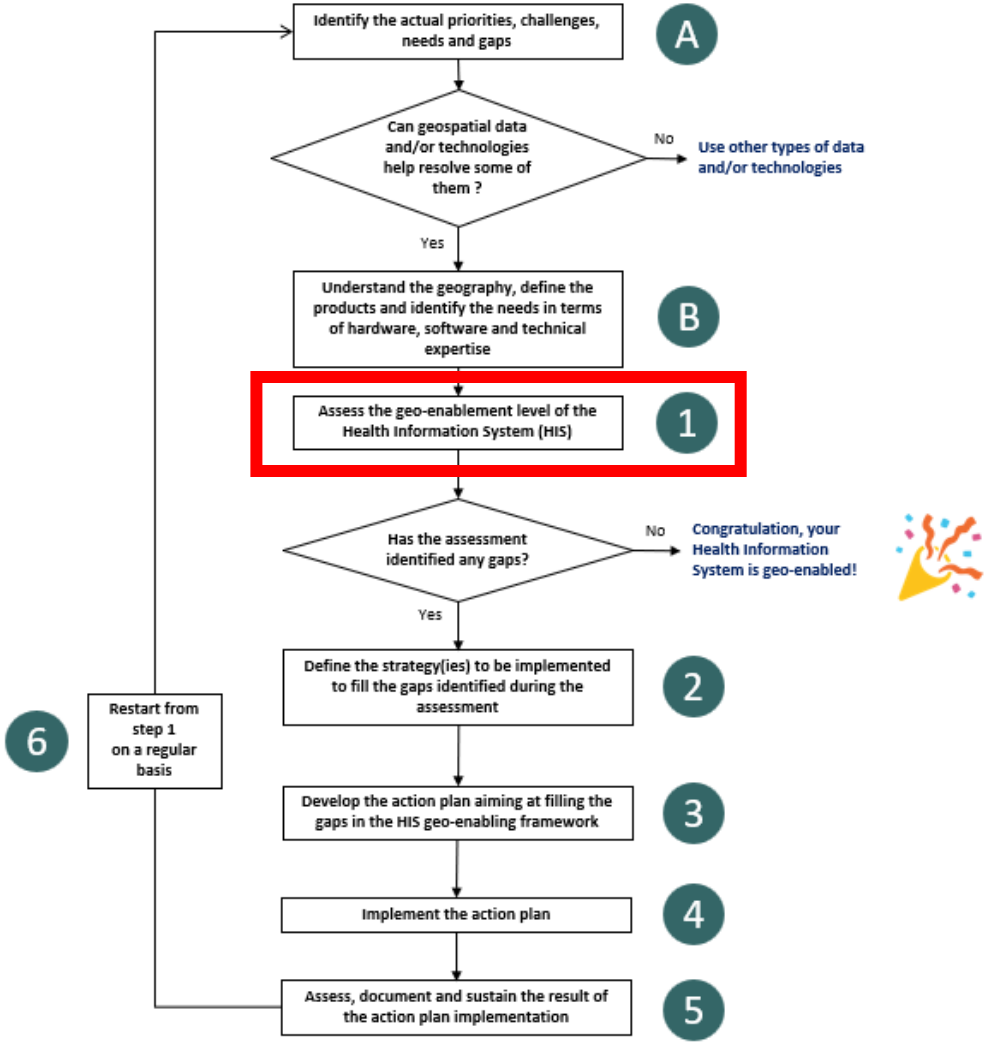
Geo-enabling the Health Information System, programs or interventions training workshop for Asia Pacific

...and beyond

Session 12: Assess the availability, accessibility and quality of data and information: the geospatial data management cycle

HIS Geo-enabling framework implementation process

Step 1 : Assess the availability, accessibility and quality of data and information

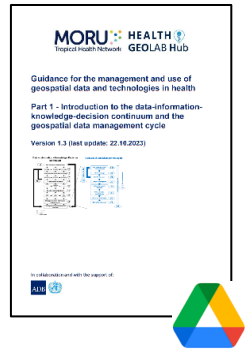


When geo-enabling the HIS, this activity only look at where the master list for a set of geographic features core to public health find themselves along a predefined continuum

- Geo-enabling a program or intervention requires for a more in-depth assessment of the availability, accessibility and quality of the data needed to generate the information products defined during step B
- Such an assessment is to take place during Step 1 and meant to identify currently existing gaps and how to fill them
- Entry point for the geospatial data management cycle

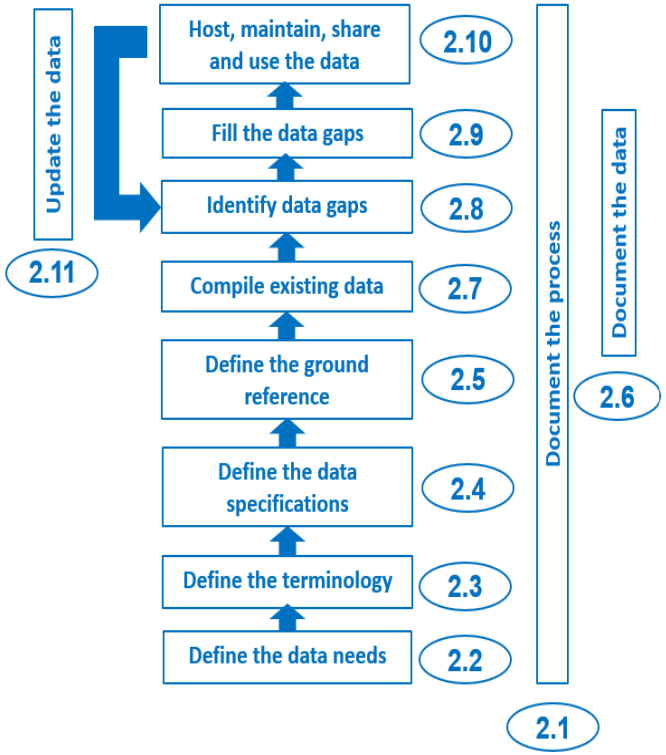
The geospatial data management cycle

1



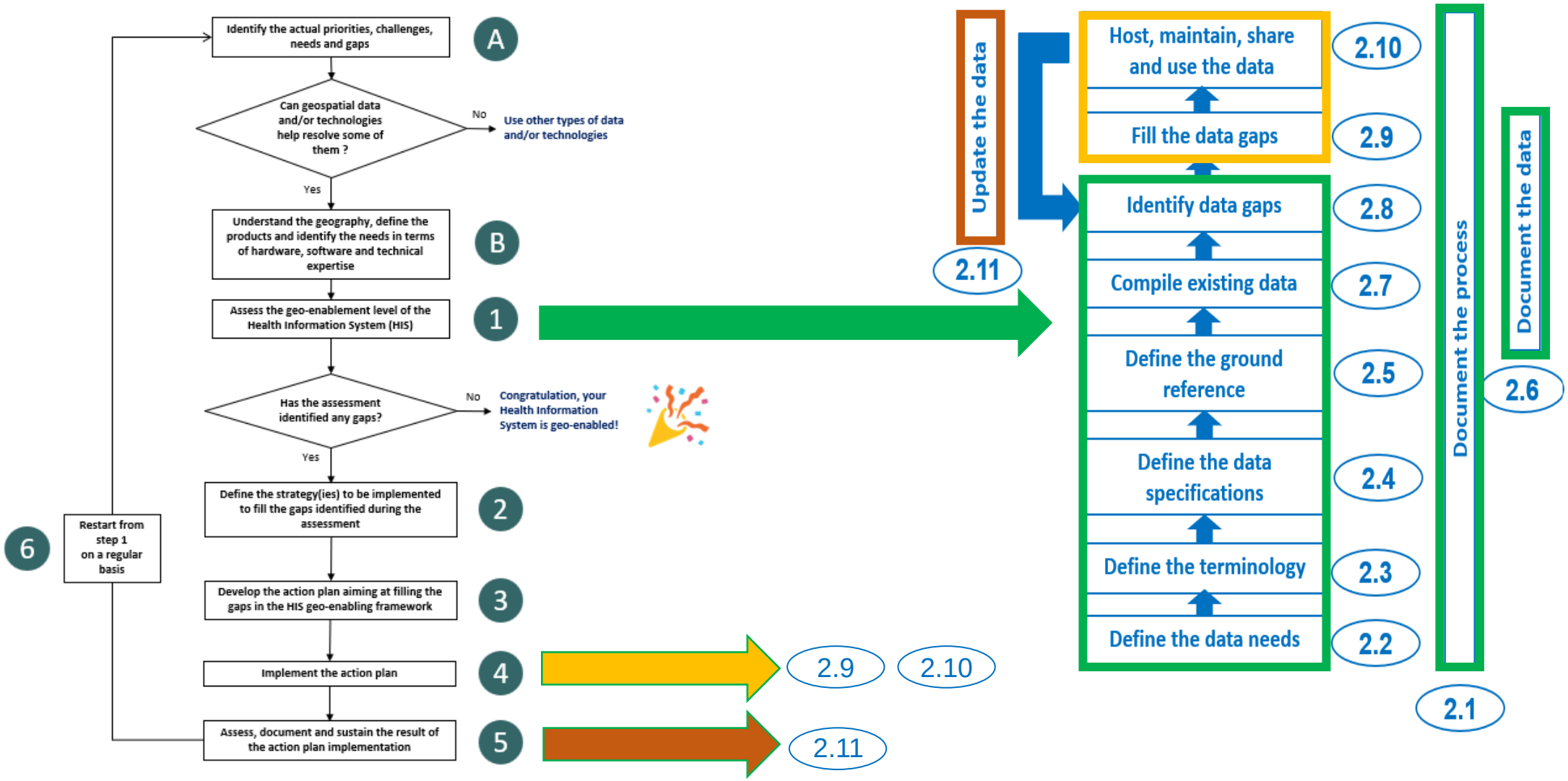
Generating and maintaining good quality data (geospatial, statistical) and products require proper data management standards, processes, and protocols to be defined and implemented.

- The **geospatial data management cycle** covers the steps to follow to define and implement proper data standards, processes, and protocols
- Ensures **data quality** and therefore **the quality of the information and information product** to be generated out of it
- Applicable to georeferenced master lists as well as geospatial and statistical data



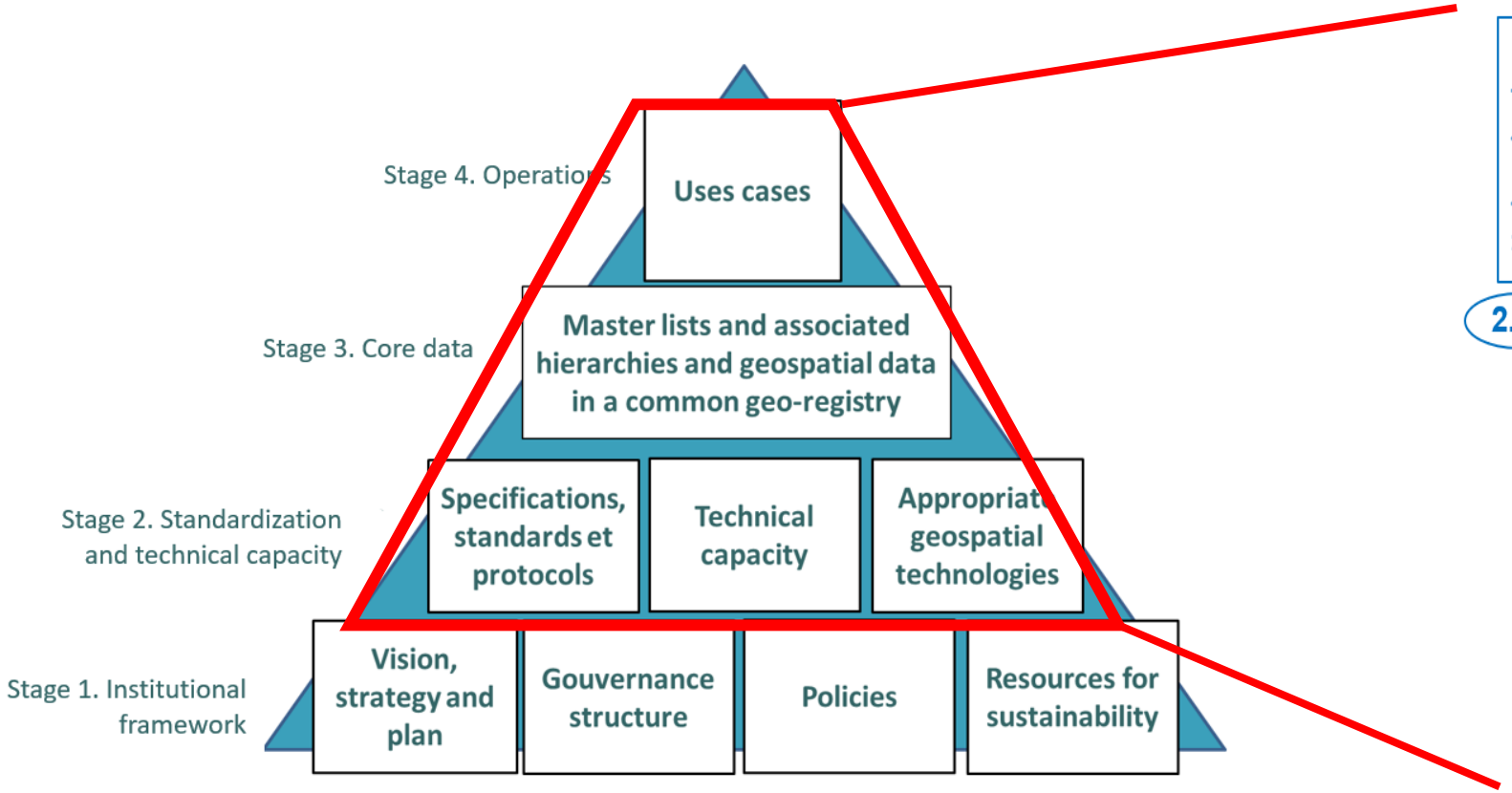
1. http://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part1.pdf

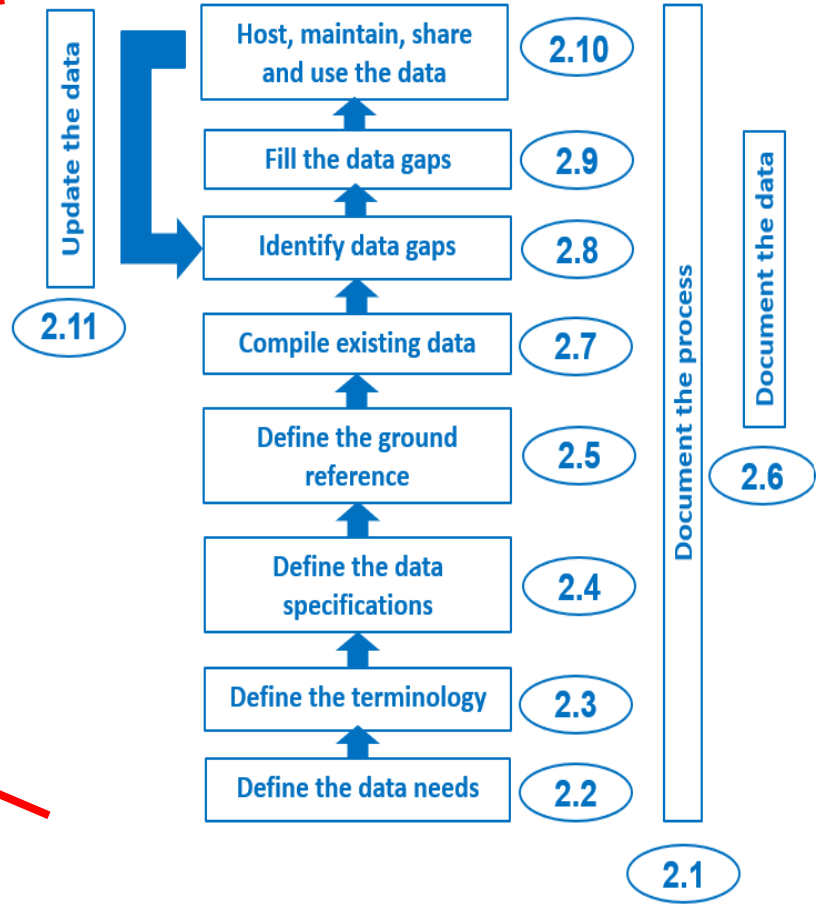
The geospatial data management cycle



The geospatial data management cycle

The technical process behind the HIS geo-enabling framework





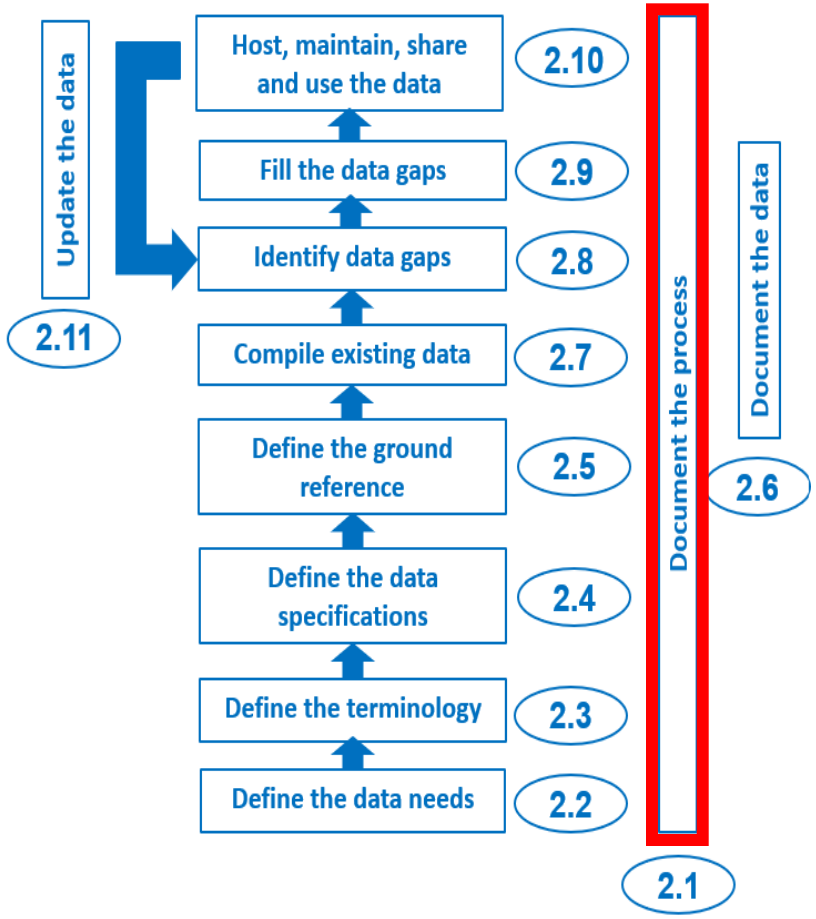
Strategic framework

Technical process

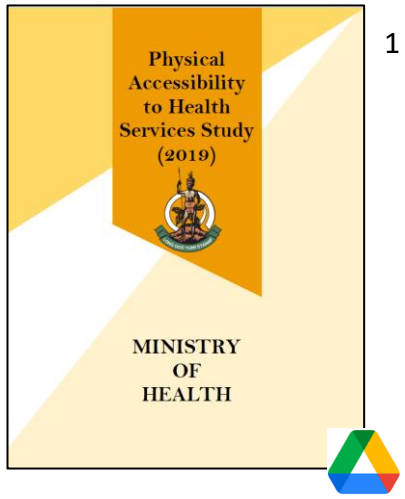
Document the process

Documenting each step from the beginning as precisely as possible ensures that the process can be replicated

- Applies to the entire geospatial data management cycle
- A critical activity often forgotten or considered as not being necessary



Good example of project implementation documentation



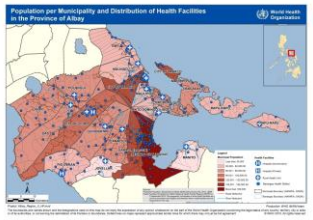
1

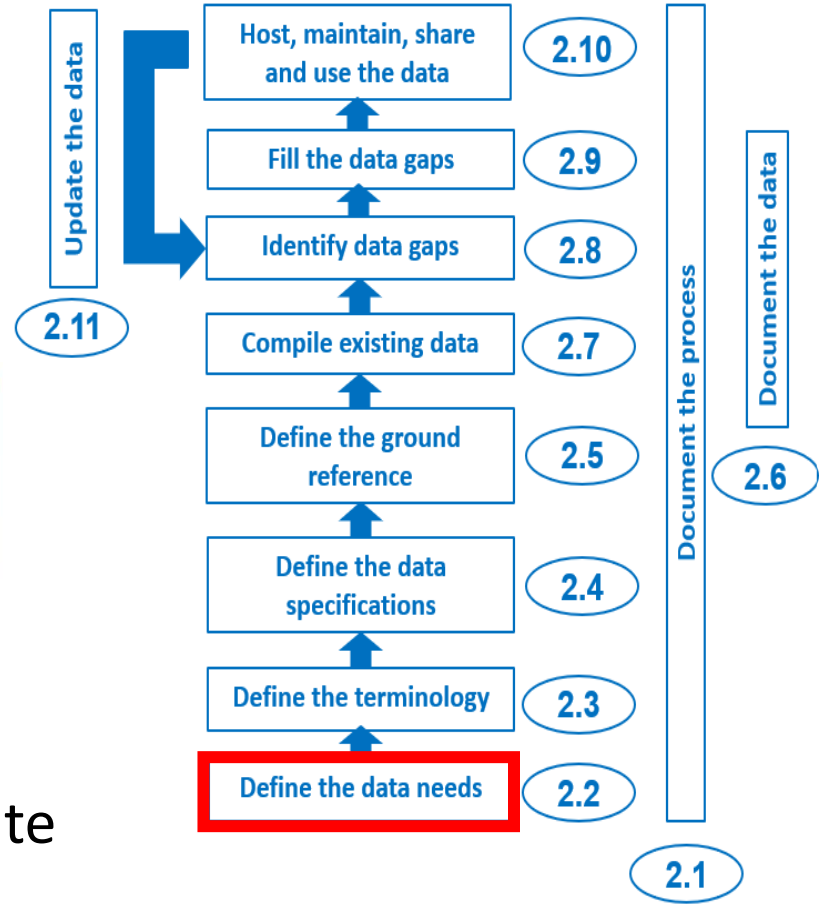
1 https://healthgeolab.net/KNOW_REP/Acc_Analysis_VUT_050224_FINAL.pdf

Define the data needs

This step builds on two activities completed during step B of the HIS geo-enabling framework implementation process:

1. Understand and document the geographic context (Session 6)
2. Define the purpose, audience, content and format of the GIS-based products to be generated to support the program or intervention (Session 8)





➡ Used to identify the data needed to generate the defined GIS-based products

Define the data needs

The content of the products identified during step B of the HIS geo-enabling process define the data that are needed

- Master lists
- Geospatial data
- Statistical data
- Input parameters

➡ Will depend on the application of geospatial data and technologies

➡ Examples on the right for microplanning

Example for microplanning

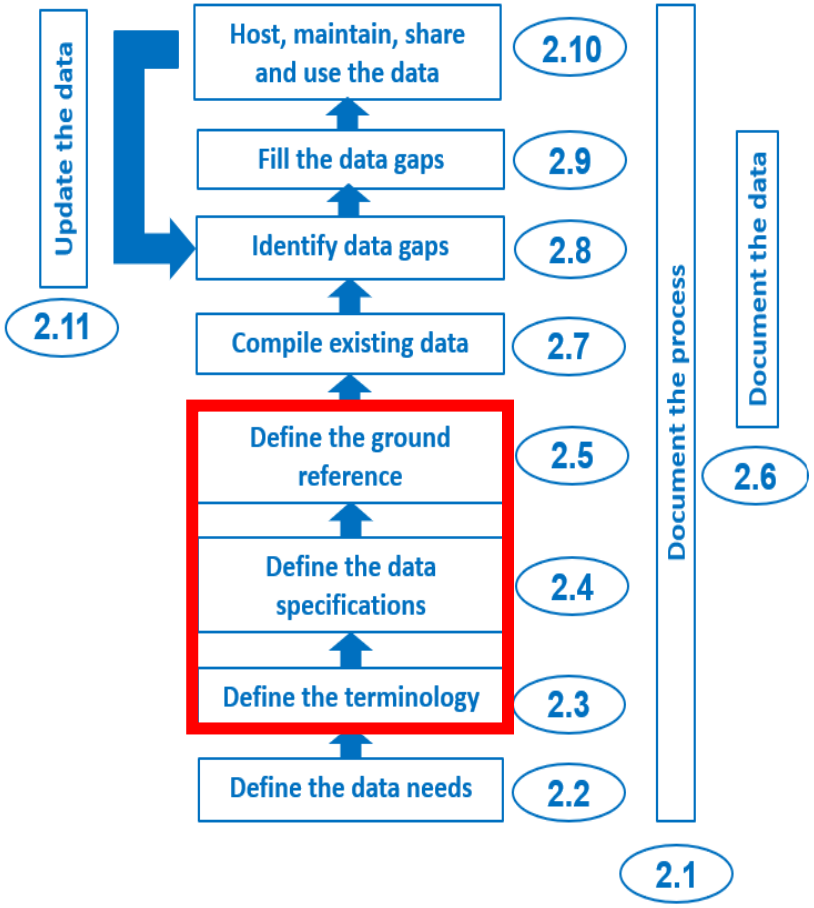
Geographic feature	Master list	Geospatial data
Service delivery points	X	X
Health areas and catchment areas	X	X
Administrative units (down to lowest level)	X	X
Human settlements (e.g. cities, villages)	X	X
Points of interest (e.g. schools, market places, landmarks)	X	X
Supply points (e.g. vaccine depots)	X	X
Transportation network	Not necessary	X
Hydrographic network	Not necessary	X
Population distribution (vector or raster)	Not applicable	X
Digital elevation model	Not applicable	X
Land cover/land use	Not applicable	X

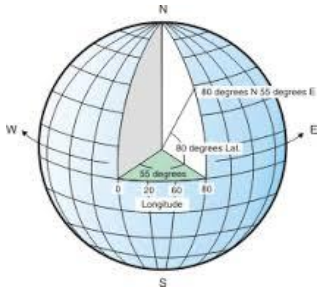
Data type	Description	Use
Population estimates and related statistics	Population estimates/spread, age, sex, employment, income, education, etc., disaggregated to lowest level possible	Population estimation and spatial distribution Contextual information on thematic maps
Service delivery point information and statistics	Services offered, capacity, type, presence of cold storage, accessibility to electricity and/or clean water	Geographic accessibility Contextual information on thematic maps
Previous campaign data	Coverage data from previous campaigns (e.g. indoor residual spraying, insecticide treated net distribution)	Determine campaign prioritization
Programmatic data	Outbreak data (e.g. malaria or measles outbreaks)	Determine campaign prioritization
Budgeted resources	Available resources for the health intervention (e.g. supplies, financial)	Determine resource allocation
Security concern	Type of security concern	Determine campaign feasibility
Travel scenario	Transportation mode, speed of travel	Input parameters for geographic accessibility, service location and route optimization modelling
Campaign target	Not necessary	Input parameters for geographic accessibility modelling

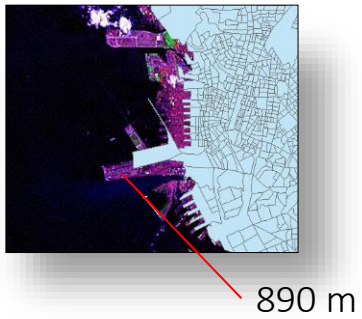
Define terminology, data specifications, and ground references

Defining the terminology ensures that all the actors involved speak the same language.

Data specifications and ground references (satellite imaging, master lists) provide the measurable criteria to assess and improve the quality of geospatial data across the 6 dimensions of data quality (completeness, uniqueness, timeliness, accuracy, validity and consistency).







We will come back to this in Session 13

Document the data (metadata)

All geospatial and statistical data used or generated as part of the implementation of the cycle must be properly documented.

This documentation provides the information needed to guarantee the quality of the data and should be started as soon as possible throughout the cycle.

The information is entered in a metadata record based on a standardized metadata profile.

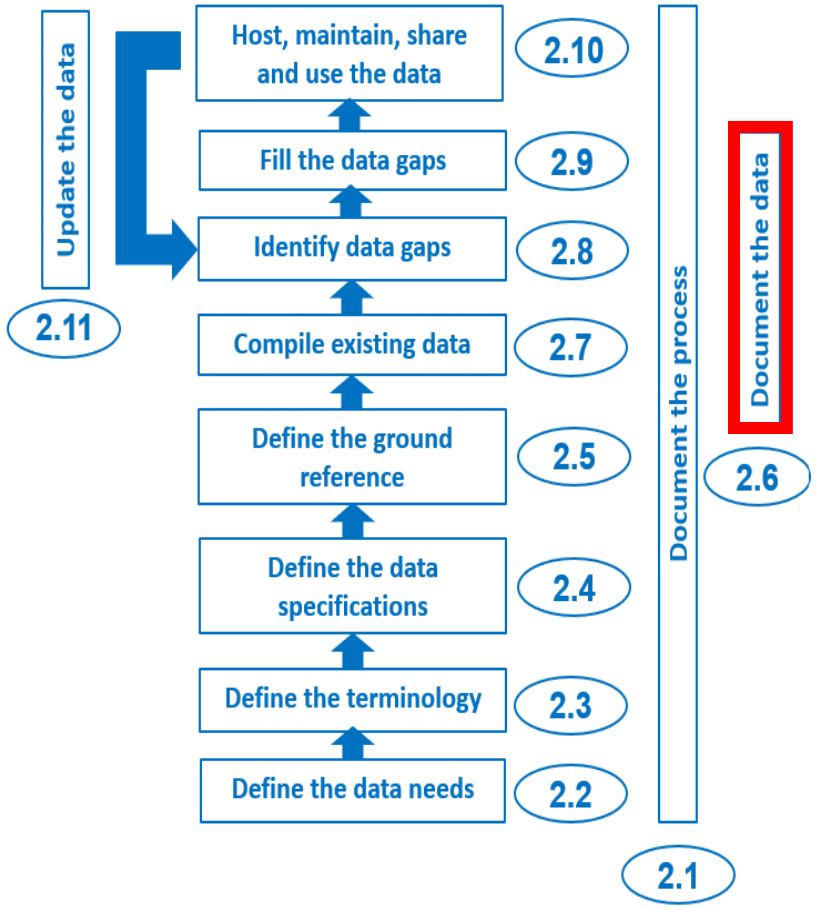
We will come back to this in Session 14

Nutrition Facts

4 servings per container
Serving size 1 1/2 cup (208g)

Amount per serving	
Calories	**240**
	% Daily Value*
Total Fat 4g	**5%**
Saturated Fat 1.5g	**8%**
Trans Fat 0g	
Cholesterol 5mg	**2%**
Sodium 430mg	**19%**
Total Carbohydrate 46g	**17%**
Dietary Fiber 7g	**25%**
Total Sugars 4g	
Includes 2g Added Sugars	**4%**
Protein 11g	
Vitamin D 2mcg	10%
Calcium 260mg	20%
Iron 6mg	35%
Potassium 240mg	6%

* The % Daily Value (DV) tells you how much a nutrient in a serving of food contributes to a daily diet. 2,000 calories a day is used for general nutrition advice.

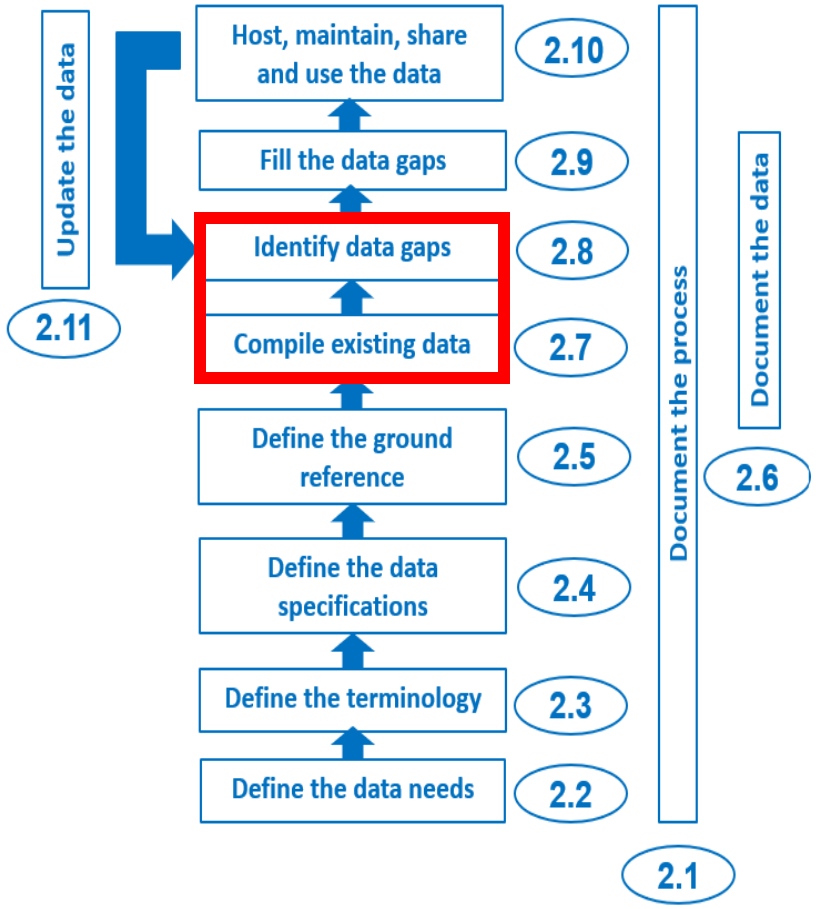


Compile existing data and identify gaps

The compilation and organization of existing data make it possible to reduce the duplication of efforts, and therefore save time and money.

The quality of the available and accessible data that has been compiled must then be assessed against the data specifications and ground reference to identify the potential gaps that will have to be filled.

Restrictions in the use and/or sharing of the compiled data can also result in the need to collect new or additional data.



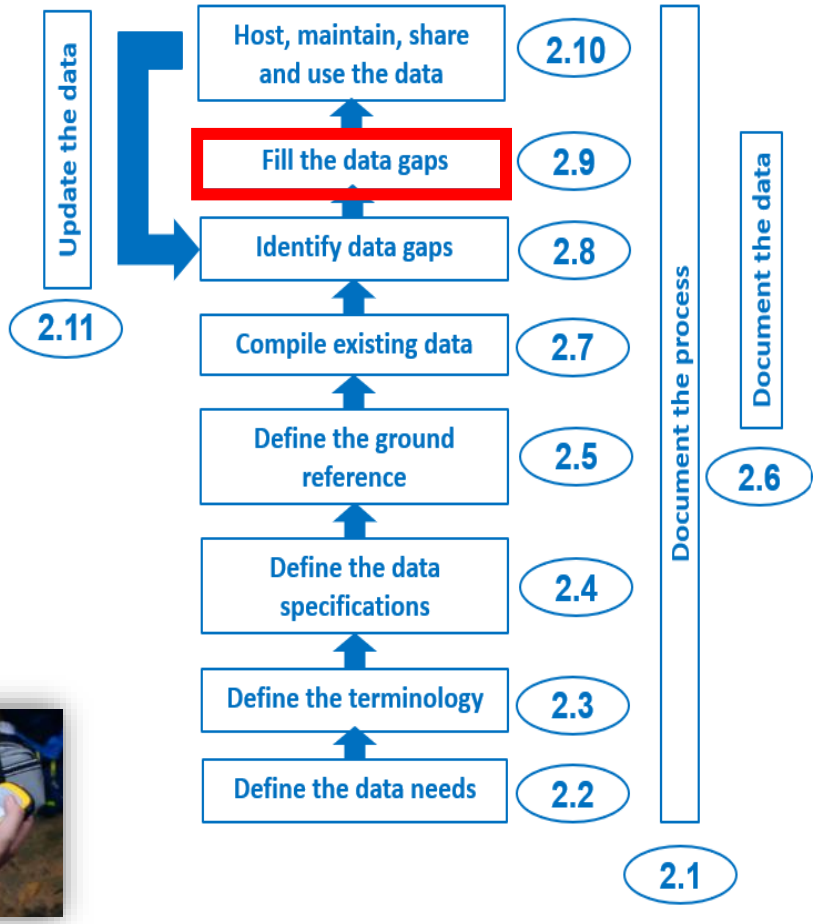
➡ We will come back to this in Session 15

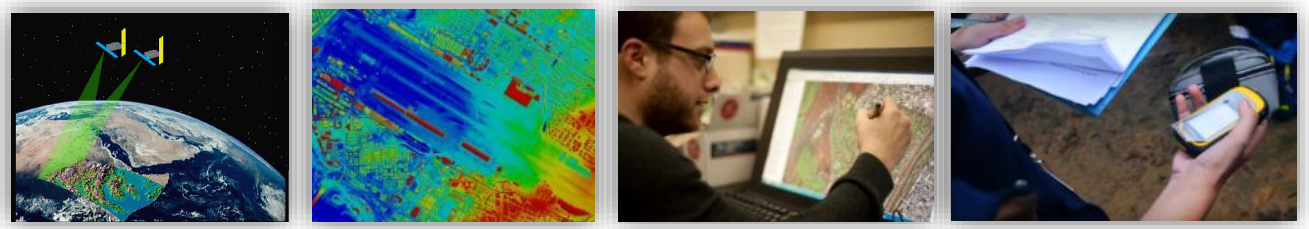
Fill the gaps in the data

Gaps in the data identified during the previous step should ideally be filled.

Different approaches can be used, including the improvement of existing data, the extraction of new data using different methods such as digitization or the collection of geographic coordinates in the field.

The data in question must also be validated before being used.





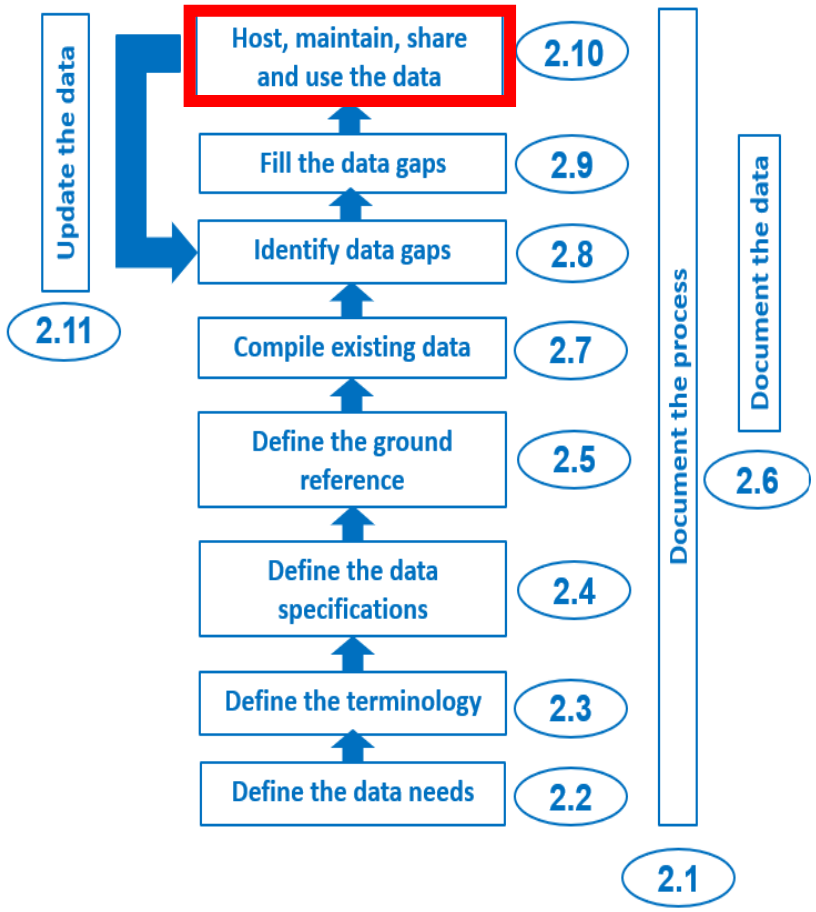
We will also cover this during Session 15

Host, manage, share, and use data

Once the data has reached the appropriate quality level, it can be used to generate the GIS-based products defined at the beginning of the process.

The data in question should ideally be properly hosted and maintained to allow for any other use.

Depending on the restrictions attached to it, the data can be distributed to the targeted and/or interested parties. The same goes for the information and/or information products which are then generated based on this data.



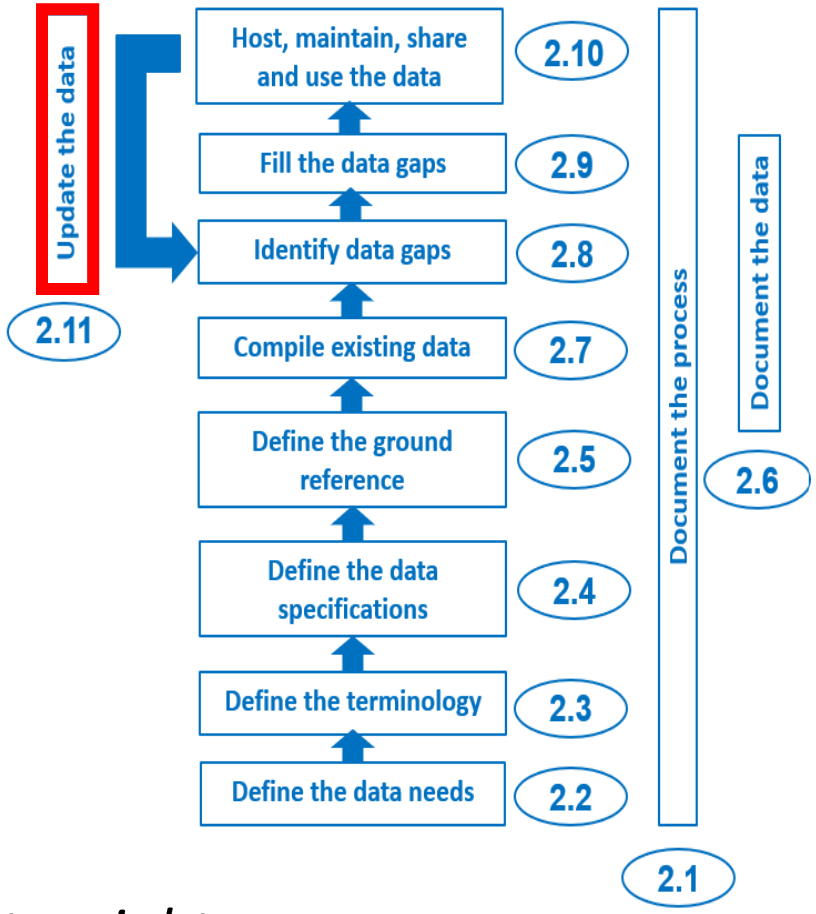
Update data

Like any other type of data, geospatial data must be updated regularly to reflect the changes that occur over time

Such an update must be done by following an established update mechanism

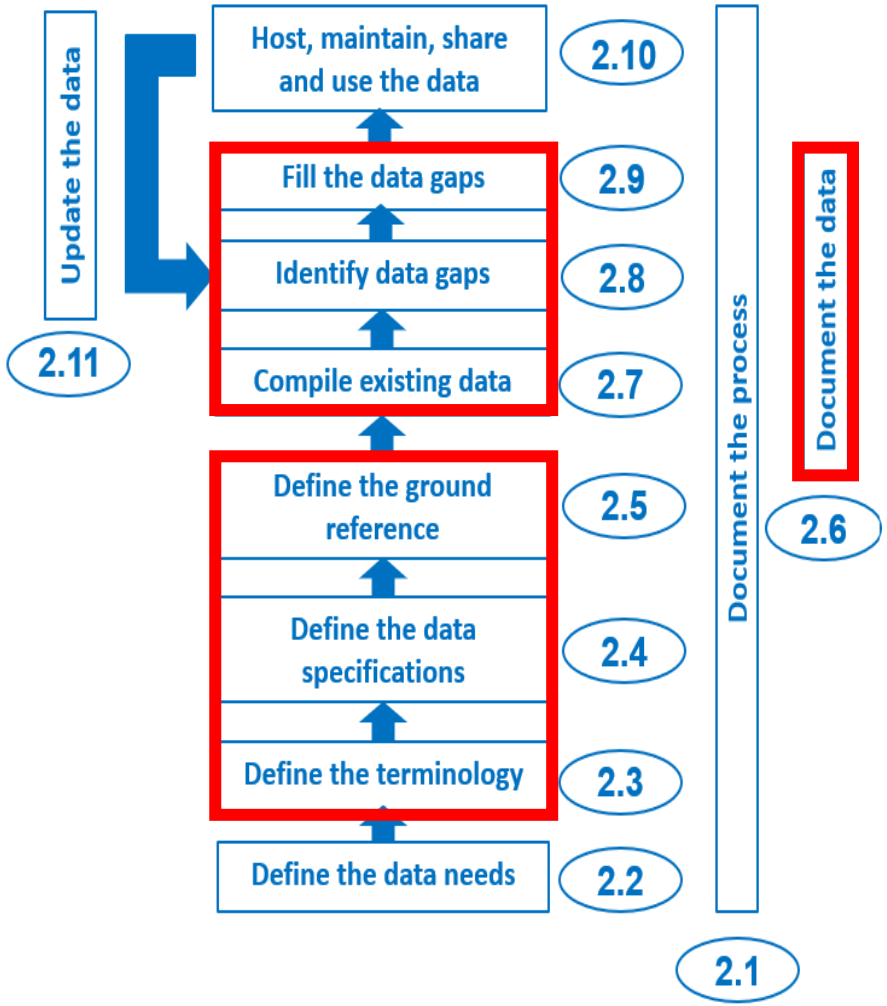
This may require the collection or extraction of additional data as well as updating the GIS-based products generated with it

➡ Generates a loop



Note: The need to generate new GIS-based products might require for the process to restart from step 2.1

The geospatial data management cycle



The remaining sessions for today will focus in more details on specific steps of the geospatial data management cycle (in red)

For the other steps please refer to:

1. The Health GeoLab guidance to improve the management and use of geospatial data and technologies in health (https://healthgeolab.net/resources/reference-materials/)
2. The HIS geo-enabling training course (https://healthgeolab.net/resources/his-geo-enabling-course/)

Free online resources

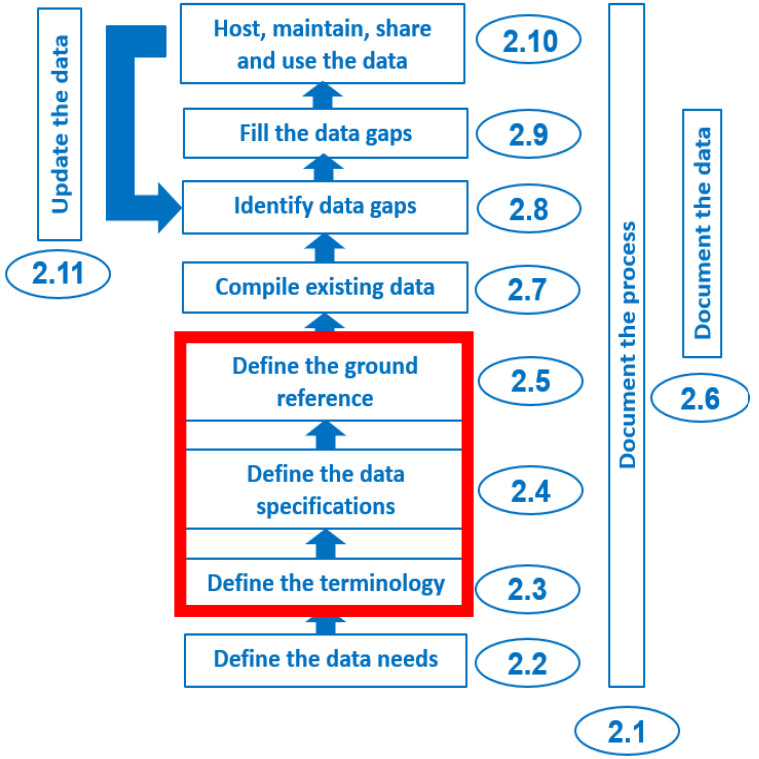
Geo-enabling the Health Information System, programs or interventions training workshop for Asia Pacific

...and beyond

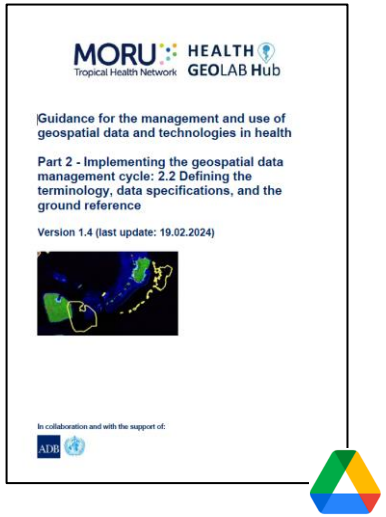
Session 13: Implement the geospatial data management cycle – Define the terminology, data specifications and the ground reference

Define the terminology, data specifications and the ground reference

Once the data needs have been defined, the next three steps in the geospatial data management cycle (define the terminology, the data specifications and the ground reference) provide the foundation that will ensure the quality of the data being collected, generated, extracted and used during the rest of the cycle



Topics of one of the HGL's guidance document (2.2)



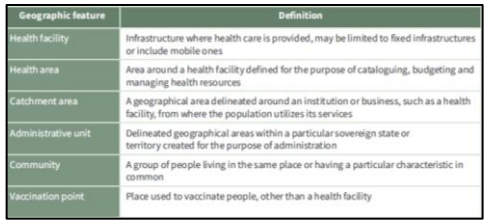
http://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_2.pdf

Define the terminology

Agreeing on and using a common terminology ensures that all actors involve understand each other by speaking the same language!

2 main glossaries to have:

1. Definition for each geographic feature (Session 6, Module 2)
2. Geospatial data and technologies terminology. Examples:
   - Esri GIS Dictionary: https://support.esri.com/en-us/gis-dictionary



Geographic feature	Definition
Health facility	Infrastructure where health care is provided, may be limited to fixed infrastructures or include mobile ones
Health area	Area around a health facility defined for the purpose of cataloguing, budgeting and managing health resources
Catchment area	A geographical area delineated around an institution or business, such as a health facility, from where the population utilizes its services
Administrative unit	Delineated geographical areas within a particular sovereign state or territory created for the purpose of administration
Community	A group of people living in the same place or having a particular characteristic in common
Vaccination point	Place used to vaccinate people, other than a health facility

It is also important to have a dictionary/glossary specific to the HIS, program or intervention being geo-enabled

https://drive.google.com/file/d/1jj779zww4herWOESAd9mXqVE1YfQehtH/view?usp=sharing;
https://healthgeolab.net/DOCUMENTS/Guidance_Common_Geo-registry_Ve2.pdf

Define the data specifications and ground reference

Addressing public health issues requires data to be of quality and this across the six (6) dimensions of data quality:

1. Completeness: No data gap
2. Uniqueness: No duplicates
3. Timeliness: Reality from the required point in time
4. Accuracy: Correctness
5. Validity: Conform to the defined format, type, range,...
6. Consistency: Absence of apparent contradictions

- Data quality is being assessed and improved based on measurable criteria captured in the data specifications (geospatial and statistical data, georeferenced master lists) and using remote sensing images and the master lists as ground reference when it comes to geospatial data
- Applying these criteria and ground reference ensure the same minimum level of quality for the data used by an organization.

Define the data specifications – Geospatial data

When it comes to geospatial data, these criteria should at least cover:

- Validity:
  - V.1 Geographic coordinate system and map projection
  - V.2 Geographic extent of the area being covered
  - V.3 Language(s) included in the data
  - V.4 File format(s) for sharing data
  - V.5 Metadata profile
- Accuracy:
  - A.1 Scale (vector/raster layers)
  - A.2 Spatial resolution (raster layers)
  - A.3 Positional accuracy (vector/raster layers)
  - A.4 Positional accuracy (GNSS reading)
  - A.5 Positional precision (GNSS reading)
- Timeliness:
  - T.1 Period for which the data is being considered as relevant

Define the data specifications – Geospatial data

When it comes to completeness and uniqueness :

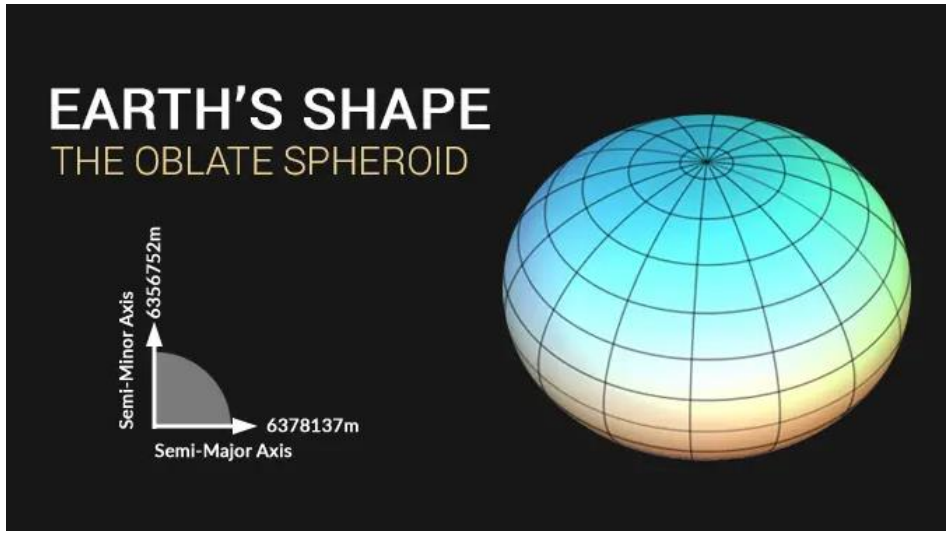
1. For vector format geospatial data associated to a master list:
   - Completeness:
     a. Records: When applicable, the data contains all the active records included in the corresponding master list
     b. Data elements: When applicable, a value is available for all the data elements included in the corresponding master list
   - Uniqueness:
     a. No duplicate geographic objects based on the master list as ground reference
2. For other geospatial data:
   - Completeness: All the features existing in the reality are included in the dataset (e.g. road network) or area completely covered (e.g. DEM)
   - Uniqueness: No duplicated geographic objects

➡ Consistency is achieved by reaching the criteria for all the other data quality dimensions

Define the data specifications – Geospatial data

V.1 Geographic Coordinate System

System in which geospatial data is defined by a 3-D surface and measured in latitude and longitude.



Earth's shape is not a perfect sphere (21,385 m difference between the two semi-axis)

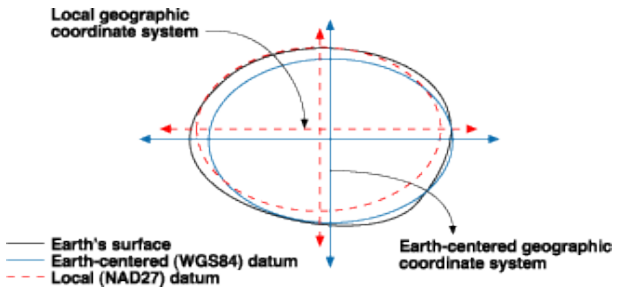
Different models exists, each of them trying to reproduce the earth's shape based on:

- Angular units: The unit of measurement on a sphere or a spheroid, usually degrees.
- Prime meridian: The zero meridian (0°) used as the reference from which longitude east and west is measured.
- Datum: Defines the position of the spheroid relative to the center of the earth.
- Spheroid: The reference spheroid for the coordinate transformation.

Define the data specifications – Geospatial data

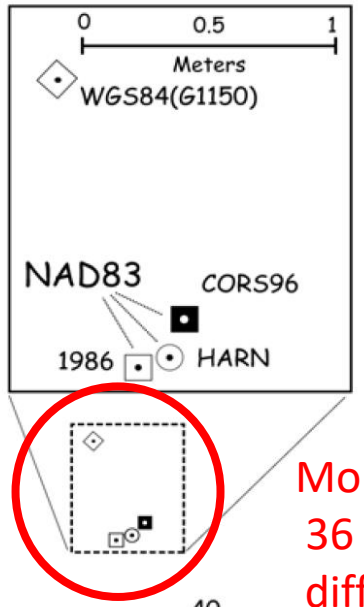
V.1 Geographic Coordinate System

IMPORTANT: All the datasets combined on a map must present the same Geographic Coordinate System



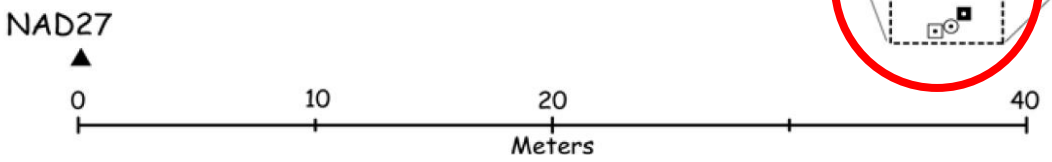
Impact of using different datum

Datum	Longitude (W)	Latitude(N)	Shift(m)
NAD27	74° 12' 3.86927"	40° 47' 0.76531"	36.3
NAD83(1986)	74° 12' 2.39240"	40° 47' 1.12726"	0.04
NAD83(HARN)	74° 12' 2.39069"	40° 47' 1.12762"	0.05
NAD83(CORS96)	74° 12' 2.39009"	40° 47' 1.12936"	0.95
WGS84(G1150)	74° 12' 2.39720"	40° 47' 1.15946"	



More than 36 metres difference

Correct Datum



Define the data specifications – Geospatial data

V.1 Projected Coordinate System

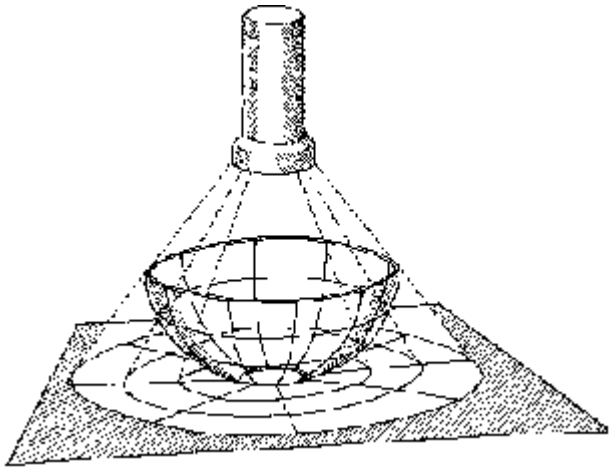
System in which geospatial data is defined by a flat 2-D surface and can be measured in units of meters and feet.

→ **Geographic Coordinate System + Map projection**

Map projection

A method by which the curved surface of the earth is portrayed on a flat surface

- The systematic transformation of points on the Earth's surface to corresponding points on a plane (flat) surface
- The earth is 3D but maps need to be flat!
- This requires distortion of some parts of the map.



Note: When you don't use a map projection, the data is unprojected. In this case, the GIS software treats the coordinate values as if they are linear values to display the data (Like the X and Y values on a graph).

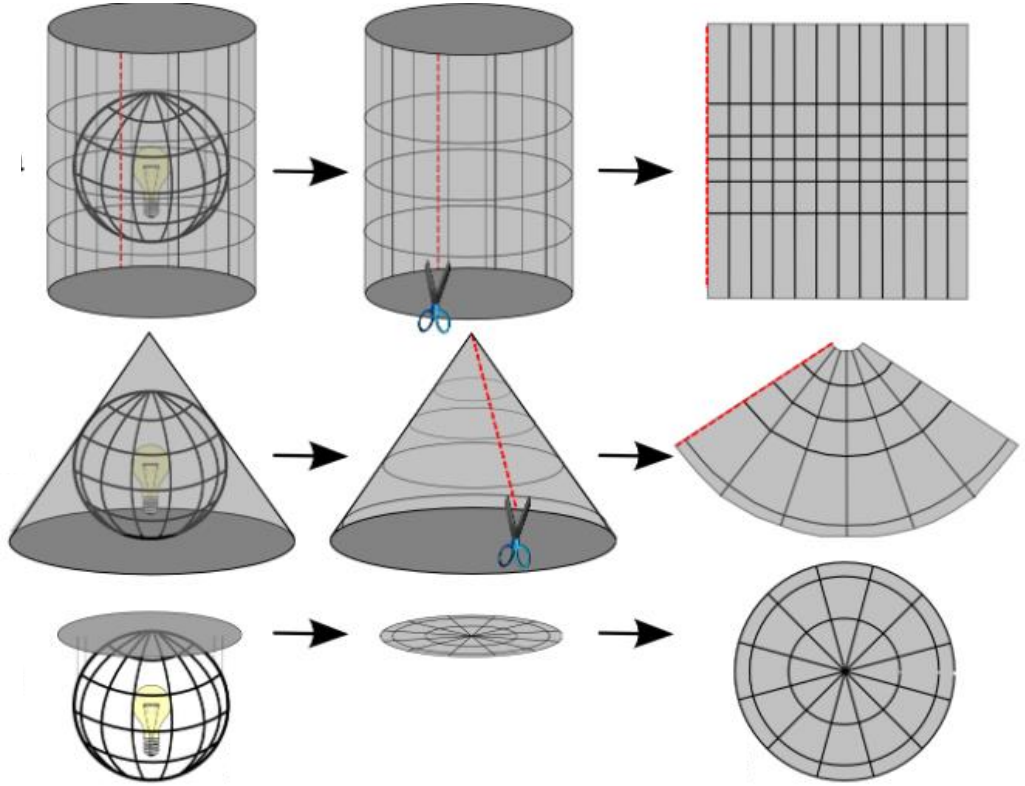
Define the data specifications – Geospatial data

V.1 Map Projection – Basic projection techniques

Cylindrical

Conical

Azimuthal



http://www.icsm.gov.au/mapping/about_projections.html#types

Define the data specifications – Geospatial data

V.1 Map Projection – Basic Projection Types

Each projection type preserves a particular relationship or characteristic:

- **Equal-Area —** correctly shows the size of features
- **Conformal —** correctly shows the shape of features
- **Equidistant —** correctly shows the distance between two features
- **True Direction —** correctly shows the compass direction between two features

➡ A printed map cannot be at the same time equal-area or conformal – it can only be one or the other, or neither.

➡ A map projection is to be chosen based on the needs before being printed

Note: GIS software perform corrections on the fly which allows for these different types of measurements to correspond to the reality

Define the data specifications – Geospatial data

V.1 Map Projection – Examples

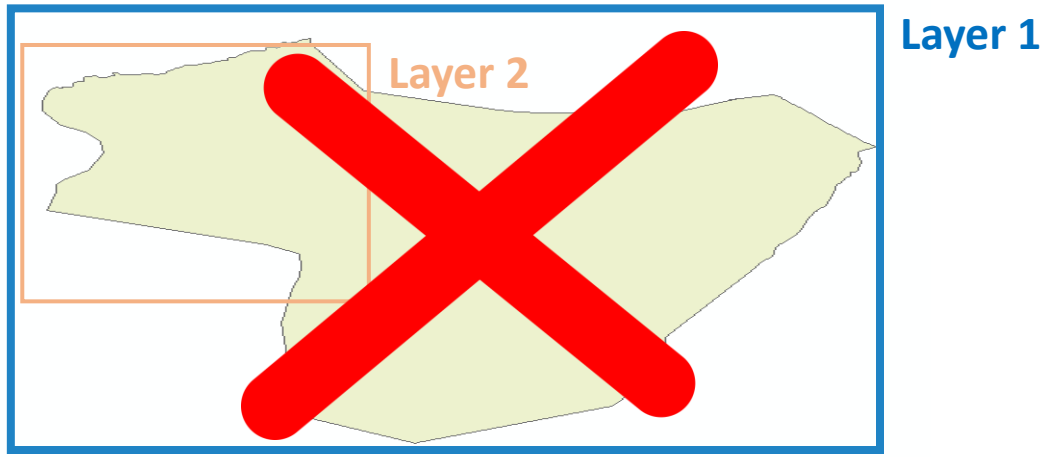
Projection	Technique	Type	Comment
Equirectangular	Cylindrical	Equidistant	Simplest geometry; distances along meridians are conserved. Plate carrée: special case having the equator as the standard parallel.
Lambert cylindrical equal-area	Cylindrical	Equal-area	
Universal Transverse Mercator (UTM)	Cylindrical	Conformal	Divides the Earth into sixty zones, each being a six-degree band of longitude
Robinson	Pseudocylindrical	Compromise (neither equal-area nor conformal)	Used to create global maps

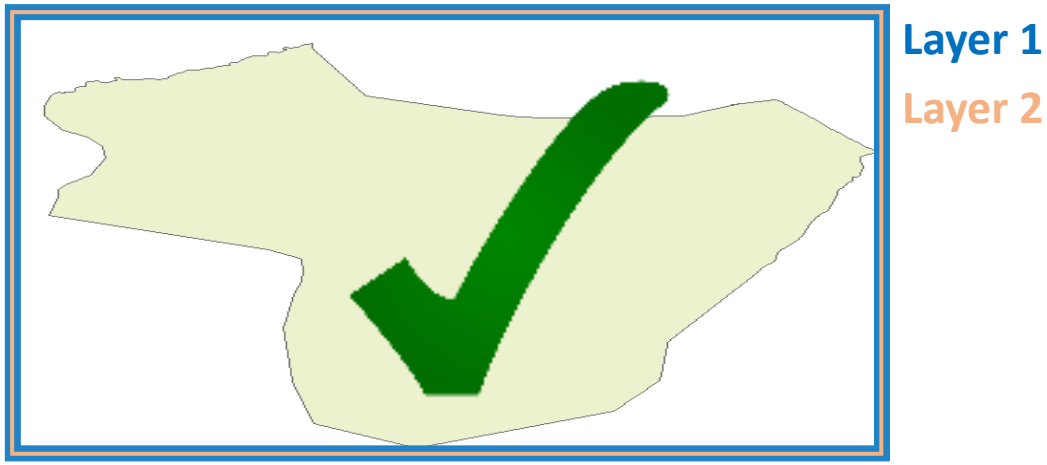
https://en.wikipedia.org/wiki/List_of_map_projections

Define the data specifications – Geospatial data

V.2 Geographic extent of the area being covered

Once all the layers presenting the same geographic coordinate system and map projection you want to make sure that they are also covering the full extent of the area of interest





Define the data specifications – Geospatial data

V.3 Language

- National and international language (e.g. English) using the Unicode text encoding standard

V.4 Geospatial and attribute data format (most used)

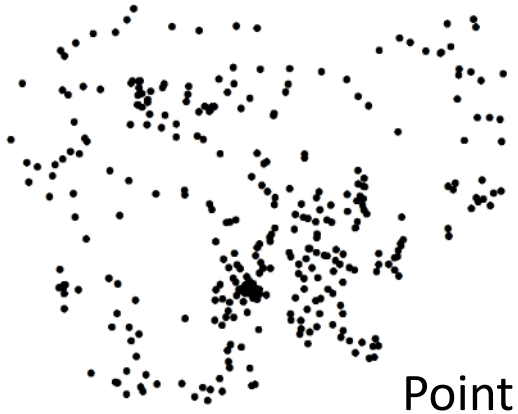
- Vector
  - Shapefile (actually composed of 3 to 8 files)
  - GeoJSON (QGIS)
- Raster
  - Georeferenced: Geotiff or GRID
  - Not georeferenced: .jpeg, .png, etc.
- Tabular
  - Spreadsheets: .xls, .dbf
- Combined vector/raster/tabular
  - Geodatabases

Recommended

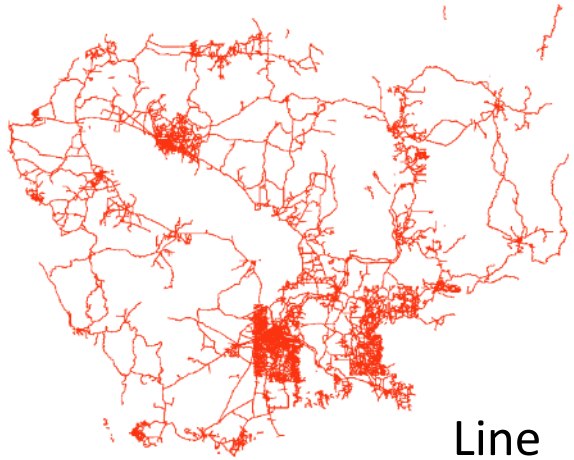
Geospatial data formats

V.4 Geospatial and attribute data format (most used)

Vector format



Point



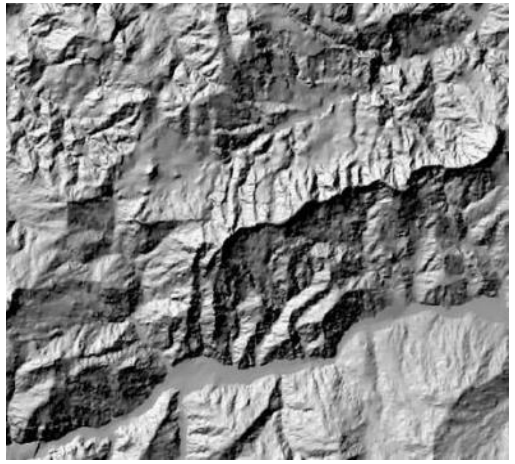
Line

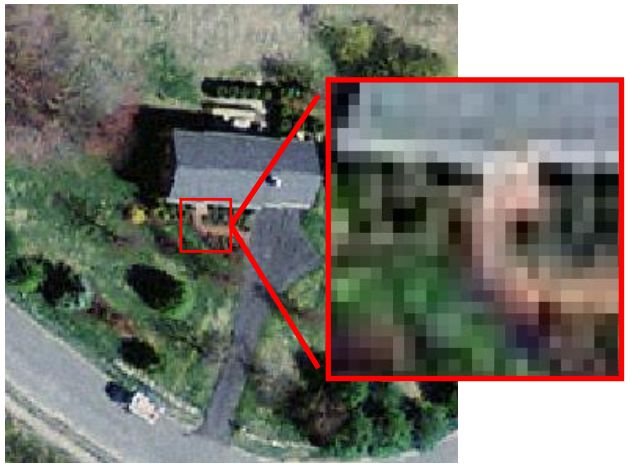
Polygon

Raster format

Grid

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	3	8	1	0	9	8	6	7	2	7	4	3	3	2	1	7	9	7	0	2
2	8	4	0	1	7	5	2	1	6	7	5	2	9	9	9	6	10	4	4	8
3	5	2	4	9	5	10	1	6	8	5	10	3	9	8	5	5	1	8	3	9
4	9	6	9	0	9	2	1	9	5	1	1	7	7	4	2	10	0	5	8	2
5	8	4	2	7	1	0	3	4	8	6	8	3	7	7	1	10	8	6	3	2
6	3	10	6	2	0	1	0	4	5	9	4	8	4	9	10	2	6	7	4	10
7	1	7	6	10	8	1	6	1	9	5	9	2	7	3	6	5	1	3	2	6
8	2	4	5	10	3	9	4	1	8	8	9	4	2	6	6	8	4	5	7	10
9	5	4	10	3	1	5	3	3	6	0	0	0	1	1	2	7	10	6	5	6
10	1	1	6	3	6	7	5	7	3	9	3	6	4	1	10	2	3	3	2	7
11	4	4	10	7	3	5	3	1	10	10	0	9	1	5	7	9	1	5	8	2
12	9	2	3	6	8	2	0	0	8	0	5	4	7	3	1	7	1	3	3	4
13	7	8	10	10	1	2	2	5	4	4	3	4	5	3	5	4	10	8	9	5
14	2	7	1	10	10	2	5	7	10	1	6	4	5	3	3	2	2	5	1	1
15	7	2	1	5	2	5	3	7	0	6	9	3	3	10	8	4	2	7	1	4
16	0	5	9	10	3	3	0	7	3	9	2	4	10	8	2	10	8	2	1	6
17	1	9	7	1	9	4	5	6	8	6	9	9	9	6	7	1	2	6	0	4
18	7	3	1	8	9	8	6	7	5	4	7	3	1	10	9	4	3	1	7	1
19	7	4	6	4	9	9	5	6	9	6	5	3	1	6	10	4	8	3	2	4





Define the data specifications – Geospatial data

V.5 Metadata

Information that describes the content, quality, condition, origin, and other characteristics of data or other pieces of information.

- Data about the data
- Allows user to find out if the data is appropriate for its intended purpose
- Applies to master lists, geospatial and statistical data

Different metadata standards exists (FGDC, ISO) and a metadata profile needs to be generated out of one of these standards

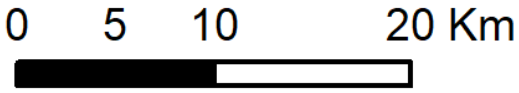
- Will be covered in more details during Session 14

Define the data specifications – Geospatial data

A.1 Scale

The ratio or relationship between a distance or area on a map and the corresponding distance or area on the ground, commonly expressed as a fraction or ratio

A map scale of 1/100,000 or 1:100,000 means that one unit of measure on the map equals 100,000 of the same unit on the earth surface.



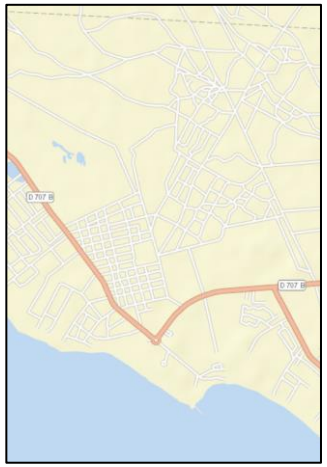
Large scale map versus small scale map

A large-scale map has a smaller ratio (1:10,000 or 1:25,000) and would contain more details such as streets and building footprints.

Small scale

Medium scale

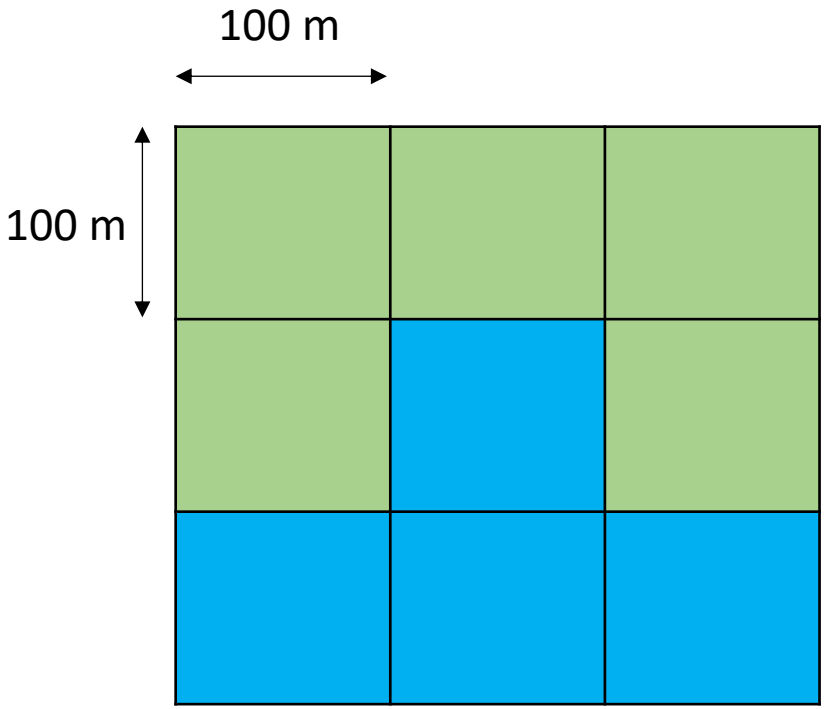
Large scale



Define the data specifications – Geospatial data

A.2 Resolution

The dimensions represented by each cell or pixel in a raster format geospatial dataset.



Scale Range	Raster resolution (m)
1:1 - 1:10,000	0.0005- 5
1:50,000 - 1:100,000	25- 50
1:250,000 - 1:500,000	125- 250
1:750,000 - 1:1,000,000	375- 500
1:1,500,000 - 1:2,000,000	750- 1,000
1:5,000,000 - 1:10,000,000	2,500- 5,000
1:25,000,000 -1:50,000,000	12,500- 25,000

Relation between scale and resolution

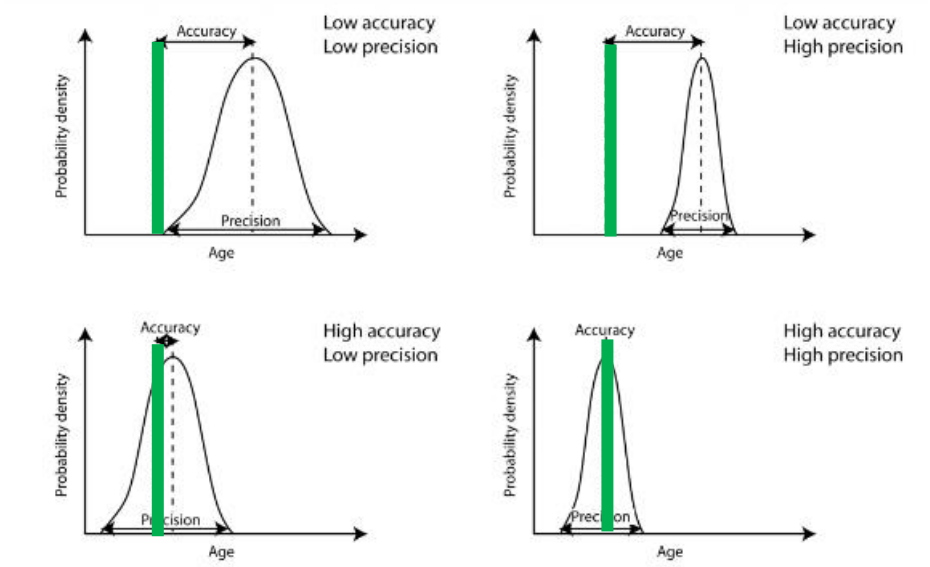
Tobler W. (1987): Measuring Spatial Resolution, Proceedings, Land Resources Information Systems Conference, Beijing, pp. 12-16

Define the data specifications – Geospatial data

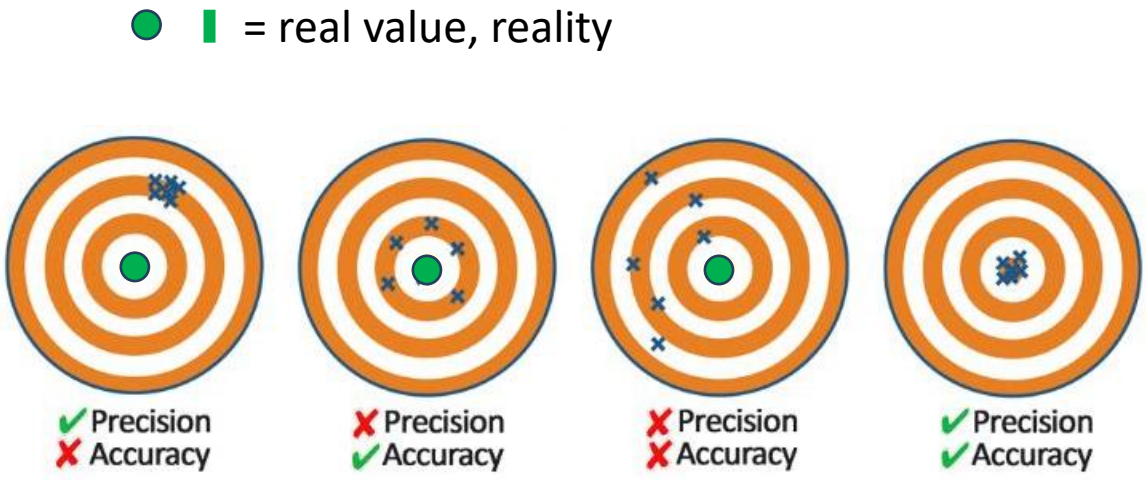
A.3 & A.4 Positional accuracy, and A.5 Precision

Positional accuracy: Quantifiable value that represents the positional difference between a geospatial layer and reality (proximity of measurements to the real value)

Precision: The number of significant digits used to store numbers, particularly coordinate values (measurement of dispersion)



Indicator



Geographic coordinates

Define the data specifications – Geospatial data

Precision (A.5)

At the equator: 360 º → 40,075 km

1 º ≈ 111,320 m

Number of captured digits	Example (Longitude)	Maximum potential error (m)	Precision level
1	120.9	11,132	
2	120.93	1,113	Kilometre
3	120.037	111	Hectometre
4	120.9376	11	Decametre
5	120.93761	1	Metre

Recommended

- During data collection in the field (GNSS enabled devices)
- When generating or extracting vector format geospatial data (precision level of vertices)

Define the data specifications – Geospatial data

Scale and positioning accuracy (A.1, A.3 et A.4)

United States Geological Survey mapping standards: "requirements for meeting horizontal accuracy as 90 per cent of all measurable points must be within 1/30th of an inch for maps at a scale of 1:20,000 or larger, and 1/50th of an inch for maps at scales smaller than 1:20,000."

Classification	Map examples	Range examples	Expected positional accuracy (m)
Large scale	Village, town or sub national level map	1:1 - 1:10,000	0 - 8
		1:50,000 - 1:100,000	26 - 52
		1:250,000 - 1:500,000	130 - 259
Medium scale	Country map	1:750,000 - 1:1,000,000	389 - 518
		1:1,500,000 - 1:2,000,000	777 - 1,036
Small scale	World map	1:5,000,000 - 1:10,000,000	2,591 - 5,182
		1:25,000,000 - 1:50,000,000	12,954 - 25,908

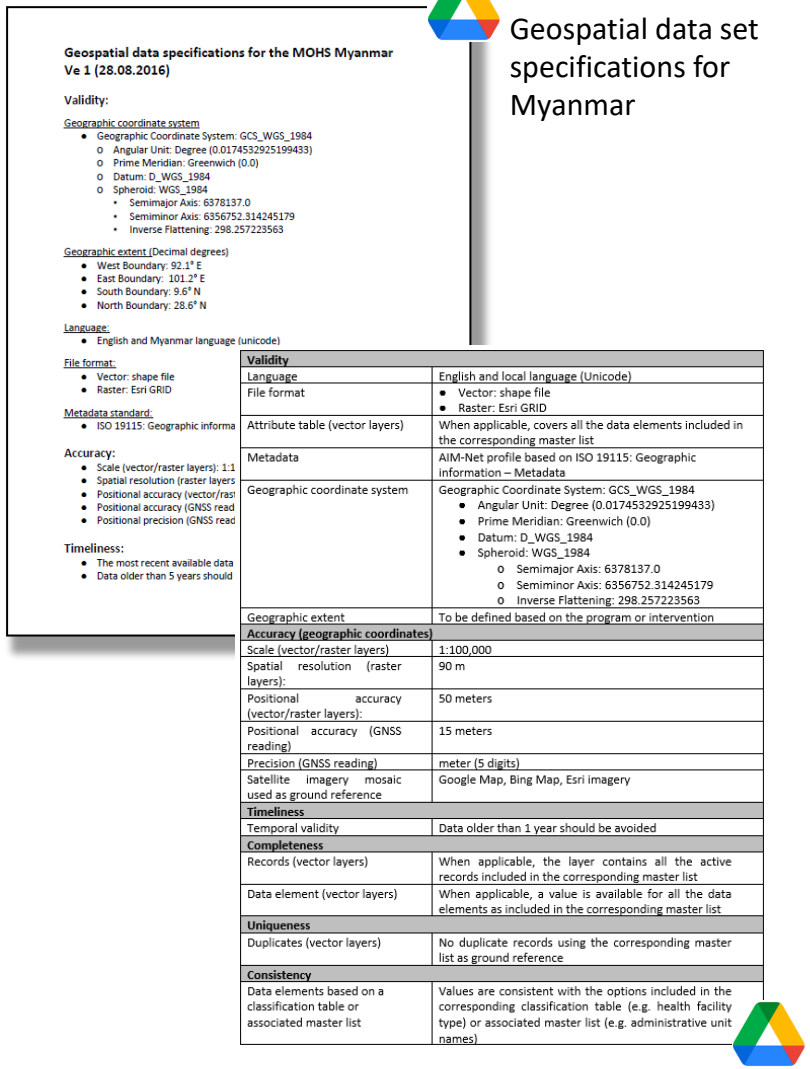
➡ Similar relationship than between scale and resolution

http://www.colorado.edu/geography/gcraft/notes/error/error_f.html

Define the data specifications – Geospatial data

Summary

- The purpose behind the use of geospatial data will guide the choice of a specific scale of work.
- This scale will directly influence the positional accuracy and spatial resolution that should be used when compiling, collecting, extracting or using geospatial data.
- The highest positional accuracy and precision possible should be sought, including when using GNSS-enabled devices, to allow for the largest use possible of the resulting data.
- A precision level down to the meter (5 digits in decimal degrees) is recommended



Geospatial data specifications for the MOHS Myanmar
Ve 1 (28.08.2016)

Validity:

Geographic coordinate system
- Geographic Coordinate System: GCS_WGS_1984
  - Angular Unit: Degree (0.0174532925199433)
  - Prime Meridian: Greenwich (0.0)
  - Datum: D_WGS_1984
  - Spheroid: WGS_1984
    - Semimajor Axis: 6378137.0
    - Semiminor Axis: 6356752.314245179
    - Inverse Flattening: 298.257223563

Geographic extent (Decimal degrees)
- West Boundary: 92.1° E
- East Boundary: 101.2° E
- South Boundary: 9.6° N
- North Boundary: 28.6° N

Language:
- English and Myanmar language (unicode)

File format:
- Vector: shape file
- Raster: Esri GRID

Metadata standard:
- ISO 19115: Geographic informa

Accuracy:
- Scale (vector/raster layers): 1:1
- Spatial resolution (raster layers
- Positional accuracy (vector/ras
- Positional accuracy (GNSS read
- Positional precision (GNSS read

Timeliness:
- The most recent available data
- Data older than 5 years should

Validity	
Language	English and local language (Unicode)
File format	• Vector: shape file • Raster: Esri GRID
Attribute table (vector layers)	When applicable, covers all the data elements included in the corresponding master list
Metadata	AIM-Net profile based on ISO 19115: Geographic information – Metadata
Geographic coordinate system	Geographic Coordinate System: GCS_WGS_1984 • Angular Unit: Degree (0.0174532925199433) • Prime Meridian: Greenwich (0.0) • Datum: D_WGS_1984 • Spheroid: WGS_1984 o Semimajor Axis: 6378137.0 o Semiminor Axis: 6356752.314245179 o Inverse Flattening: 298.257223563
Geographic extent	To be defined based on the program or intervention
Accuracy (geographic coordinates)	
Scale (vector/raster layers)	1:100,000
Spatial resolution (raster layers):	90 m
Positional accuracy (vector/raster layers):	50 meters
Positional accuracy (GNSS reading)	15 meters
Precision (GNSS reading)	meter (5 digits)
Satellite imagery mosaic used as ground reference	Google Map, Bing Map, Esri imagery
Timeliness	
Temporal validity	Data older than 1 year should be avoided
Completeness	
Records (vector layers)	When applicable, the layer contains all the active records included in the corresponding master list
Data element (vector layers)	When applicable, a value is available for all the data elements as included in the corresponding master list
Uniqueness	
Duplicates (vector layers)	No duplicate records using the corresponding master list as ground reference
Consistency	
Data elements based on a classification table or associated master list	Values are consistent with the options included in the corresponding classification table (e.g. health facility type) or associated master list (e.g. administrative unit names)

Geospatial data set specifications for Myanmar

Define the data specifications – Georeferenced Master list

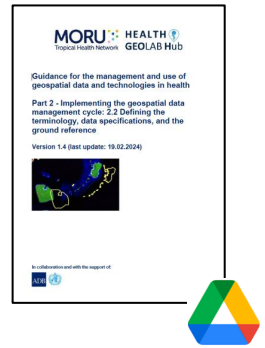
Validity	
Language	English and local language (Unicode)
File format	MS Excel
Data catalog	Covers the minimum set of data elements included in the corresponding master list data dictionary
Metadata	Covers the minimum set of fields included in the corresponding master list template
Geographic coordinate system	Geographic Coordinate System: GCS_WGS_1984 • Angular Unit: Degree (0.0174532925199433) • Prime Meridian: Greenwich (0.0) • Datum: D_WGS_1984 • Spheroid: WGS_1984 o Semimajor Axis: 6378137.0 o Semiminor Axis: 6356752.314245179 o Inverse Flattening: 298.257223563
Accuracy (geographic coordinates)	
Positional accuracy	15 meters
Positional precision	meter (5 digits)
Timeliness	
Temporal validity	Data older than 1 year should be avoided
Completeness	
Records	All the currently active records are included in the list
Data element	A value is available for all the data elements included in the data catalog
Uniqueness	
Duplicates	No duplicate records
Consistency	
Data elements based on a classification table or associated master list	When applicable, data elements values are consistent with the options included in the corresponding classification table (e.g. health facility type) or associated master list (e.g. administrative unit names)

Fake example

→ Data dictionary

→ Master list template

Same criteria as for the geospatial data



→ Classification tables

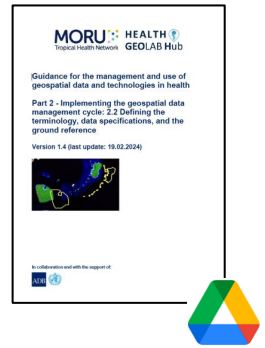
→ Other master lists

Define the data specifications – Statistical data

Validity	
Language	English and local language (Unicode)
File format	MS Excel
Data catalog	Covers all the data elements included in the file
Metadata	Covers the minimum set of fields included in the defined metadata profile
Timeliness	
Temporal validity	Data older than 1 year should be avoided
Completeness	
Records	Statistics available for all the active records included in the corresponding master list
Data element	A value is available for all the statistical indicators included in the data catalog
Uniqueness	
Duplicates	No duplicate records using the corresponding master list as ground reference
Consistency	
Values	Values are captured according to the indicator reference sheet (format, unit)

Fake example

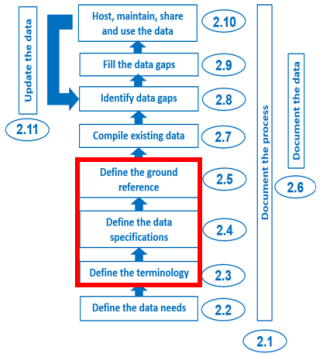
→ Metadata profile



→ Indicator reference sheet

AIM-Net disaster statistical indicator reference sheet	
Indicator name:	Number of affected people
Indicator code:	Nbr_Aff_People
Purpose/rationale:	Indicator used to monitor the number of affected people during a disaster event
INDICATOR DESCRIPTION	
Precise Definition(s):	Number of people requiring immediate assistance during a period of emergency (i.e. requiring basic survival needs such as food, water, shelter, sanitation and immediate medical assistance)
Unit of Measure:	Number of people
Data Type:	Integer
Disaggregated by:	Districts impacted by the disaster
INDICATOR COLLECTION INFORMATION	
Source:	District disaster management unit
Collection method	The district-level response teams use an online form designed by the National Disaster Management Agency. The data being collected through this form is automatically aggregated on the Disaster Response Dashboard accessible not only to the national and province-level EOC but also to the field response command center.
Collection frequency:	Daily during the first 7 days of the response and then weekly
DATA QUALITY ISSUES	
Known limitations:	The indicator is known to be under-estimating the number of affected peopl in the first 3 days of the response
CHANGES TO INDICATOR	
Changes to Indicator:	Until 2023, missing and deceased people were also included in this indicator. Since January 2024, these numbers are captured in separated indicators
Other Notes (optional):	The indicator reference sheet is accessible from the Disaster Response Dashboard
SHEET LAST UPDATED ON:	16.03.2024

Define the ground reference



The closest reference to the reality against which data quality can be measured

In the context of the HGL guidance, the ground reference is provided by the following :

- High resolution orthorectified remote sensing imagery (applies to geospatial data)
- Georeferenced master lists (applies to geospatial and statistical data)

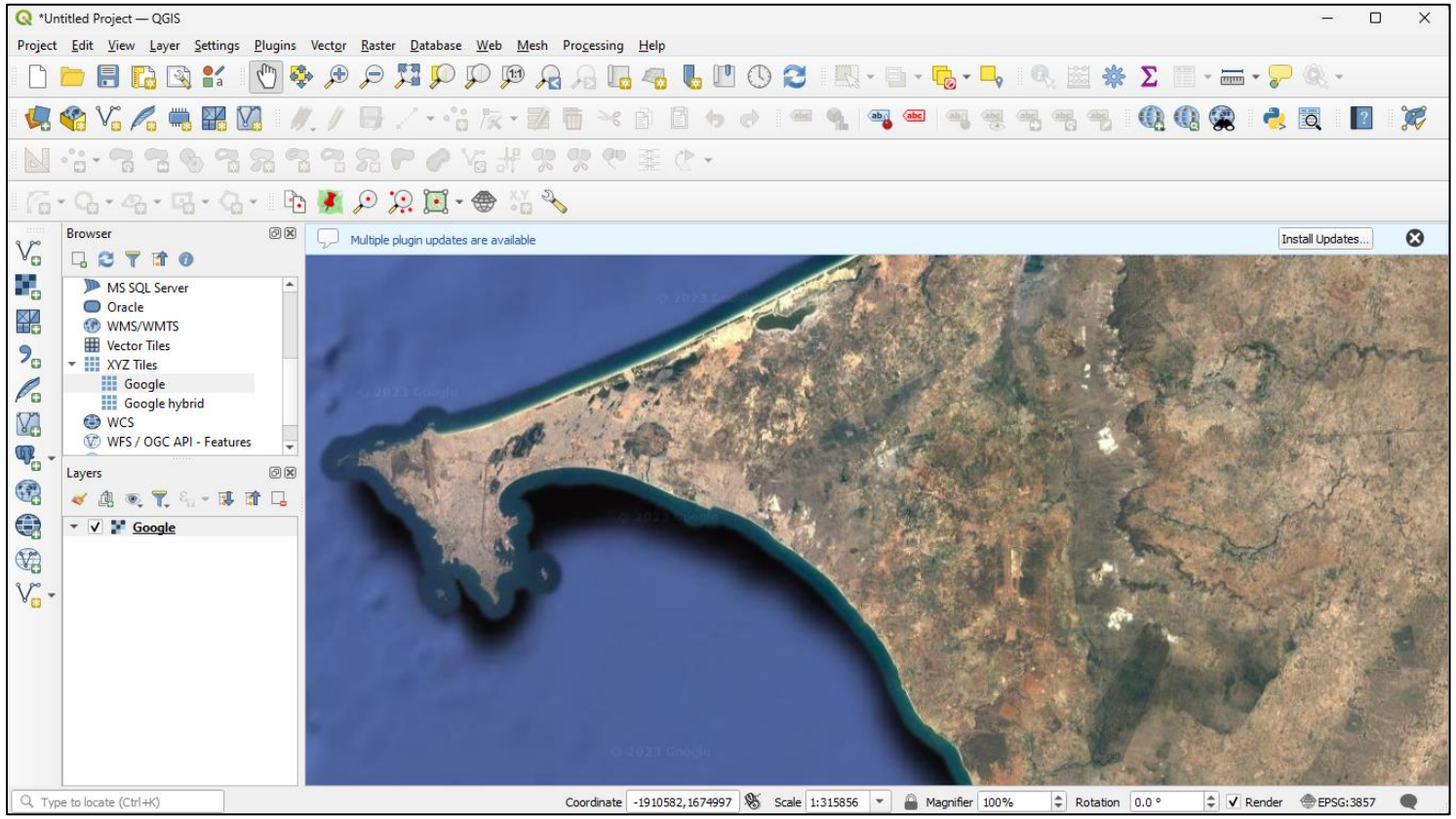
Data quality dimension	Ground reference	
	Remote sensing images	Master lists
Completeness	X	X
Uniqueness		X
Timeliness	X	X
Validity		X
Accuracy	X	X
Consistency	X	X

Emphasizes once again the key role that master lists are playing

Define the ground reference – Remote sensing imagery

Imagery collected by a satellite, an aircraft, or a drone

Medium to high resolution satellite imagery are now available for free through online applications such as Google Map or Bing Map or directly in GIS software through what we call a web mapping service.



We will come back to this when talking about data quality assessment in Session 15

Define ground references – Georeferenced master Lists

Unique, authoritative, officially curated by the mandated agency, complete, up-to-date and uniquely coded list of all the active (and past active) records for a given type of geographic feature/object (e.g. health facilities, administrative units, villages)

The information that allows to do the following for each of the records in the master list:

- Uniquely identify (unique identifier, name)
- Classify (type, ownership,…)
- Locate (address, administrative division, geographic coordinates)
- When it applies, contact (head name, phone number, email address,…)

Example for health facilities

Unique ID	Health facility name
HF0013	San Juan Referral Hospital

Health facility type	Owernship
Referral Hospital	Government

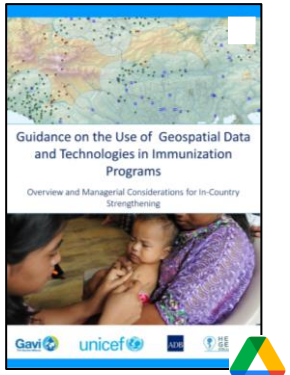
Address	Province code	Province Name	Latitude	Longitude
20, St Andreas Street	TLK01	Andustar	14.412830	121.033090

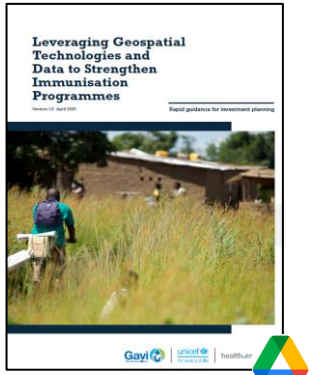
Head name	Head position	Phone number
Horm Mada	Director	+99 97 11477917

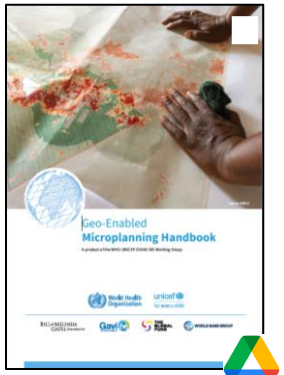
Any other data element is to be considered programmatic attributes and managed outside the master list

Define ground references – Georeferenced master Lists

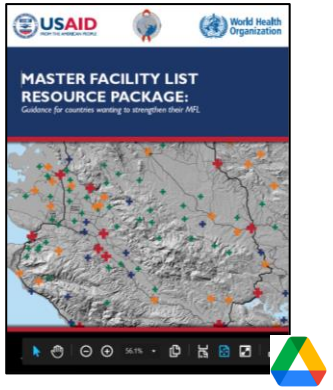
Key elements mentioned in several health program guidelines



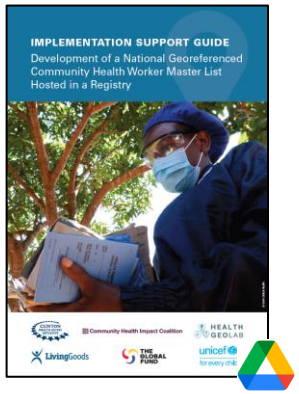




…the main subject of a growing number of master list-specific documents

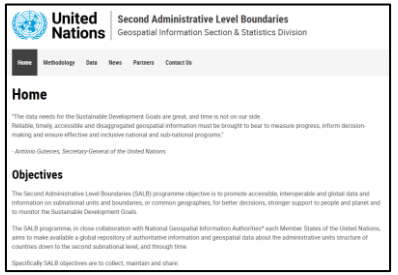


Administrative units



Health facilities

…and at the origin of global level initiatives



https://salb.un.org/en

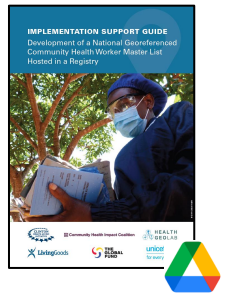
https://www.who.int/data/GIS/GHFD

Define ground references – Georeferenced master Lists

Key elements of a master list

1. Definition → Covered during Session 6 (Module 2)
2. Content
   - Data elements and dictionary
   - Coding scheme
   - Naming convention
   - Classification tables (types, Owner)
   - Locations (addresses, administrative divisions, geographic coordinates)
   - Contact information (establishment manager, phone number, etc.)
3. Registry or Common Geo-Registry (CGR) → Will be covered during Session 19 (Module 5)
4. Process

Georeferenced master Lists – Data dictionary



A collection of names, description, and attributes about the data elements that are being used or captured in a database or information system

Such a dictionary should contain the following for each data element:

Example of how the information for a set of data elements could be captured as part of the CHWML data dictionary.

Unique ID

Contextual definition: A unique code that identifies a CHW and distinguishes her/him from others

Applicability: All individuals included in the list

Suggested updating frequency: Not applicable (applied once)

Data type: String
Format: Alphanumeric
Maximum character length: Variable

Example values: Value — CHW0098758

Notes: This may be a government issued unique ID, or an MOH ID (see the Define the CHV for more recommendations regarding the definition of the unique ID coding scheme)

CHW Full Name (English and local language)

Contextual definition: The full name of the CHW in English and in the local language

Applicability: All individuals included in the list

Suggested updating frequency: As necessary (when a change in name occurs)

Data type: String
Format: Alphabetical
Maximum character length: Variable

Permissible values:

Value	Meaning	Range
Last name	Family name	Variable characters
Middle name	Middle name	Variable characters
First name	First name	Variable characters
Other name	Other name	Variable characters

Example values:

English: Henri Bernard Steelwood
Local language:

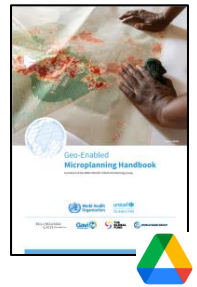
Notes: First, middle, family and other name should be captured in full and as much as possible in separated fields

- A clear contextual definition
- Its applicability (applied to all records in the list or only some of them)
- Its format (alphanumeric, numeric, date, other)
- Its maximum character length
- The values that the data element can take, especially when these are bound to a limited number of options (classification tables)
- Whether the data element should be considered as mandatory when a new record is being added to the master list (all the core data elements should be considered as mandatory by default).

→ Needs to be developed for each master list and must contain the information for all the data elements to be included in each of them

→ The list of data elements to be included is developed through a consultative and collaborative involving the concerned stakeholders

Georeferenced master list – Data elements

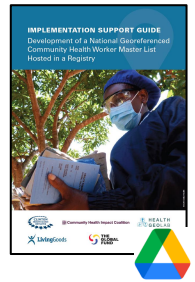


Examples of core data elements for different types of geographic features

Group	Example of data elements for different geographic features		
	Health facility	District	Vaccination post
Uniquely identify	⊙ Official unique identifier ⊙ Official name	⊙ Official unique identifier	⊙ Official unique identifier ⊙ Official name
Classify	⊙ Health facility type ⊙ Ownership	⊙ Not applicable	⊙ Vaccination post type ⊙ Vaccination strategy
Locate	⊙ Administrative unit in which health facility is located (unique identifier and name) ⊙ Address (e.g. street name and number) ⊙ Geographic coordinates (latitude and longitude)	⊙ Upper-level administrative unit in which district is located ⊙ Geometry stored in a GIS vector format layer (link with the master list through the unique identifier)	⊙ Administrative unit in which the vaccination post is located ⊙ Geographic coordinates (latitude and longitude)
Contact	⊙ Full name of health facility head ⊙ Phone number (health facility, head) ⊙ Email address (health facility, head)	⊙ Not applicable	⊙ Not applicable

Depends on how the concept of vaccination post is defined

Georeferenced master list – Data elements

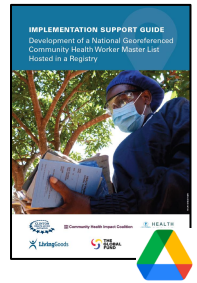


Core data elements for establishing a community health worker master list

Group	Data element	Definition	Considerations
Uniquely identify	Official unique ID	A unique code that officially identifies a CHW and distinguishes from others	This may be a government-issued unique ID, MOH ID, etc.
	CHW full name (English)	Complete official name of the CHW in English	First, middle and family names should be captured in full and, as much as possible, in separated fields taking into account that names follow different structures in different countries.
	CHW full name (official local language)	Complete official name of the CHW in the official local language	First, middle and family names should be captured in full and, as much as possible, in separated fields.
	CHW birth date	Date when the CHW was born	The date should follow an agreed upon format and be consistent across all the records in the list. The format in question should cover cases when only the birth year is known.
Classify	CHW gender	Self-identified CHW gender	The way gender is defined and classified depends on the country context. This should be defined in country and be inclusive.[52]
	CHW employment occupation category (CHW type)	Classification of the CHW by type	Ideally defined based on the country's official taxonomy of CHWs (e.g., Health Extension Worker, Community Health Agent, Community Based Volunteer). If an official taxonomy does not exist, this can be developed by the governance structure.
Locate[52]	CHW human settlement of residence (unique ID and name)	Unique identifier and name of the human settlement within which the CHW resides (community, village, etc.)	Each country might define the concept of human settlement differently. Information ideally sourced from the human settlement master list and stored in separated fields.
	CHW human settlement of residence (administrative structure)	Unique identifier and name of the administrative unit in which the CHW place of residence is located across the levels of the administrative structure (province, district, etc.)	Each country has a different administrative structure. Information ideally sourced from the human settlement master list.
	CHW human settlement of residence (geographic coordinates)	Latitude and longitude as well as associated information of the human settlement within which the CHW resides	It is important to identify which landmark within the human settlement has been used as reference to collect its geographic coordinates and if this respects potential data privacy regulations in place in the country Information ideally sourced from the human settlement master list (latitude, longitude and associated information stored in separated fields).
	CHW primary place of work (unique ID and name)	Unique identifier and name of the primary human settlement within which the CHW works	Each country might define the concept of human settlement differently. Ideally sourced from the human settlement master list and stored in separated fields.
	CHW primary place of work (administrative structure)	Unique identifier and name of the administrative unit in which the CHW primary place of work is located across the levels of the administrative structure (province, district, etc.)	Each country has a different administrative structure. Ideally sourced from the human settlement master list.
	CHW primary place of work (geographic coordinates)	Latitude and longitude as well as associated information of the primary human settlement within which the CHW works	Ideally sourced from the human settlement master list (latitude, longitude and associated information stored in separated fields).
	Health facility (unique ID and name)	Unique identifier and full name of the fixed health facility to which the CHW reports	Should include official unique ID and name of health facility Ideally sourced directly from the health facility master list and stored in different fields.
	Health facility (geographic coordinates)	Geographic coordinates as well as associated information of the fixed health facility to which the CHW reports	Ideally sourced directly from the health facility master list (latitude, longitude and associated information stored in separated fields)
Contact	CHW mobile number	Main mobile phone number at which the CHW can be reached	Mobile phone numbers should be formatted in an agreed upon form, verified and consistent across CHWs in the list.
Status	CHW employment status	Employment status of the CHW	The other information linked to the employment are themselves part of the additional data elements included in Table 4.
	CHW functional status	Whether or not the CHW has submitted a report within a given temporal interval (e.g., in the last month or quarter)	Ideally, the submission of a periodic report is an excellent way to assess the CHW functional status. Where supervision is carried out, the supervisor can inform about the functional status based on recent interactions with the CHW.

Core data elements necessary to uniquely identify, classify, locate and contact each Community Health Worker in the country as well as know their current contractual and reporting status.

Georeferenced master list – Data elements



Additional data elements for establishing a community health worker master list

Data element	Definition	Considerations
CHW photograph	Recent picture of the CHW	To be useful, the picture in question should comply to the general requirements for travel documents
CHW secondary place(s) of work	Unique ID and name of the human settlement(s) within which the CHW works but not as his/her primary place of work within the catchment area the CHW covers	Each country might define the concept of human settlement differently. Ideally the official unique identifier and name of the human settlement are coming from the settlement master list. The sum of the primary and secondary places of work corresponds to the extent of the CHW catchment area.
CHW education level	Level of formal education completed by the CHW	For example, primary school, secondary school, university. Should include provision for accredited/ recognized adult learning and tertiary institution courses that a CHW might have completed.
CHW training	Modules on which CHW has received training	Should list all the training modules the CHW has received training on, their duration of training and when training was conducted.
CHW accreditation status	Annual CHW accreditation examination performance	Should indicate the CHW performance or test scored for each exam taken, and accreditation received.
CHW contract duration	Start and end date of current CHW contract	According to the contractual status, especially when wages are provided, and based on the duration of the grant, some CHWs can have time-bound contract durations.
CHW contract type	Type of arrangement under which the CHW is contracted	To reflect the type of contract a CHW has with its managing organization, which can be formal, informal, permanent or temporary or short term.
CHW contracting authority	Official name of the organization that contracts the CHW	Depending on the context, the CHW can be contracted by a government department, specific donor, or NGO.

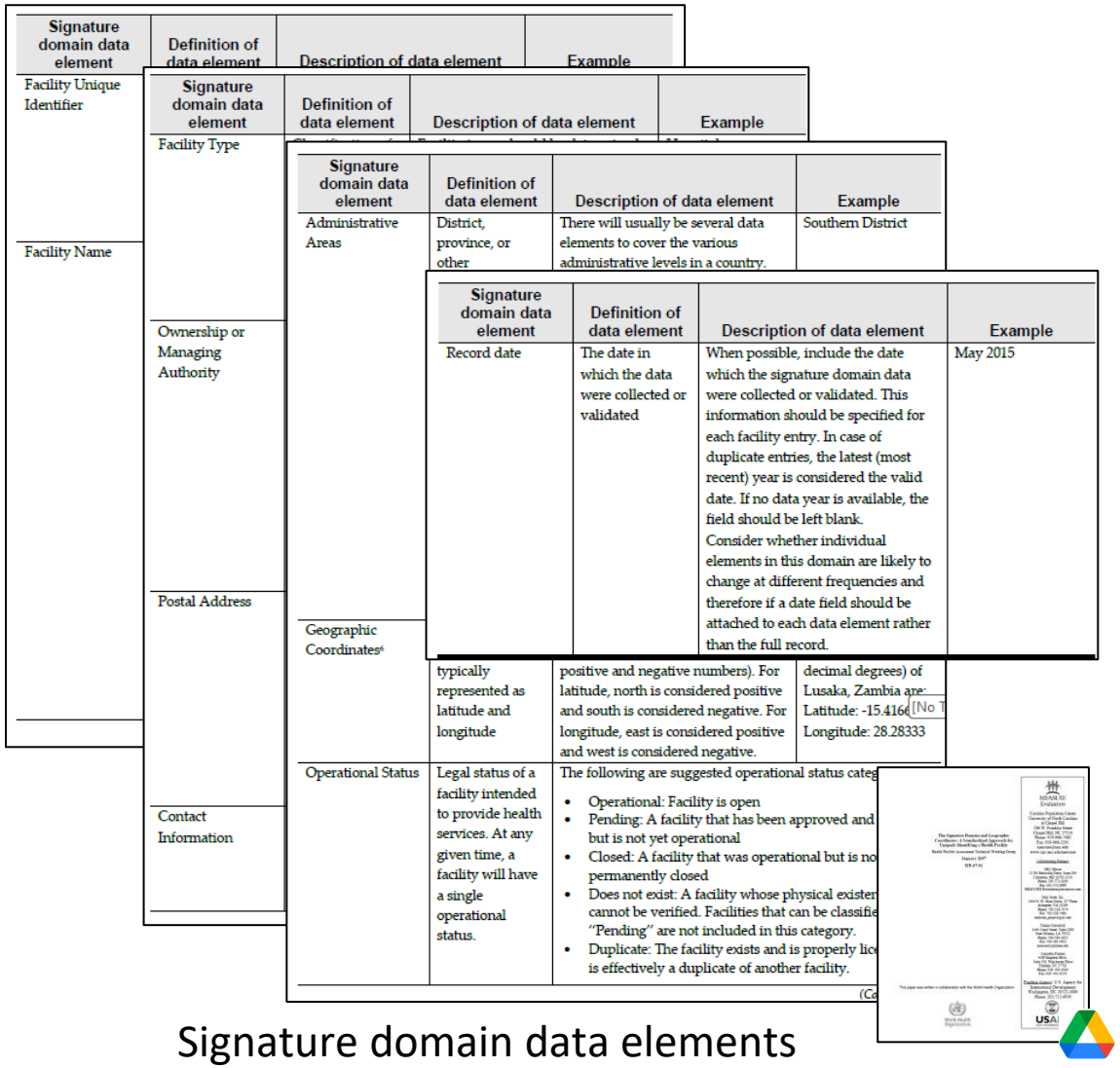
Data element	Definition	Considerations
CHW managing authority	Official name of the organization that manages the CHW	Should reflect the complete official name of the organization that manages the CHW.
Salary or other incentives received by the CHW	The salary amount, and/or other regular incentives (e.g., financial, material, recognition) provided to the CHW	CHW incentives/remuneration can differ according to the purpose of the CHW programme, the training received, roles and responsibilities, and the context in which they operate.
CHW supervisor	Unique ID and/or name of CHW's supervisor	Where CHWs are supervised by the upper level (health facility), the Unique ID of the supervisor drawn from HRH can be used.
CHW commodities	Treatment and diagnostics commodities with which the CHW is meant to be equipped to administer assigned health services	This may include rapid diagnostics (e.g., malaria RDT) and treatment (e.g., ORS, Zinc, ACT, antibiotics). Ideally this is sourced from a master list of essential medicines or supplies for the health system.
CHW equipment	Resources and tools with which a CHW is supposed to be equipped to help support their work functions	This may include a bicycle, mobile phone, job aids, backpack, Mid-Upper Arm Circumference (MUAC) tape, height board, etc.
CHW language	Languages spoken by CHW	Languages could be limited to a maximum of two per CHW, depending on country context.
Alternate phone number	Alternate phone number at which the CHW can be reached	Alternate phone numbers could be limited to a maximum of two per CHW.
Services offered	Services provided by the CHW	Can list services for which the CHW has been accredited.

Example of **additional data elements** to be prioritized in order to define which ones to be included in the minimum set for the master list

Note: the more data elements you include the more difficult it is to maintain the quality of the master list

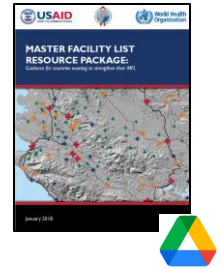
Georeferenced master list – Data elements

Health facility master list (HFML) vs Master Facility List (MFL)



Signature domain data element	Definition of data element	Description of data element	Example
Facility Unique Identifier			
Facility Name			
Facility Type			
Ownership or Managing Authority			
Postal Address			
Contact Information			
Administrative Areas	District, province, or other	There will usually be several data elements to cover the various administrative levels in a country.	Southern District
Geographic Coordinates[6]	typically represented as latitude and longitude	positive and negative numbers). For latitude, north is considered positive and south is considered negative. For longitude, east is considered positive and west is considered negative.	decimal degrees) of Lusaka, Zambia are: Latitude: -15.4166 [No T Longitude: 28.28333
Operational Status	Legal status of a facility intended to provide health services. At any given time, a facility will have a single operational status.	The following are suggested operational status categ • Operational: Facility is open • Pending: A facility that has been approved and but is not yet operational • Closed: A facility that was operational but is no permanently closed • Does not exist: A facility whose physical exister cannot be verified. Facilities that can be classifie "Pending" are not included in this category. • Duplicate: The facility exists and is properly lic is effectively a duplicate of another facility.	
Record date	The date in which the data were collected or validated	When possible, include the date which the signature domain data were collected or validated. This information should be specified for each facility entry. In case of duplicate entries, the latest (most recent) year is considered the valid date. If no data year is available, the field should be left blank. Consider whether individual elements in this domain are likely to change at different frequencies and therefore if a date field should be attached to each data element rather than the full record.	May 2015

Signature domain data elements



Service domain data element	Definition of data element	Description of data element	Example
Services Offered	Types of services offered by facility	A series of data elements listing key health services is included in the MFL; facilities are categorized as "Yes" providing the service or "No" not providing the service. Information must be adapted at a country level to include the package of services offered through the country's health system, and that are of interest to data consumers.	Family Planning Antiretroviral Therapy (ART) Labor and Delivery
Human Resources	Number of medical personnel by type	The categorization of health personnel is specific to the country. Types include, but are not limited to: physicians, non-physician clinicians, registered nurses, and registered midwives. For each type of health personnel, the facility reports the number available.	Number of Midwives: 4
Infrastructure	Number of inpatient and maternity beds and cots in facility	For the MFL, it is suggested that only information on impatient beds/cots (including maternity beds) be collected. Other equipment and infrastructure details should be collected through a separate health facility assessment (SAM, SARA, SPA, HFA, etc.). However, additional equipment and infrastructure data may be added to the MFL, if desired.	Number of Inpatient Beds: 15

Service domain data elements

Included in the Master Facility List concept

https://www.who.int/publications/i/item/-9789241513302

Georeferenced master list – Data elements

Health facility master list (HFML) vs Master Facility List (MFL)

Including service domain data elements in the HFML should be carefully considered because:

- Covering the 4 groups of data elements considered as being core to master lists (uniquely identify, classify, locate and contact) can already result in a long list of data elements to include in the data dictionary and therefore the master list
- Service domain data elements can be under the curation mandate of different entities and have their proper life cycle
- Service domain data elements can be more sensitive than the signature domain one

➡ Synchronization of the information system containing the service domain data elements with the HFML through the integration of the health facility unique identifier

Groupe	Libellé de l'élément de données	Description
Identifier de manière unique	HF_ID	Identifiant unique officiel de l'établissement de santé
	HF_N_E	Nom officiel ou commercial de l'établissement de santé (anglais)
	HF_N_M	Nom officiel ou commercial de l'établissement de santé (mongol)
	OP_DATE	Date à laquelle l'établissement de santé a été officiellement ouvert pour la première fois
Classifier	OP_STATUS	Statut opérationnel de l'établissement de santé
	HF_TYPE	Type d'établissement de santé épelé en toutes lettres selon la classification officielle
	HF_OWN	Propriété majeure épelée en toutes lettres selon la classification définie
Localiser	HF_ADD	Numéro et nom de la rue
	HF_ZIP	Code postal attaché à l'adresse de l'établissement de santé
	PR_C	Code officiel de la province dans laquelle l'établissement de santé est situé
	PR_N_E	Nom officiel de la province dans laquelle l'établissement de santé est situé (anglais)
	PR_N_M	Nom officiel de la province dans laquelle se trouve l'établissement (mongol)
	MU_C	Code officiel du District Municipal dans lequel l'établissement de santé est situé
	MU_N_E	Nom officiel du District Municipal dans lequel l'établissement de santé est situé (anglais)
	MU_N_M	Nom officiel du district municipal dans lequel l'établissement de santé est situé (mongol)
	KO_C	Code officiel du Khoroo dans lequel l'établissement de santé est situé

Example: Master list of health facilities – Mongolia (33 proposed signature domain data elements)

Groupe	Libellé de l'élément de données	Description
Localiser	KO_N_E	Nom officiel du Khoroo dans lequel l'établissement de santé est situé (anglais)
	KO_N_M	Nom officiel du Khoroo dans lequel l'établissement de santé est situé (mongol)
	LAT	Latitude de l'établissement de santé
	LONG	Longitude de l'établissement de santé
	COOR_SO	Source des coordonnées géographiques basées sur la table de classification définie
	COOR_ME	Méthode pour collecter les coordonnées géographiques basées sur la table de classification définie
	COOR_AC	Mesure qualitative du niveau d'exactitude des coordonnées géographiques basée sur la table de
Contact	HEAD_N	Nom complet du responsable de l'établissement de santé
	HEAD_POS	Position du responssable de l'établissement de santé
	LAND_NBR_1	Premier numéro de téléphone fixe officiel auquel l'établissement de santé peut être contacté
	LAND_NBR_2	Deuxième numéro de téléphone fixe officiel auquel l'établissement de santé peut être contacté
	MOB_NBR_1	Premier numéro de téléphone portable officiel auquel l'établissement de santé peut être contacté
	MOB_NBR_2	Deuxième numéro de téléphone portable officiel auquel l'établissement de santé peut être contacté
	FAX_NBR	Numéro de fax officiel auquel l'établissement de santé peut être contacté
	EMAIL_1	Adresse e-mail officielle de l'établissement de santé
	EMAIL_2	Adresse e-mail alternative utilisée par l'établissement de santé
	WEB	Site internet officiel de l'établissement de santé

Georeferenced master list – Classification tables

Table organizing and categorizing the values for a data element according to predefined criteria

Health Facility Type Code	Acronym	Health Facility Type in English	Health Facility type in Khmer	Definition	Source of the definition
6	NH	National Hospital	មន្ទីរពេទ្យជាតិ	Health facility used to train health personnel and undertake research studies, in addition to providing specialised referral services.	Modified from MOH (1998): Guide for developping operational districs
5	PH	Provincial Hospital	មន្ទីរពេទ្យ បង្អែកខេត្ត	Health facility under the administration and technical supervision of the PHD. In the Health Coverage Plan, provincial hospitals function as operational district referral hospitals and will also Q provide CPA services.	Modified from MOH (1998): Guide for developping operational districs
4	RH	Referral Hospital	មន្ទីរពេទ្យ បង្អែកស្រុក	Health facility providing the Complimentary Package of Activites (CPA	Modified from MOH (1998): Guide for developping operational districs
2	HC	Health Center	មណ្ឌលសុខភាព	Health facility delivering primary health care through the Minimum Package of Activites (MPA)	Modified from MOH (1998): Guide for developping operational districs
3	HC/B	Health Center with beds	មណ្ឌលសុខភាពមានគ្រែ	Health facility delivering primary health care through the Minimum Package of Activites (MPA) and having beds for patients	Modified from MOH (1998): Guide for developping operational districs
1	HP	Health Post	ប៉ុស្តិ៍សុខភាព	Health posts are located in remote areas and function as the lowest level within the district health system and thus the first point of contact with the population in low density provinces	Modified from MOH (1998): Guide for developping operational districs

- To be defined for each of the data elements where the values are limited to a set of options (example: type, ownership, accuracy, etc.)
- Ensures consistency across records

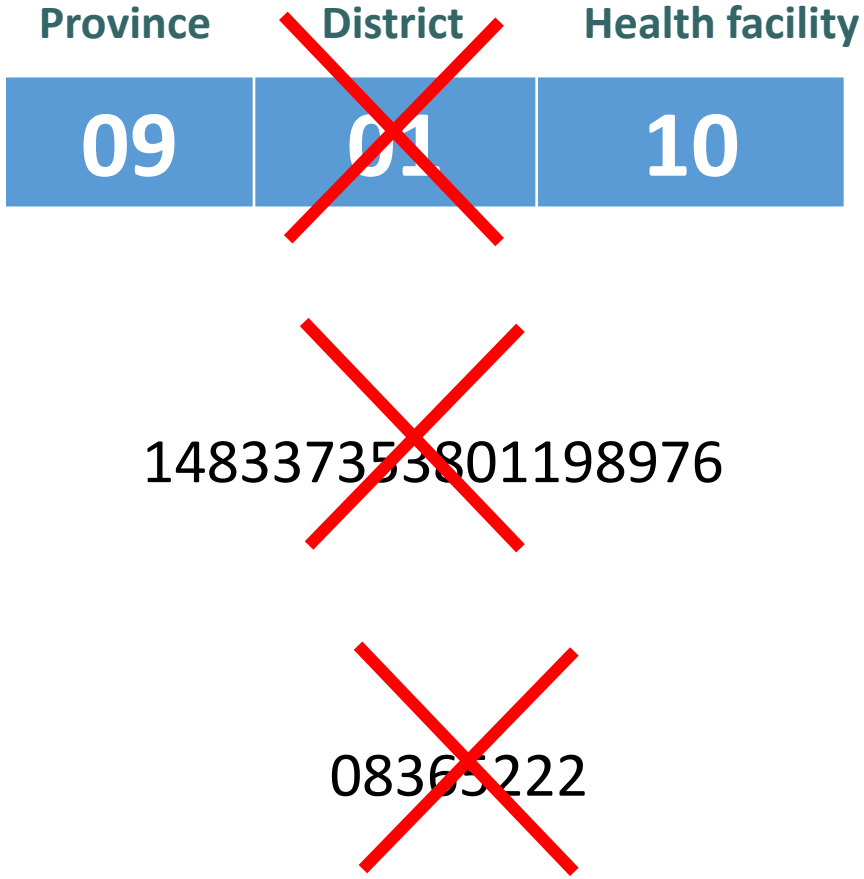
Georeferenced master list – Coding scheme

Core rules to choose an appropriate coding scheme to uniquely identify each record in the master list:

1. Ensure for the code schemes to be **meaningless** when applied to an infrastructure => never include information that could change through time in an identifier (e.g., admin division code, health facility type, etc)

2. Use a sequence as short as possible but considering the number of changes that could take place over the coming decades (function of the current number of health facilities observed in the country for example) to avoid having to modify the structure of the coding scheme at some point

3. Use a specific set of characters at the beginning of the sequence (e.g. “HF”) to avoid the problems linked to having a “0” in front of it (disappears or format not considered in the same way (e.g., text vs integer) in some application/software



Georeferenced master list – Coding scheme

Ensure for the unique identifier to remains the same from the start until the end of existence of each geographic feature (example: from the opening until the closing of a health facility and this even if its name, type (upgrade) or location changes (because of a flood for example))

UNIQUE_ID	HF_NAME_EN	HF_TYPE	PRO_CODE	PRO_NAME_EN	DIS_CODE	DIS_NAME_EN	COM_CODE	COM_NAME_EN	VIL_CODE
HF000001	Banteay Neang	HC	01	Banteay Meanchey	0102	Mongkol Borei	010201	Banteay Neang	01020103
HF000002	Chamnaom	HC	01	Banteay Meanchey	0102	Mongkol Borei	010203	Chamnaom	01020305
HF000003	Kouk Ballangk	HC	01	Banteay Meanchey	0102	Mongkol Borei	010204	Kouk Ballangk	01020401
HF000004	Koy Maeng	HC	01	Banteay Meanchey	0102	Mongkol Borei	010205	Koy Maeng	01020503
HF000005	Ou Prasat	HC	01	Banteay Meanchey	0102	Mongkol Borei	010201	Banteay Neang	01020113
HF000006	Phnum Touch	HC	01	Banteay Meanchey	0102	Mongkol Borei	010207	Phnum Touch	01020701
HF000007	Rohat tuek	HC	01	Banteay Meanchey	0102	Mongkol Borei	010208	Rohat Tuek	01020801

2017

2010 HF000005 Ou Prasat

2005 HF000005 Ou Prasat

2010 10204 Kouk Ballangk

2005 10205 Koy Maeng

Example for Cambodia

All the other information are stored in the master list and past values are kept in the IT solution serving as health facility registry

➡ Allows following the changes occurring through time for any data element of a given record

Georeferenced master list – Coding scheme

“Geo-tag”

If users want to be able to quickly identify which administrative unit a health facility is located in, they can always use what we call a “geo-tag”.

➡ Concatenate the unique identifier of the administrative unit with that of the health facility on the fly based on the contents of the health facility master list

UNIQUE_ID	HF_NAME_EN	HF_TYPE	PRO_CODE	PRO_NAME_EN	DIS_CODE	DIS_NAME_EN	COM_CODE	COM_NAME_EN	VIL_CODE
HF000001	Banteay Neang	HC	1	Banteay Meanchey	102	Mongkol Borei	10201	Banteay Neang	1020103
HF000002	Chamnaom	HC	1	Banteay Meanchey	102	Mongkol Borei	10203	Chamnaom	1020305
HF000003	Kouk Ballangk	HC	1	Banteay Meanchey	102	Mongkol Borei	10204	Kouk Ballangk	1020401
HF000004	Koy Maeng	HC	1	Banteay Meanchey	102	Mongkol Borei	10205	Koy Maeng	1020503
HF000005	[illegible]	HC	1	Banteay Meanchey	102	Mongkol Borei	10201	Banteay Neang	1020113

1020113 + HF000005

➡ Geo-tag: 1020113HF000005

➡ Not to be used as a unique identifier as this “tag” has meaning and it will change over time!

Georeferenced master list – Naming convention

- Good practice:
  - In local language and English
  - Include the name in full (avoid acronyms)
  - When applicable, avoid name duplicates by including:
    - the infrastructure type
    - The name of the administrative unit in which the feature is located

name_vn
Bß╕ çnh viß╕ çn –Éa khoa Khu vß╕ ▒c T ├óy Bß ‖ »c Tß╕ ënh Nghß╕ ç An
Bß╕ çnh viß╕ çn –Éa khoa Huyß╕ çn Quß╕ | Hß╕ úp
Bß╕ çnh viß╕ çn –Éa khoa Huyß╕ çn Kß╕ | S ├ín

A unique name for each feature

HF009530	Tha Pyay Pin SRHC (Ingapu in Leik Paung Swea RHC)
HF009538	Tha Pyay Pin SRHC (Ingapu in Zay Di Khon RHC)

HF008370	Zay Kone SRHC (Kalaw)
HF003467	Zay Kone SRHC (Thegon)

Examples for Myanmar

Georeferenced master list – Location

Address

- Street number and name (if applicable)
- Possibly the postal code if useful

Location in the administrative structure

Official code and name of the administrative unit of 1st, 2nd, 3rd,… level in which the geographic feature is located

➡ The administrative structure of the country also evolves over time

➡ Requires a regularly updated master list of administrative units and settlements

➡ Such master list should come from the governmental entity having the curation mandate over this information

Georeferenced master list – Location

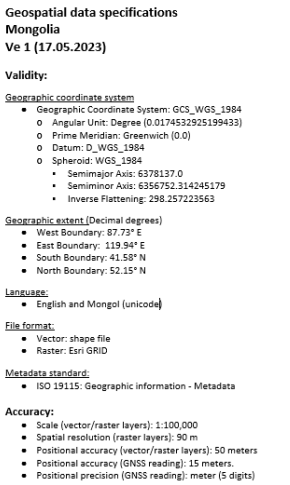
Geographic coordinates

Latitude and longitude of each geographic feature expressed in decimal degrees as well as the indication of the source, method and level of accuracy attached to the coordinates (a separate data element for each information)

Example:

LAT	LONG	COOR_SO	COOR_ME	COOR_AC
47.36039427	102.4659928	MOH	GNSS	HIGH
47.19471094	102.8254849	CPRA	GOOGLE MAP	MODERATE
47.4487403	103.1499235	CPRA	GOOGLE MAP	MODERATE
46.89904492	102.7681493	WHO	GNSS	HIGH
46.81401853	102.2465582	CPRA	GOOGLE MAP	MODERATE
47.194902	102.826179	MNUMS	UNKNOW	LOW
46.759302	103.516541	CPRA	GOOGLE MAP	MODERATE

COOR_SO: Source
COOR_MU: Method used
COOR_AC: Qualitative assessment of accuracy



Geospatial data specifications
Mongolia
Ve 1 (17.05.2023)

Validity:

Geographic coordinate system
- Geographic Coordinate System: GCS_WGS_1984
  - Angular Unit: Degree (0.0174532925199433)
  - Prime Meridian: Greenwich (0.0)
  - Datum: D_WGS_1984
  - Spheroid: WGS_1984
    - Semimajor Axis: 6378137.0
    - Semiminor Axis: 6356752.314245179
    - Inverse Flattening: 298.257223563

Geographic extent (Decimal degrees)
- West Boundary: 87.73° E
- East Boundary: 119.94° E
- South Boundary: 41.58° N
- North Boundary: 52.15° N

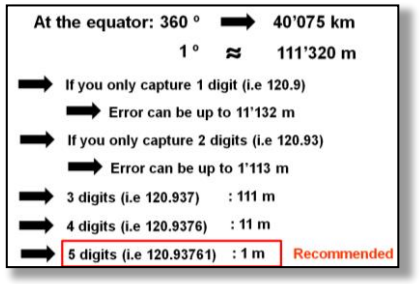
Language:
- English and Mongol (unicode)

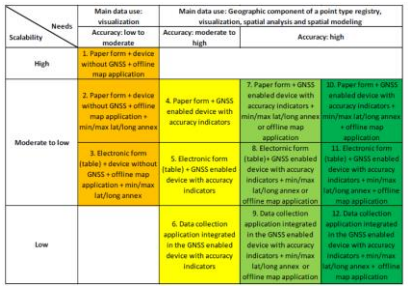
File format:
- Vector: shape file
- Raster: Esri GRID

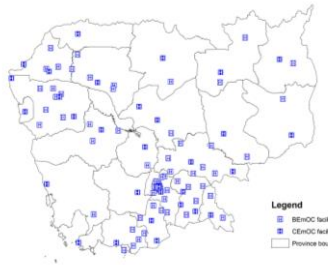
Metadata standard:
- ISO 19115: Geographic information - Metadata

Accuracy:
- Scale (vector/raster layers): 1:100,000
- Spatial resolution (raster layers): 90 m
- Positional accuracy (vector/raster layers): 50 meters
- Positional accuracy (GNSS reading): 15 meters.
- Positional precision (GNSS reading): meter (5 digits)

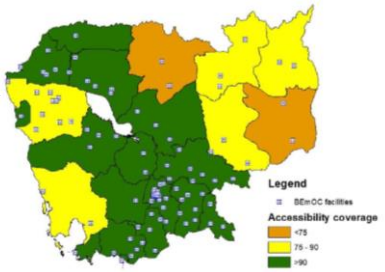
Timeliness:
- The most recent available data should be used.
- Data older than 5 years should be avoided.



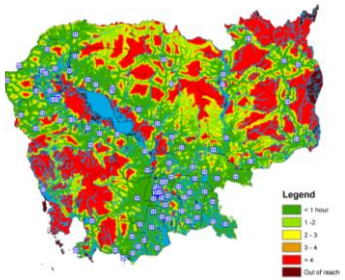




Thematic mapping



Spatial analysis



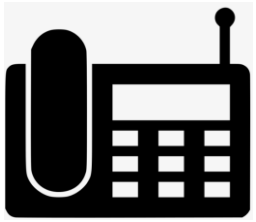
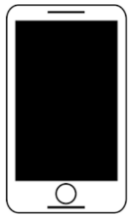
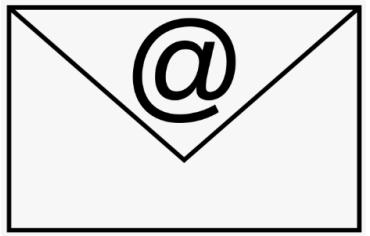
Spatial modeling

Examples for Cambodia

Georeferenced master list – Contact information

Information needed to contact the healthcare facility through various media, including:

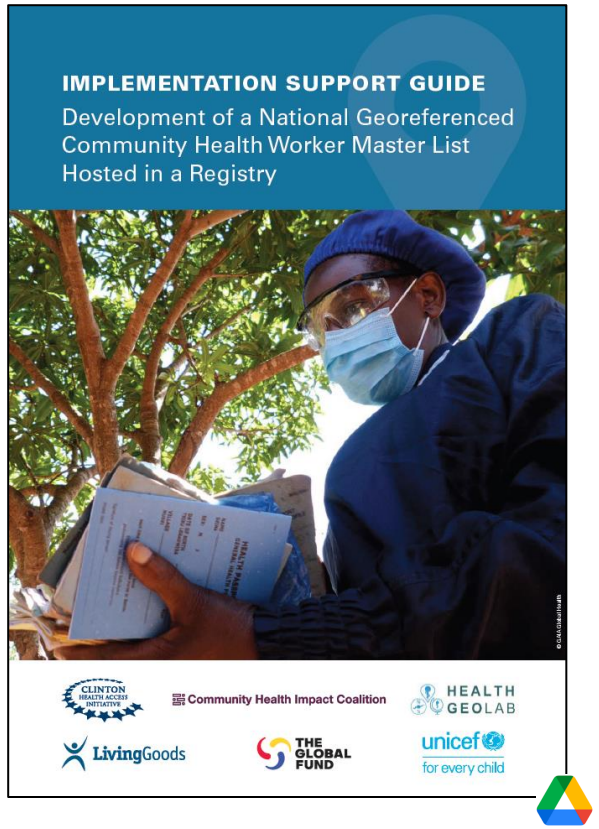
- The full name and position of the head of the infrastructure
- Telephone numbers (mobile, landline)
- E-mail address

➡ Not applicable to all the geographic features

Georeferenced master list – Process

1. Assess the current state
2. Establish the governance structure
3. Define the target state
4. Generate the first version of master list
5. Establish the registry (or common geo-registry) to host, manage and regularly update the master list
6. Share and use the master list
7. Maintain the master list and registry



Short break

10 min

Geo-enabling the Health Information System, programs or interventions training workshop for Asia Pacific

...and beyond

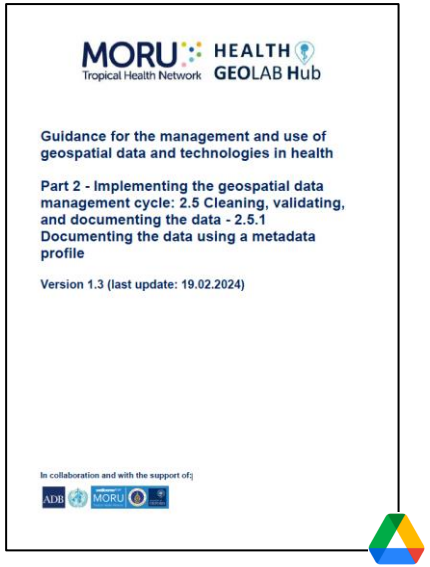
Session 14: Implement the geospatial data management cycle – Document the data

Document the data - Metadata

Documentation is a key activity in the geospatial data management cycle. It applies not only to processes (Step 2.1), but also to individual data that has been compiled and/or created.

- Georeferenced master lists, geospatial and statistical data must be properly documented using what we call metadata.

- Topics of one of the HGL's guidance document (2.5.1)


MORU Tropical Health Network HEALTH GEOLAB Hub

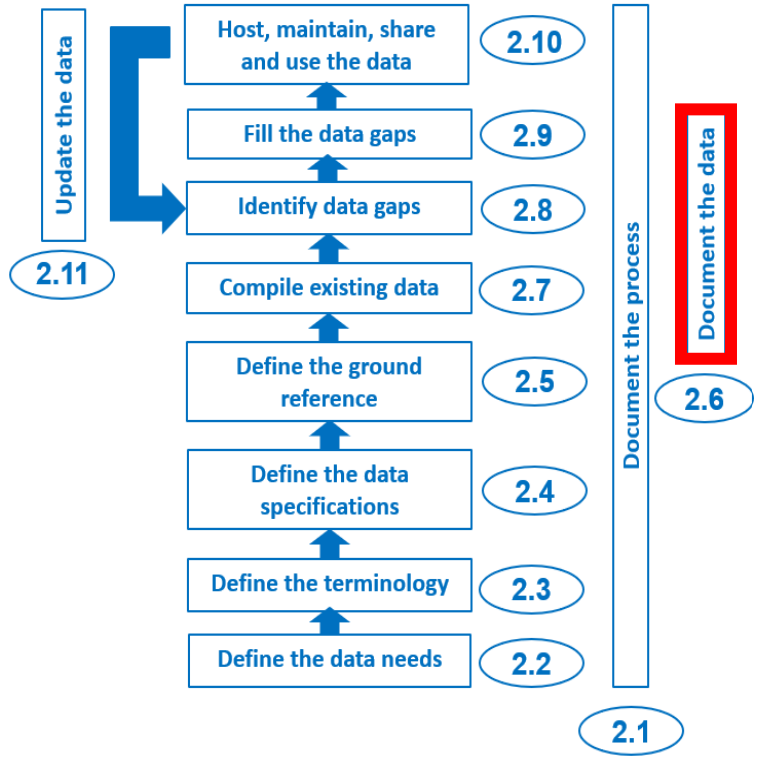
Guidance for the management and use of geospatial data and technologies in health

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)

In collaboration and with the support of:

https://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_5_1.pdf



Document the data - Metadata

Information that describes the content, quality, condition, origin, and other characteristics of data or other pieces of information.

- Data about the data
- Allows user to find out if the data is appropriate for its intended purpose
- Applies to georeferenced master lists, geospatial and statistical data

Nutrition Facts

Serving Size 172 g

Amount Per Serving	
Calories 200	Calories from Fat 8
	% Daily Value*
Total Fat 1g	1%
Saturated Fat 0g	1%
Trans Fat	
Cholesterol 0mg	0%
Sodium 7mg	0%
Total Carbohydrate 36g	12%
Dietary Fiber 11g	45%
Sugars 6g	
Protein 13g	

Vitamin A	1%	• Vitamin C	1%
Calcium	4%	• Iron	24%

*Percent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs.

NutritionData.com

The information captured in the metadata profile should be captured as much as possible during data collection and completed before data dissemination

- Begin as soon as the data specifications are defined.

Document the data - Metadata

Good metadata must at least make it possible to answer the following questions:

- What is the data about?
- Who created the data?
- When was the data created, collected, last updated or released?
- How was the data created?
- What are the data specifications associated with the data (geographic coordinate system/projection system, extent, scale, accuracy, precision, language, etc.)
- Is the data subject to any use or redistribution restrictions?
- Who can I contact if I have questions about the data?

Document the data - Metadata

Information is to be captured into a metadata record based on a **metadata profile** (contextualization of an existing metadata standard).

A **metadata standard** is a requirement to establish a common understanding of the meaning or semantics of data, and to ensure correct and adequate use and interpretation of data by its owners and users.[1]

The most widely used metadata standards for geospatial data are those produced by:

- The International Organization for Standardization (ISO)
- The Federal Geographic Data Committee (FGDC)

FGDC
FEDERAL GEOGRAPHIC DATA COMMITTEE

➡ The ISO standard is obtained by international consensus which promotes a wider use

➡ The FGDC is even encouraging federal agencies to transition to the ISO metadata standard [2]

[1] https://en.wikipedia.org/wiki/Metadata_standard

[2] https://www.fgdc.gov/metadata/selecting-a-geospatial-metadata-standard

Document the data – Metadata for geospatial data

The minimum metadata profile for geospatial data recommended by the HGL is based on the ISO 19115 metadata standard.

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 ha to serve as the offici governmental health facilitie				
	Generic keyword associated to the geospatial data layer	Health				
	Keywords describing the area covered by the geospatial data layer	Region VIII, Philippines				
Who created the data?	Complete name of the Institution who generated the geospatial data layer	Department of Health of the				
When was the data created/collected/last updated?	Date at which the geospatial data layer has been created/collected	01-05-15				
	Development status of the geospatial data layer	Completed				
	Frequency at which the geospatial data layer is being updated	Continual				
How was the data created?	General description of the steps that have been followed to create the geospatial data layer	This geospatial data layer h by compiling different sou coordinates from within anc These coordinates have th before linking them to the in the National health Facilit				

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				
	Denominator of the scale at which the geospatial data layer has been generated	100,000 (meaning 1:100,000 scale)				
	Format of the geospatial data layer	Vector				
	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)				
	Scope that will allow for the system to define the quality information that should be captured in the metadata	Dataset				
	Type of quality report that is being provided about the geospatial data layer	Absolute External Positional Accuracy				
	Identifies the axe (horizontal or vertical) to which the spatial quality information applies	horizontal				
	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				
	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				
	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				

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

30 fields to fill in to cover the key questions mentioned earlier

Document the data – Metadata for georeferenced master list

There are no specific metadata standards for georeferenced master lists.

A simple and effective way is to add a “ data catalogue” and a "metadata“ worksheets into the Excel file containing the georeferenced master list when sharing it.

Field	Description of the content
REG_C	Official Region code
REG_N_E	Official Region name (English)
PRO_C	Official Province code
PRO_N_E	Official Province name (English)
MUN_C	Official Municipality code
MUN_N_E	Official Municipality name (English)
VIL_C	Official Village code
VIL_N_E	Official Village name (English)

Data catalogue

Metadata

Title:	Administrative units master list for Tolkien Province
Originator:	National Mapping Agency of Middle Earth (NMAME)
Publication date:	January 2025
Abstract:	This master list has been created to be used during the Introduction to geospatial data management and technologies for Malaria Programs training workshop
Process:	The master list has been obtained from PSA
Progress:	Ongoing (updated regularly)
Access constraints:	The access to this data is limited to the participants attending the above mentioned training workshop
Use constraints:	The use of this data is limited to the participants attending the above mentioned training workshop
Acknowledgment:	National Mapping Agency of Middle Earth (NMAME)
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 NMAME be liable for damages arising from its use.
Primary Contact	
Contact Name	Elrond Luna
Organization	NMAME
Contact Telephone number:	899-0000
Contact Email Address:	e.luna@nmame.gov

Document the data – Metadata for statistical data

There are different metadata standards for statistical data, including the EDSM (Exchange of Statistical Data and Metadata) initiative. However, these standards tend to be difficult for non-IT professionals to understand and focus more on how data and metadata are exchanged.

- The same approach as the one used for georeferenced master list can also be applied for files containing statistical data

- “Data catalogue” and "metadata“ worksheets included into the Excel file containing the statistical data .

Field	Description
INF_ID	Official unique identifier of the infrastructure
INF_N_E	Official or business name of the infrastructure (English)
NBR_IDP	Number of internally displaced people residing currently in the evacuation center

Data catalogue

Title:	Evacuation center-level number of internally displaced people
Originator:	National Disaster Agency of the Middle Earth (NDAME)
Publication date:	20/03/2025
Temporal vailidity:	20/03/2025
Abstract:	This data set contains the evacuation center level number of internally displaced people due to the January 1, 2025 earthquake
Process:	This data has been collected in the field by the emergency response team deployed in the Tolkien province since January 1st, 2025
Progress:	Ongoing (updated daily)
Access constraints:	Access to this data is limited to the authorized personnel
Use constraints:	Use of this data is limited to the authorized personnel
Acknowledgment:	National Disaster Agency of the Middle Earth (NDAME) Response team
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 NDAME be liable for damages arising from its use.
Primary Contact	
Full name	Miguel Da Silva
Organization	NDAME
Phone number:	877-0001
Email address:	m.dasilva@ndame.gov

Metadata

Document the data – Metadata for georeferenced master list and statistical data

Microsoft Excel is the recommended file format for exchanging georeferenced master lists and statistical data because:

- It is widely used in countries
- It allows the storage of data and information in multiple worksheets within the same file, which is not the case with other file formats such as .csv files.

➡ This allows the georeferenced master list or statistical dataset as well as the data catalogue and metadata to be stored in a single file, ensuring that none of this information is lost when shared.

Geo-enabling the Health Information System, programs or interventions training workshop for Asia Pacific

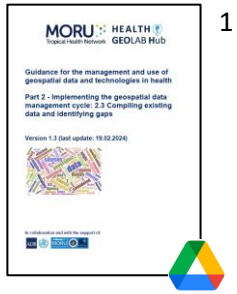
...and beyond

Session 15: Implement the geospatial data management cycle – Compile existing data, identify and fill data gaps

Compile existing data, identify and fill data gaps

The next steps consist in compiling existing data and identifying potential gaps against the needs that have been defined at the beginning of the process

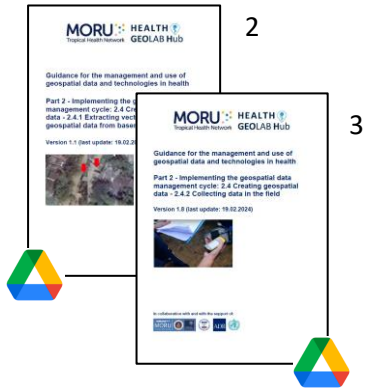
➡ Topics of one of the HGL's guidance document (2.3)



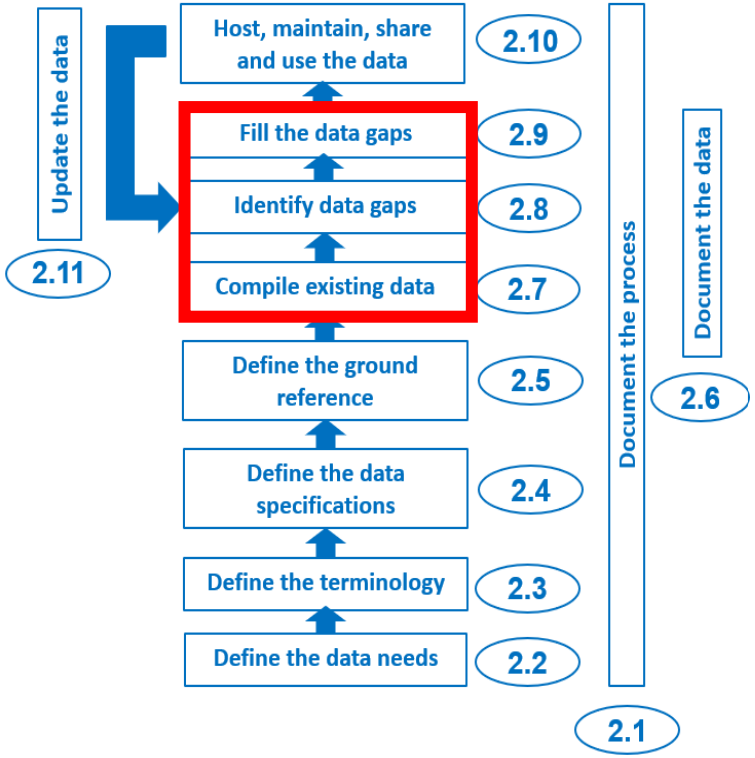
1

The identified gaps is then to be filled as much as possible using different methods depending on the type of data and the resources at this disposal

➡ Topics of two other HGL guidance document (2.4.1 and 2.4.2)



2 3



1 https://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_3.pdf

2 https://healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_4_1.pdf

3 https://healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_4_2.pdf

Compile existing data

What to compile?

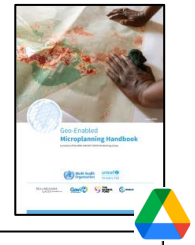
All data identified during the step which defined the data needs

➡ Depending on the applications of geospatial data and technologies considered, and therefore the GIS-products to be generate, this may include:

- All data and information that could be used to establish one or more master lists
- If they already exist, the master list for all the geographic features included in the data model that has been documented
- The geospatial data containing the location (point) or extension (polygons, lines) of these same geographic entities
- Statistical data that will either be represented on thematic maps or used during geospatial analyses or modeling
- The parameters that will be used during certain analyses (example: travel scenario)

It is critical to obtain, and to keep, the metadata for this data and information

Compile existing data



Source of the data to be compiled – Master lists and geospatial data

Master lists should only come from the government entities having the official curation mandate over each considered geographic feature.

This table provides the list of the governmental entities generally in charge of the master list and associated geospatial data for the geographic features core to public health.

Geographic feature	Master list	Geospatial data	Governmental entity
Health facilities	✓	✓	Ministry of Health
Health districts or other reporting divisions	✓	✓	Ministry of Health
Administrative divisions and villages	✓	✓	Ministry of Interior, National Statistical Agency, National Mapping Agency
Transportation network	Not necessary	✓	Ministry of Public Works, Ministry of Transportation
Hydrographic network	Not necessary	✓	Ministry of Environment/Agriculture
Climate data (temperature, precipitation, etc.)	Not applicable	✓	Ministry of Meteorology, Meteorological agency
Digital Elevation Model (DEM)	Not applicable	✓	National Mapping Agency
Land cover	Not applicable	✓	National Mapping Agency, Ministry of Environment/Agriculture

➡ A master list is not required for certain geographic features

Compile existing data

Other potential sources of data to compile

Non -governmental sources of geospatial data can also be considered according to the availability, accessibility and quality of the governmental data

For example:

- **NGOs** (UN, etc.) **and voluntary groups** (e.g. OSM): administrative boundaries, road network, hydrographic network, populated places, etc.
- **Research groups/universities**: Population distribution grids (e.g. WorldPop), land cover, DEM

➡ The quality and usability (e.g. use and/or redistribution rights) of this data must first be identified

Compile existing data

Other potential sources of data to compile – Example of global data

Geographic feature	Database name	Temporal validity/ last update	Format	Website	Access rights	Use restrictions	Redistribution restrictions
Digital elevation model	SRTM90	2000	Raster	https://cmr.earthdata.nasa.gov/search/concepts/C1214622194-SCIOPS	Open access	Commercial use	Commercial use
Land cover	Copernicus Global Land Cover Layers	2015 - 2019	Raster	https://land.copernicus.eu/global/products/lc	Open access	None	None
Distribution of the population	WorldPop	2020	Raster	https://hub.worldpop.org/geodata/summary?id=50236	Open access	None	None
Hydrographic network	OpenStreetMap (OSM)	2024	Vector	http://download.geofabrik.de/	Open access	None	None
Transport network							
Building footprint	Google Open building	2021	Vector	https://sites.research.google/open-buildings/	Open access	None	None

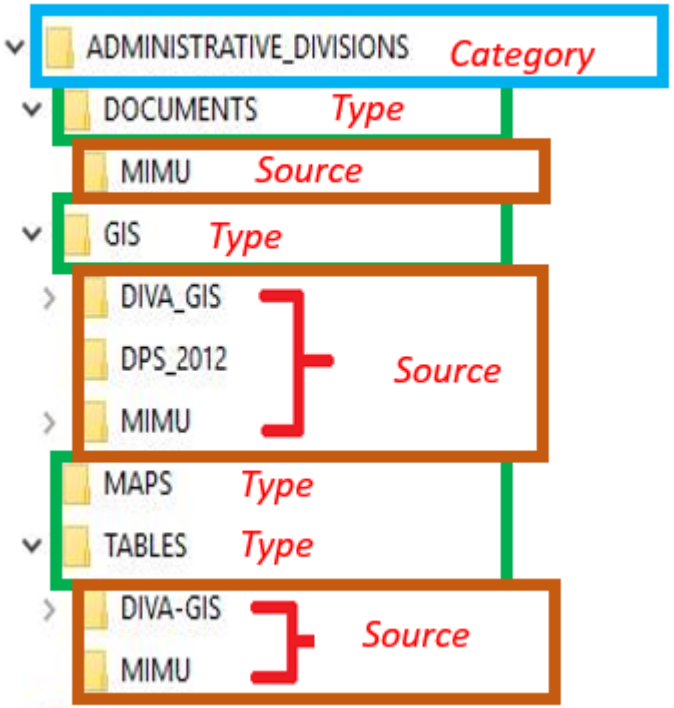
Compile existing data

Organization of the compiled data

It is important to organize the data compiled so that it can be easily found, including by other users in the event of sharing it within a group.

Example using 3 levels of folders:

1. Data category corresponding to geographic characteristics (for example administrative divisions)
2. Data type
   Four types to consider:
   - *Documents: reports, publication and other narrative documents*
   - *GIS: For geospatial data*
   - *Maps: for maps saved in PDF, MS Word, or other formats*
   - *Tables: data recorded in the form of a table (Excel, CSV, DBF, etc.)*
3. Data source with a file by source, including metadata



➡ The compiled data file is then placed in the corresponding subfolder

Clean and homogenize compiled data

Lists and geospatial data

Once organized the data will have to be cleaned and homogenized in terms of validity (format, structure) and uniqueness (duplicate)

Lists

The cleaning phase, including data homogenization, can then be implemented as follows on each of the lists identified as containing data information.

sufficient quality:

- Delete data elements not part of the defined minimum set;
- Remove values identified as being of poor quality (e.g. incomplete telephone numbers or imprecise geographic coordinates);
- If necessary, deconcatenate information that should be stored in different fields (eg names of human settlements and administrative units);
- If necessary, adjust the values of different data elements to the format and standards agreed in Step 3 (e.g. telephone number format, geographic coordinates in decimal degrees, etc.);
- Adjust the wording of each data element to facilitate merging and combining lists while keeping track of the source of the information as well as its temporal validity (e.g. NOM_ASC_DSIS2_2017; NOM_ASC_MS_2020);
- If the list does not include a UI for each ASC, add one temporarily to make it easier to identify each record once the merge and combination has been completed.

Once the different lists have been cleaned and homogenized, identify potential duplicates in each list, decide which record should be kept and whether certain data elements should be transferred from the duplicates to the main record (e.g.: some of the data elements were only entered in one of the duplicates or the value is more up-to-date in a duplicate than the record that will be retained). Then remove duplicates.

The lists (if more than one) can then be merged (rows) and/or combined (columns)

➡ First draft of master list for which gaps can be identified

Geospatial data

Data specifications

Validity:

Geographic coordinate system

- Geographic Coordinate System: GCS_WGS_1984
  - Angular Unit: Degree (0.0174532925199433)
  - Prime Meridian: Greenwich (0.0)
  - Datum: D_WGS_1984
  - Spheroid: WGS_1984
    - Semimajor Axis: 6378137.0
    - Semiminor Axis: 6356752.314245179
    - Inverse Flattening: 298.257223563

Geographic extent (Decimal degrees)

- West Boundary: 92.1° E
- East Boundary: 101.2° E
- South Boundary: 9.6° N
- North Boundary: 28.6° N

\+ When applicable, align the geospatial data attribute table with the contents of the corresponding master list

➡ Geospatial data for which gaps can be identified

Clean and homogenize compiled data

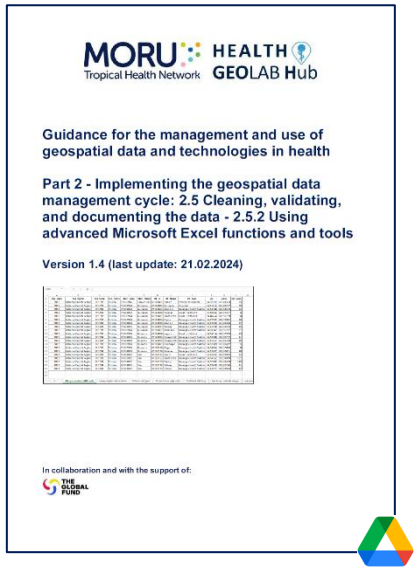
Using MS Excel

An important part of the cleaning and standardizing work for the compiled lists will happen in MS Excel before the first version of the master list is managed in a registry or common geo-registry

- It is important to know how to use a set of functions and tools available in MS Excel
- Subject of one of the guides prepared by the Health GeoLab (2.5.2)

https://healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_5_2.pdf

Step-by-step

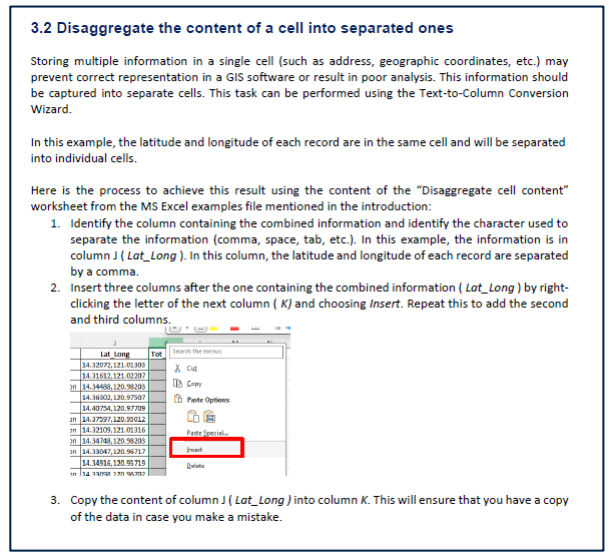

MORU Tropical Health Network HEALTH GEOLAB Hub

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 and tools

Version 1.4 (last update: 21.02.2024)

In collaboration and with the support of:


3.2 Disaggregate the content of a cell into separated ones

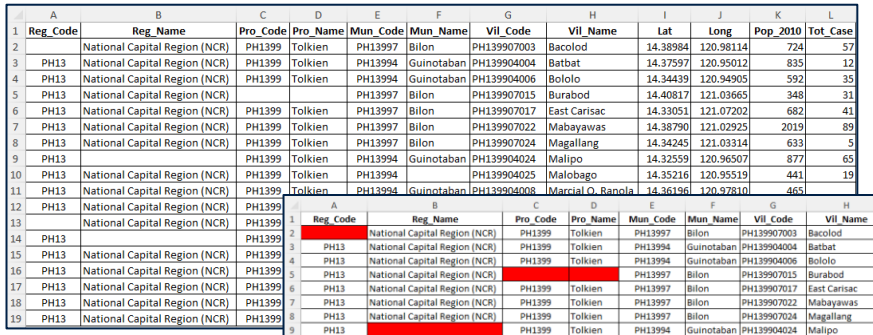
Storing multiple information in a single cell (such as address, geographic coordinates, etc.) may prevent correct representation in a GIS software or result in poor analysis. This information should be captured into separate cells. This task can be performed using the Text-to-Column Conversion Wizard.

In this example, the latitude and longitude of each record are in the same cell and will be separated into individual cells.

Here is the process to achieve this result using the content of the "Disaggregate cell content" worksheet from the MS Excel examples file mentioned in the introduction:

1. Identify the column containing the combined information and identify the character used to separate the information (comma, space, tab, etc.). In this example, the information is in column J (*Lat_Long*). In this column, the latitude and longitude of each record are separated by a comma.
2. Insert three columns after the one containing the combined information (*Lat_Long*) by right-clicking the letter of the next column (*K)* and choosing *Insert*. Repeat this to add the second and third columns.

3. Copy the content of column J (*Lat_Long)* into column *K*. This will ensure that you have a copy of the data in case you make a mistake.



Reg_Code	Reg_Name	Pro_Code	Pro_Name	Mun_Code	Mun_Name	Vil_Code	Vil_Name	Lat	Long	Pop_2010	Tot_Case
	National Capital Region (NCR)	PH1399	Tolkien	PH13997	Bilon	PH139907003	Bacolod	14.38984	120.98114	724	57
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13994	Guinotaban	PH139904004	Batbat	14.37597	120.95012	835	12
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13994	Guinotaban	PH139904006	Bololo	14.34439	120.94905	592	35
PH13	National Capital Region (NCR)			PH13997	Bilon	PH139907015	Burabod	14.40817	121.03665	348	31
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13997	Bilon	PH139907017	East Carisac	14.33051	121.07202	682	41
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13997	Bilon	PH139907022	Mabayawas	14.38790	121.02925	2019	89
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13997	Bilon	PH139907024	Magallang	14.34245	121.03314	633	5
PH13		PH1399	Tolkien	PH13994	Guinotaban	PH139904024	Malipo	14.32559	120.96507	877	65
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13994		PH139904025	Malobago	14.35216	120.95519	441	19
PH13	National Capital Region (NCR)	PH1399	Tolkien	PH13994	Guinotaban	PH139904008	Marcial O. Ranola	14.36196	120.97810	465	
PH13	National Capital Region (NCR)	PH1399									
	National Capital Region (NCR)	PH1399									
PH13		PH1399									
PH13	National Capital Region (NCR)	PH1399									
PH13	National Capital Region (NCR)	PH1399									
PH13	National Capital Region (NCR)	PH1399									
PH13	National Capital Region (NCR)	PH1399									

Exercise file

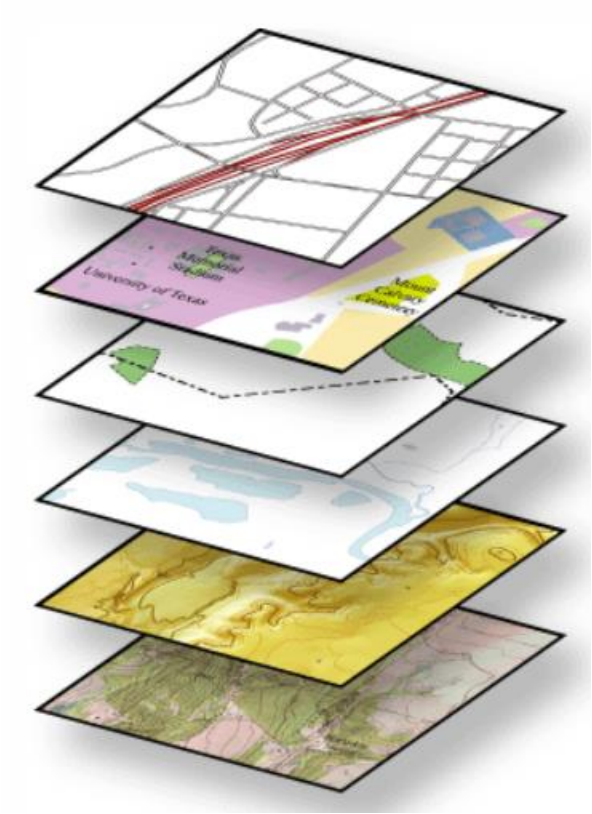
Result file

Clean and homogenize compiled data

Identify temporal disparities

Data collected at, or representative of, different points in time

- The GIS software will not report this problem
- This should be identified as part of the cleaning and homogenization process
- Importance of metadata



Roads (2023)

Land cover (2000)

Administrative boundaries (2015)

Hydrographic network (2020)

Population (2009)

Satellite imagery (2022)

Clean and homogenize compiled data

Identify data gaps

Data gaps can be of two types:

1. No source is available, accessible and/or usable (e.g. restriction of use)

2. There are gaps across some or all the six dimensions of data quality. For example:

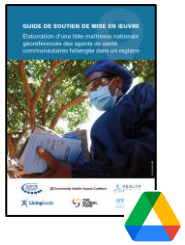
<u>1st draft of the master list</u>

- Lack of value/information for certain data elements
- Uncertainty about values/information entered (e.g. accuracy of geographic coordinates)
- Outdated information

<u>Geospatial data</u>

- Geographic features not captured or missing
- Tracing of geographic features not aligned with satellite imagery (example: roads)
- Inconsistencies between databases (administrative boundaries not aligned with the hydrographic network)

Assess the quality of compiled data



Lists and Master lists

Questions to assess the quality of the master list, or lists to be combined, across the 6 dimensions of data quality

Several cases may arise after answering these questions:

1. None of the lists contains information considered to be of sufficient quality (too old, too much gaps, many duplicates, etc.).
   ➡ Create the master list from scratch
2. Only one of the lists contains information considered to be of sufficient quality.
   ➡ Process based solely on this list (no fusion is necessary).
3. Several lists contain information considered to be sufficient quality.
   ➡ Merge the different lists together

Dimension quality	Questions to answer	Method for respond to the question	Information result / measurement
Timeliness	When was the last update?	Access to metadata and/or interview with the source of the data	Date (YYYY-MM-DD) of the last update of the list
	Have all data elements been updated or just some of them?		List of data elements updated during the last update
Completeness	Does the list cover all data elements included in the minimum set?	Visual list analysis	List of minimum data elements covered by the list
	The value of each data element is it available for all records in the list?	Manual identification or pseudo-automatic empty records	% missing values for each data element
Uniqueness	Does the list contain duplicates?	Manual identification or pseudo-automatic duplicates	% of duplicates identified in the list
Accuracy	Does the information entered in the list correspond to reality ?	Access to SOPs used for data collection, random verification, comparison between sources, visual verification of geographic coordinates	% of records whose values do not match reality
Validity	The values of each data element are they seized according to the format and to agreed standards?	Manual verification of format and standards against which the values are entered	List of elements data for which the values do not match the format and/or Standards
Consistency	Are there significant inconsistencies in the how certain data elements are entered between the lists?	Visual comparison between lists	List of elements data for which significant differences were observed on the different lists

Assess the quality of compiled data

Geospatial data

Questions to assess the quality of compiled geospatial data

Quality dimension	Applicability		Questions to be answered	Method to answer the question	Resulting information/ measurement
	Vector format	Raster format			
Timeliness	X	X	What is the temporal representativity of the dataset?	Access to metadata and/or interview data source	Date or period of validity matching or not the data specifications
Completeness	X		With master list: Does the geospatial data contain all the geographic objects contained in the master list?	Compare the content of the geospatial data with the content of the master list	% of geographic objects from the master list missing in the geospatial data
			Without master list: Does the geospatial data contain all the features observed on the satellite images used as ground reference?	Visually assess the level of completeness using satellite imagery as ground reference	Estimated % of missing geographic objects
Uniqueness	X		With master list: Does the geospatial data contain duplicates based on the master list?	Compare the content of the geospatial data with the content of the master list	% of identified duplicates
			Without master list: Does the geospatial data contain duplicates that can be identified based on the content of the attribute table and/or geographic location or extent?	Visually check content of attribute table as well as location or geographic extent	% of identified duplicates
Accuracy	X		Is the scale at which the geospatial data has been created matching the one defined in the data specifications?	Access to metadata and/or interview data source, SOP used for data creation	Difference in scale been the geospatial data and the data specifications
			Are the geographic objects in the geospatial data located with the expected positional accuracy defined in the data specifications?	Visually assess the level of accuracy using satellite imagery as ground reference; access to SOP used for data collection, random check, comparison between sources,	Estimated % of geographic objects that are not located with the expected horizontal accuracy
		X	Is the resolution of the geospatial data matching the one defined in the data specifications?	Check the properties of the geospatial data	Difference in resolution between the geospatial data and the data specifications
Validity	X	X	Is the metadata for the geospatial data available ?	Access to metadata and/or interview data source	Availability of metadata
			Are the geographic coordinate system and map projection known?	Access to metadata and/or interview data source	Availability of projection information
			Is the geospatial data available in a format that is compatible with the ones defined in the data specification or can it be converted accordingly	Check the data format and/or interview data source if unknow	Compatibility of format
			Does the geospatial data cover the study area as defined in the data specifications?	Visually assess the coverage of the geospatial data using the satellite images as ground reference	% coverage of the study area

Consistency is itself reached once the reference criteria set for the other dimensions are reached

Assess the quality of compiled data

Statistical data

When it comes to statistic data, the questions to answer are mainly the following from the point of view of their use in an application of geospatial data and technologies:

- Timeliness:
  - What is the temporal validity of statistical data?
  - Does it correspond to the temporal validity defined in the data specifications?
- Completeness:
  - Is a value available for each of the records in the corresponding master list?
- Uniqueness:
  - Does the data set contain duplicates?
- Validity:
  - Is the data set accompanied by a data catalogue and metadata containing all the information necessary to use it correctly?
  - Is the data set available in a format that allows its use or could be converted to such a format?

Assess the quality of compiled data

Identified data gaps

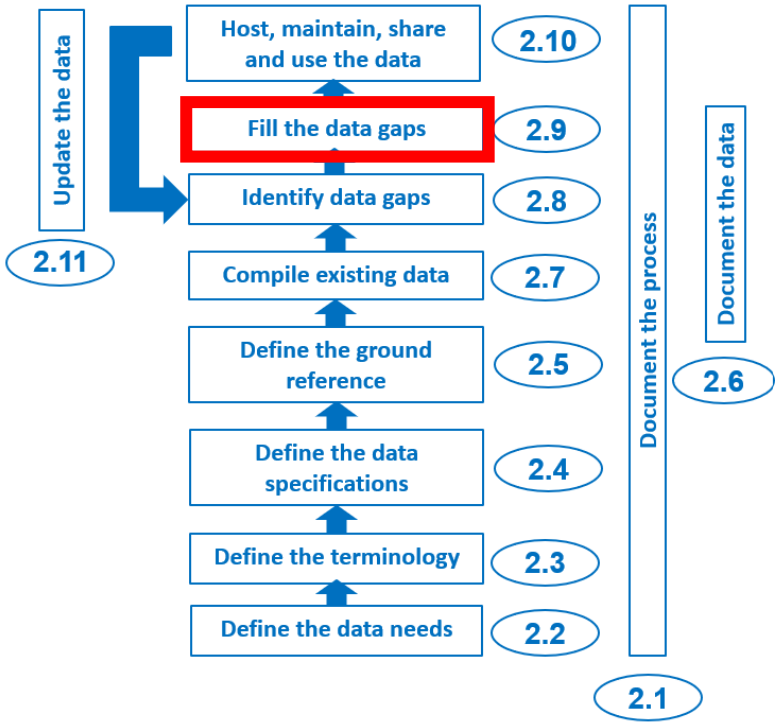
- All the gaps should be documented and ideally filled as part of the action plan implementation
- Before that, it is always worth to first make sure you have not missed any data sources during the data compilation exercise
- If this is not the case, the gaps need to be filled

Fill data gaps

Data gaps don't need to be filled at this stage, but it is important to estimate if and how this can be done in preparation for the development of the action plan

Filling data gaps (georeferenced master lists, geospatial and statistical data) can be done in the following main ways:

- By improving existing data (geospatial data)
- By collecting or extracting additional data (lists, geospatial data, statistical data)
- Modelling or projecting values (statistical data)



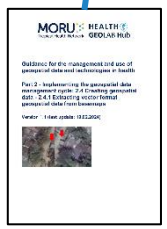
It is important to start by filling the gaps in the master lists because they serve as the ground reference for geospatial and statistical data

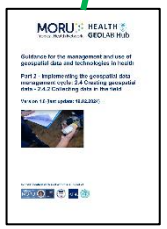
Independently from the method, the data in question should be validated to ensure the highest possible quality

Fill data gaps

Possible approaches to fill data or information gaps

	Manual extraction/correction			Interactive or automatic extraction	Field data collection	Expert consultation
	Participatory mapping	Crowd mapping	Single operator			
Geospatial data (vector)						
Points	✓	✓	✓	✓	✓	
Lines	✓	✓	✓	✓		
Polygons	✓	✓	✓	✓		
Other						
Data elements (list)					✓	✓
Statistical data					✓	✓
Input parameters (e.g. travel scenario)						✓





➡ HGL guides covering part of these methods (2.4.1 and 2.4.2)

Fill data gaps – Master list

	Manual extraction/correction			Interactive or automatic extraction	Field data collection	Expert consultation
	Participatory mapping	Crowd mapping	Single operator			
Geospatial data (vector)						
Points	✓	✓	✓	✓	✓	
Lines	✓	✓	✓	✓		
Polygons	✓	✓	✓	✓		
Other						
Data elements (list)					✓	✓
Statistical data					✓	✓
Input parameters (e.g. travel scenario)						✓

Before collecting data to fill the gaps identified in the first draft of a master list, it is important to validate the list in question with the government entity in charge of curating this list to ensure that the list:

1. Contains all currently active geographic features – Record-level completeness
2. Contains the indication of the administrative unit (from the highest to the lowest level in the hierarchy) in which each geographic feature is currently located
3. Does not contain duplicates that might have been missed during the cleaning and homogenization step

- It is important to identify at what level of disaggregation the validation should take place to involve people with situational awareness on the ground
- Depending on the situation, and the type of geographical feature, validation can constitute the first step when collecting data in the field

Fill data gaps – Master list

	Manual extraction/correction			Interactive or automatic extraction	Field data collection	Expert consultation
	Participatory mapping	Crowd mapping	Single operator			
Geospatial data (vector)						
Points	✓	✓	✓	✓	✓	
Lines	✓	✓	✓	✓		
Polygons	✓	✓	✓	✓		
Other						
Data elements (list)					✓	✓
Statistical data					✓	✓
Input parameters (e.g. travel scenario)						✓

Once the completeness of the records of the 1st draft of the master list validated by the curating government authority, and this applies to any geographic feature (infrastructure, health personnel, etc.), the collection of missing, uncertain or outdated data can take place

Two types of data can be distinguished at this level because they may require a different collection method:

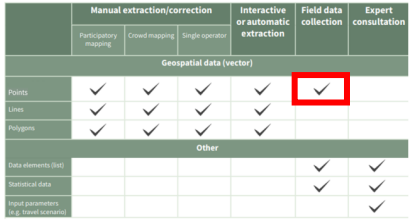
- Geographic coordinates
- Other data elements contained in the list

➡ In both cases, a Standard Operating Procedure (SOP), including the validation step, should be defined and documented and the data collectors trained accordingly to ensure the effectiveness of the exercise and the quality of the data collected

➡ Important note: the official unique identifier of each geographic feature as captured in the master list must be included in the data collection instrument being used to make the link with the master list

Fill data gaps – Collecting geographic coordinates

Twelve (12) methods for collecting geographic coordinates have been identified and documented by the Health GeoLab (guidance 2.4.2):



This table can be used to determine the most appropriate method based on the need for scalability and main data use

Scalability \ Needs	Main data use: visualization	Main data use: Geographic component of a point type registry, visualization, spatial analysis and spatial modeling	
	Accuracy: low to moderate	Accuracy: moderate to high	Accuracy: high
High	1. Paper form + device without GNSS + offline map application		
Moderate to low	2. Paper form + device without GNSS + offline map application + min/max lat/long annex	4. Paper form + GNSS enabled device with accuracy indicators	7. Paper form + GNSS enabled device with accuracy indicators + min/max lat/long annex or offline map application
	3. Electronic form (table) + device without GNSS + offline map application + min/max lat/long annex	5. Electronic form (table) + GNSS enabled device with accuracy indicators	8. Electornic form (table)+ GNSS enabled device with accuracy indicators + min/max lat/long annex or offline map application
Low		6. Data collection application integrated in the GNSS enabled device with accuracy indicators	9. Data collection application integrated in the GNSS enabled device with accuracy indicators + min/max lat/long annex or offline map application

Fill data gaps – Collecting geographic coordinates

When selecting a method from this table, it is important to remember that:

1. These methods consider that there is no pre-existing electronic field data collection system in place in the country. If a system is already in place, and it provides the necessary functionalities to collect quality data then using it should be considered
2. The use of an electronic form, with predefined content when applicable, can considerably reduce data collection time as well as data entry errors. The use of such form should therefore be preferred over paper ones
3. The higher the positional accuracy and precision of the collected coordinates, the greater the number of use cases. It is therefore recommended to always aim to achieve the highest possible level of positional accuracy and precision
4. Collecting geographic coordinates through the sole use of an offline map application (methods 1-3) requires that data collectors know how to navigate on a map or imagery and to recognize the terrain from an aerial view

Fill data gaps – Collecting geographic coordinates

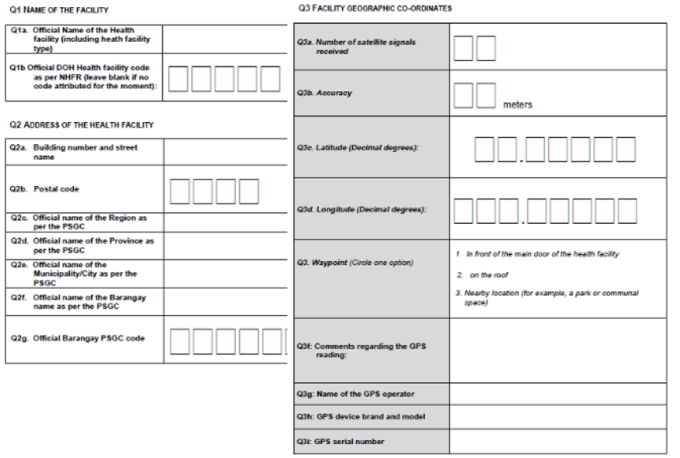
The following items are required for each of these methods:

- Data collection form
- A standard operating procedure (SOP)
- The master list of the geographic features for which geographic coordinates are to be collected

Data collection form

A proper collection form should include fields allowing to capture:

- The official name and unique identifier of the geographic feature taken from the master list
- The address and location in the administrative structure (official name and unique identifier of the administrative unit in which the geographic feature is located) of the feature
- The geographic coordinates of the geographic feature as well as the fields used to evaluate the quality of these coordinates (number of signals, instrumental accuracy).



Q1 NAME OF THE FACILITY

Q1a. Official Name of the Health facility (including heath facility type)	
Q1b Official DOH Health facility code as per NHFR (leave blank if no code attributed for the moment):	☐☐☐☐☐

Q2 ADDRESS OF THE HEALTH FACILITY

Q2a. Building number and street name	
Q2b. Postal code	☐☐☐☐
Q2c. Official name of the Region as per the PSGC	
Q2d. Official name of the Province as per the PSGC	
Q2e. Official name of the Municipality/City as per the PSGC	
Q2f. Official name of the Barangay name as per the PSGC	
Q2g. Official Barangay PSGC code	☐☐☐☐☐

Q3 FACILITY GEOGRAPHIC CO-ORDINATES

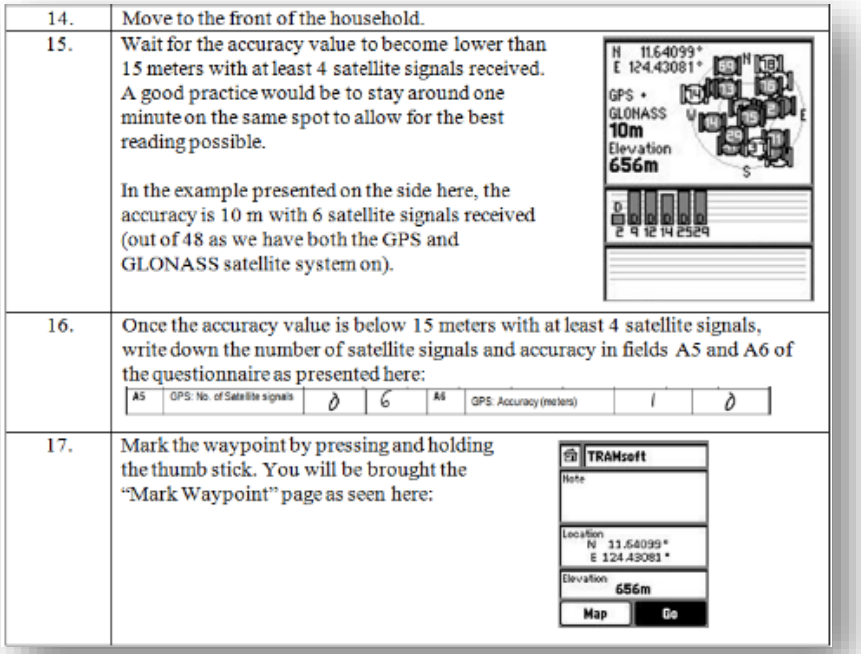
Q3a. Number of satellite signals received	☐☐
Q3b. Accuracy	☐☐ meters
Q3c. Latitude (Decimal degrees):	☐☐.☐☐☐☐☐
Q3d. Longitude (Decimal degrees):	☐☐☐.☐☐☐☐☐
Q3. Waypoint (Circle one option)	1. In front of the main door of the health facility 2. on the roof 3. Nearby location (for example, a park or communal space)
Q3f: Comments regarding the GPS reading:	
Q3g: Name of the GPS operator	
Q3h: GPS device brand and model	
Q3i: GPS serial number	

Fill data gaps – Collecting geographic coordinates

Standard Operating Procedure (SOP)

A SOP is a document that:

- Contains the details of the steps to follow to collect and verify the quality of geographic coordinates
- Should be used during the training of data collectors and supervisors as well as during the data collection exercise.
- Focuses on providing only essential information, being as clear and simple as possible, covering all the fields in the form, and including illustrations that will help the data collector complete each step



14.	Move to the front of the household.
15.	Wait for the accuracy value to become lower than 15 meters with at least 4 satellite signals received. A good practice would be to stay around one minute on the same spot to allow for the best reading possible. In the example presented on the side here, the accuracy is 10 m with 6 satellite signals received (out of 48 as we have both the GPS and GLONASS satellite system on). N 11.64099° E 124.43081° GPS • GLONASS 10m Elevation 656m
16.	Once the accuracy value is below 15 meters with at least 4 satellite signals, write down the number of satellite signals and accuracy in fields A5 and A6 of the questionnaire as presented here: A5 GPS: No. of Satellite signals 06 A6 GPS: Accuracy (meters) 10
17.	Mark the waypoint by pressing and holding the thumb stick. You will be brought the "Mark Waypoint" page as seen here: TRAMsoft Note Location N 11.64099° E 124.43081° Elevation 656m Map Go

Fill data gaps – Collecting geographic coordinates

Master Lists

The data collection exercise is based on the use of the master list of:

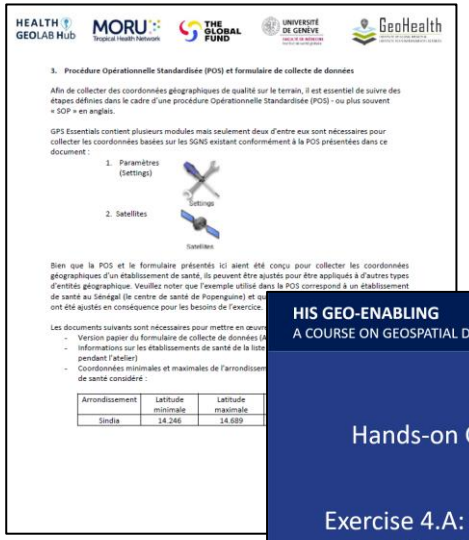
- **Geographic features that need to be located:** typically, infrastructure (e.g., health facility, school, household) or place (e.g. village)
- **Administrative units:** For the area covered by data collection exercise (total or part of the country)

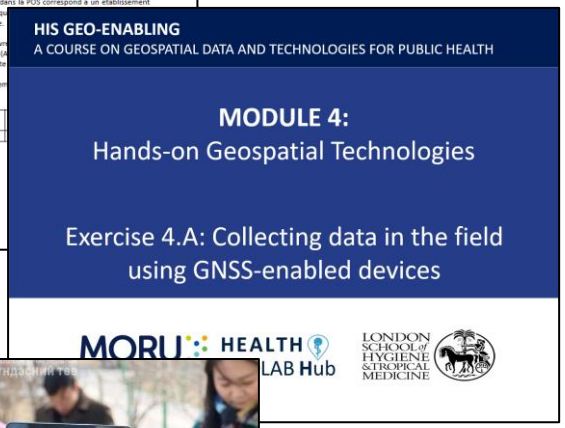
The following items are also required to implement some of the method being used:

- Geospatial data containing the boundaries of the administrative units
- A Geographic Information System (GIS) software
- GNSS-enabled devices
- An Offline map application
- Anticipated minimum/maximum values for geographic coordinates
- Purpose-designed data collection application

Fill data gaps – Collecting geographic coordinates

Several important steps must be implemented before, during and after data collection.





Before data collection:

- Prepare the materials needed to implement the selected method and to train data collectors and supervisors.
- Select the data collectors and their supervisor(s)
- Train data collectors and supervisors so that they can be as independent as possible once in the field

➡ A good training including practical exercises is key to a successful data collection exercise!

Fill data gaps – Collecting geographic coordinates

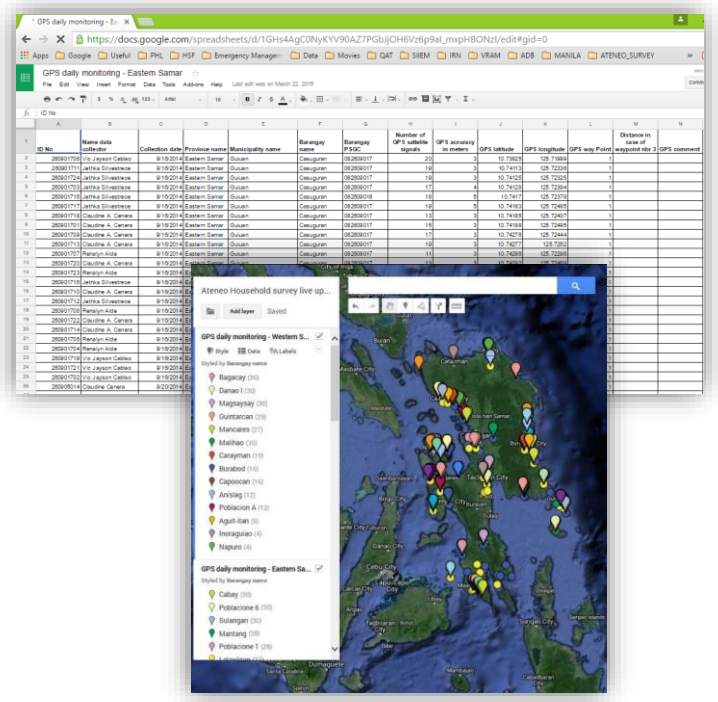
<u>During data collection:</u>

The most important is to ensure that the data collectors follow the standard operating procedure for which they have been trained.

It is also important to rapidly detect and resolve any unexpected issues while the data collectors are in the field.

The following measures can be implemented depending on available resources and the geographic extent of the area being covered:

- On-site spot checks of data accuracy and completeness conducted by the data collection supervisor
- Remote verification of the collected data submitted using a Google spreadsheet for example



Fill data gaps – Collecting geographic coordinates

After data collection:

If this has not already been done as part of the data verification process implemented during the data collection, it is important to ensure that the data collected is organized into a structured table that can be saved or exported under the form of a spreadsheet (for example, in Microsoft Excel).

If the data is collected in electronic form, it must be possible to export it in a similar structure. This dataset must contain all fields included in the data collection form.

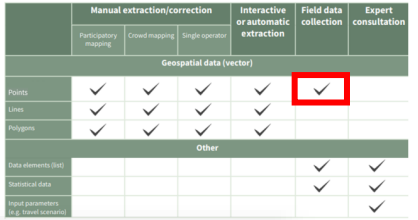
It is important that the worksheet containing the data is accompanied by the following two additional worksheets:

- The data catalogue
- The metadata

Field	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 DOH 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

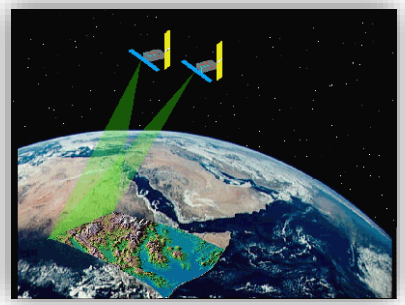
Title:	Administrative units master list for Tolkien Province
Originator:	National Mapping Agency of Middle Earth (NMAME)
Publication date:	January 2025
Abstract:	This master list has been created to be used during the Introduction to geospatial data management and technologies for Malaria Programs training workshop
Process:	The master list has been obtained from PSA
Progress:	Ongoing (updated regularly)
Access constraints:	The access to this data is limited to the participants attending the above mentioned training workshop
Use constraints:	The use of this data is limited to the participants attending the above mentioned training workshop
Acknowledgment:	National Mapping Agency of Middle Earth (NMAME)
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 NMAME be liable for damages arising from its use.
Primary Contact	
Contact Name	Elrond Luna
Organization	NMAME
Contact Telephone number:	899-0000
Contact Email Address:	e.luna@nmame.gov

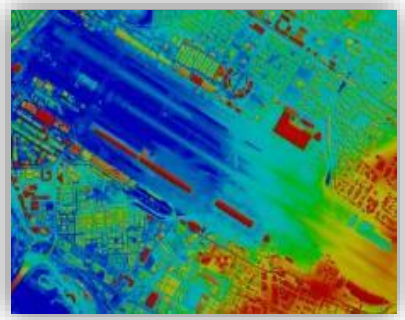
Fill data gaps – Geospatial data



Filling gaps in geospatial data can be done in different ways depending on the expected format:

1. Raster format:
   - Use of remote sensors placed on a satellite, airplane or drone (remote sensing)
2. Raster or vector format:
   - Processing images from remote sensing (image processing)
3. Vector format:
   - Adjust, create or extract point, line, or polygon type geographic features from a basemap





➡ The method you are most likely going to use

Fill data gaps - Adjust, create or extract point, line, or polygon type geographic features from a basemap

Adjust GIS data in vector format (editing)

Data specifications


Annex 1 - Data specifications for the MOH, Cambodia

Validity:

Geographic coordinate system
- Geographic Coordinate System: GCS_WGS_1984
  - Angular Unit: Degree (0.0174532925199433)
  - Prime Meridian: Greenwich (0.0)
  - Datum: D_WGS_1984
  - Spheroid: WGS_1984
    - Semimajor Axis: 6378137.0
    - Semiminor Axis: 6356752.314245179
    - Inverse Flattening: 298.257223563

Geographic extent (Decimal degrees)
- West Boundary: 102.3° E
- East Boundary: 107.7° E
- South Boundary: 10.3° N
- North Boundary: 14.7° N

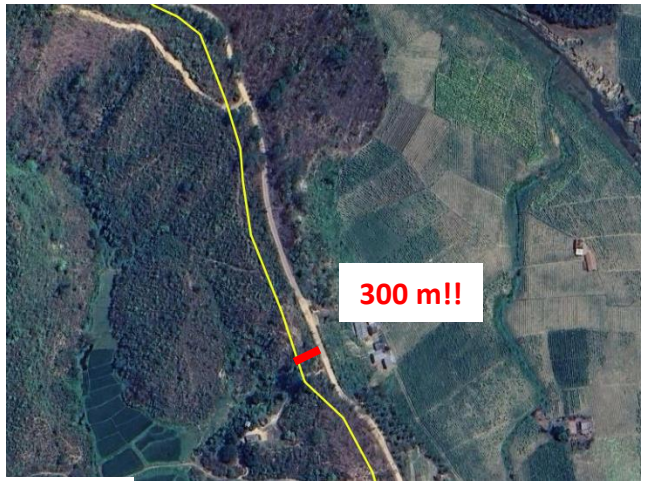
Language:
- Khmer and English (unicode)

File format:
- Vector: shape file
- Raster: Esri GRID
- Tabular: Mdb/xsd

Metadata standard:
- ISO 19115: Geographic information - Metadata
- Metadata profile: the one recommended by the HGLC

Accuracy/Precision:
- Scale (vector/raster layers): 1:100,000 - 1:250,000
- Spatial resolution (raster layers): 90 m
- Positional accuracy (vector/raster layers): 50 meters
- Positional accuracy (GNSS reading): 15 meters
- Precision: meter (5 digits)

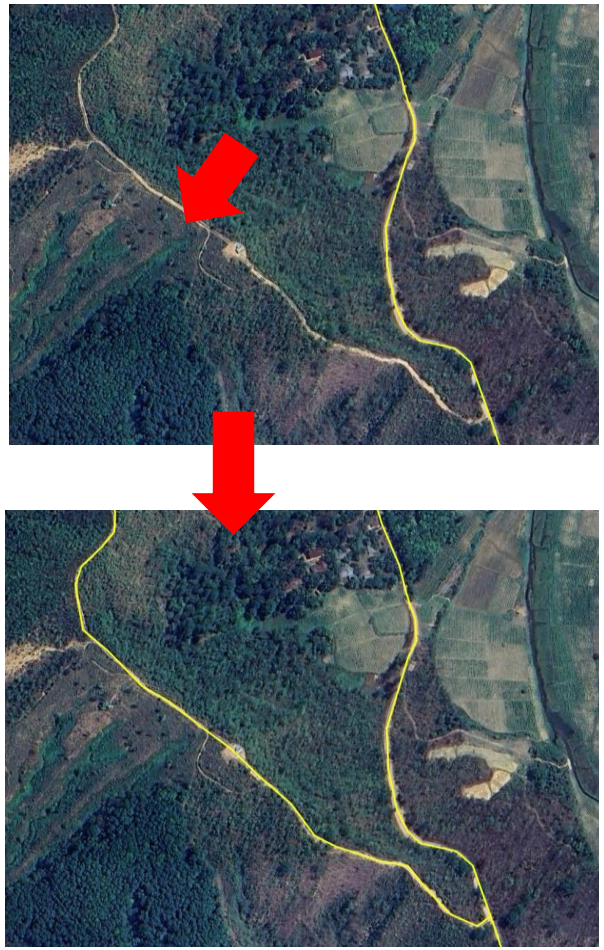
Timeliness:
- The most recent available data should be used
- Data older than 5 years should be avoided

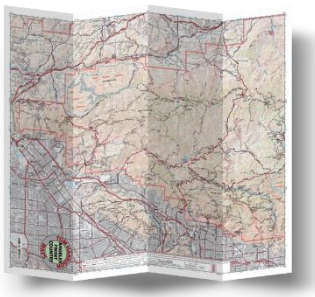


Accuracy/Precision:

- Scale (vector/raster layers): 1:100,000 - 1:250,000
- Spatial resolution (raster layers): 90 m
- Positional accuracy (vector/raster layers): 50 meters
- Positional accuracy (GNSS reading): 15 meters

Create/extract data in vector format (digitizing)





Fill data gaps - Adjust, create or extract point, line, or polygon type geographic features from a basemap

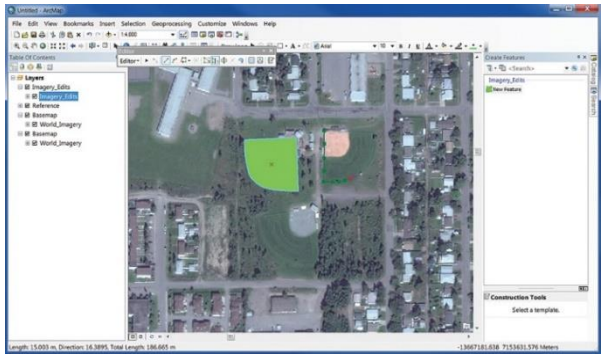
When editing and/or extracting vector data from a basemap, it is important to choose

1. The appropriate basemap
   - Respect the data specifications that have been defined:
     - Imagery: temporal validity, resolution, positional accuracy
     - Scanned map: temporal validity, scale, resolution
   - Images should be free of clouds wherever possible to ensure completeness.
   - The images must be well orthorectified to avoid "duplicates" at the border between two scenes.

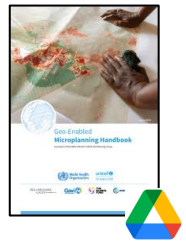
Fill data gaps - Adjust, create or extract point, line, or polygon type geographic features from a basemap

When editing and/or extracting vector data from a basemap, it is important to choose

2. The most appropriate method:
   - Manual Extraction: human-guided digitizing of geographic features from a basemap (usually done on-screen these days)
   - Interactive extraction: manual extraction of features assisted by the automatic snapping to raster cells functionalities provided by the GIS software or application being used
   - Automatic Extraction: An automatic, computer-guided method that converts raster data into vector data using image processing software that uses pattern recognition technology

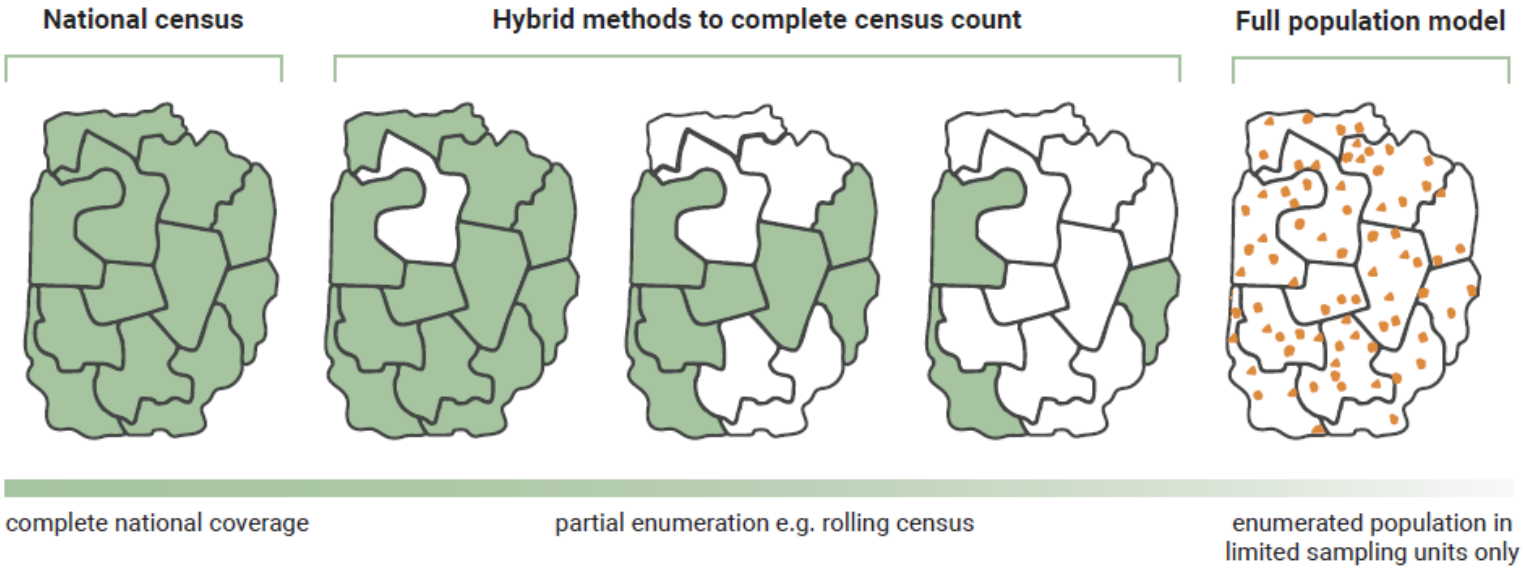


Fill data gaps – Population estimates and spatial distribution



Different types of gaps might have to be filled:

1. Need to expand the geographic extent of the available population estimates to obtain full coverage. The method being used will depend on the available data:



→ Bottom-up estimation to fill gaps

2. Need to increase the resolution of the population estimate (disaggregate) → Top-down estimation to fill gaps
3. Project the population estimates to be representative of the point in time for which it is needed

2015 → 2020 → 2024

Fill data gaps - …and in the end…

It may well be that some of the gaps could not be filled due to lack of time and/or resources

If this is the case, a decision will have to be taken regarding the creation of the GIS-based products:

- Option 1: The quality of the data remains sufficient for generating good quality GIS-based products (! Be careful of the impact this could have on decision-making!). In this case, it is necessary to precisely document the limitations that this implies, including for the GIS-based products that will be created.
- Option 2: The quality of the data is insufficient, and this is a major limitation. In this case, the creation of the GIS-based products may need to be postponed until the necessary time and/or resources are available.
- Option 3: The quality is insufficient, but the gaps could be addressed as part of the action plan implementation or the operationalization of the updating mechanism

Module 5 – Schedule and agenda

Schedule Module 5 28 August 2024 (Bangkok 12pm / Geneva 6am / Fiji 6pm)
15 min - Recap of Module 4 and agenda of Module 5
30 min – **Session 16:** Introduction to geospatial technologies
30 min - **Session 17:** Introduction to Global Navigation Satellite System (GNSS)
30 min - **Session 18:** Introduction to Geographic Information System (GIS)
30 min - **Session 19:** Introduction to the concepts of registry and Common Geo-Registry (CGR)

Geospatial technologies

Thank you for your attention and see you all again soon!