



# HIS geo-enabling: Guidance on the establishment of a common geo-registry for the simultaneous hosting, maintenance, update and sharing of master lists core to public health

Version 1.0

14 July 2017

# Table of Contents

<b>Table of Contents</b>	<b>1</b>
<b>Acknowledgements</b>	<b>3</b>
<b>Authors</b>	<b>3</b>
<b>Preface</b>	<b>4</b>
<b>Target audience</b>	<b>4</b>
<b>Background</b>	<b>4</b>
<b>Objectives of the document</b>	<b>5</b>
<b>Use Cases, Benefits, Challenges and Opportunities</b>	<b>6</b>
Use Cases	7
Benefits	9
Challenges and how they can be overcome	9
Opportunities	12
<b>Geographic objects and relationships core to public health</b>	<b>12</b>
Core geographic objects	12
Core geographic relationships	14
Unified view of core geographic objects and relationships	14
The example of Cambodia	15
<b>Recommended specifications for the master lists and associated data</b>	<b>16</b>
Specifications for master lists	17
Health facility master list	17
Administrative divisions master list	20
Reporting divisions master list	22
Specifications for the associated data	23
Geographic hierarchies	23
Historic Changes	24
Available boundary layers in GIS format	28
<b>Data flow and requirements for the Common geo-registry</b>	<b>28</b>
Data Flow	29
Requirements from the perspective of the different user roles	32
Registry Administrator	32
Registry Maintainer	32
Registry Contributor	34

API Consumer	34
Anonymous User	35
Requirements specific to each node in the data flow	36
Initial setup of the Registry	36
Upload the Master Lists and define the geographic hierarchies	36
Curate the Master Lists	36
Publish the Master Lists and associated information	39
Access the Master Lists and associated information	40
Update the master list after curation of the request	40
<b>Concluding words</b>	<b>41</b>
<b>Appendix 1 - Glossary of Terms</b>	<b>42</b>
<b>Appendix 2 - Conceptual data model for public health in Myanmar (Will be in landscape format in the final version)</b>	<b>44</b>
<b>Appendix 3 - Example data dictionary for the health facility master list (Myanmar)</b>	<b>45</b>
<b>Appendix 4 - Example data dictionary for the administrative divisions master list</b>	<b>46</b>

## Acknowledgements

First and foremost, we would like to express our gratitude to the Asian Development Bank (ADB) for the financial support provided to this project.

Our gratitude also goes to the participants of the Core Geo-registries for Health Information Exchanges Design Workshop which took place in Phnom Penh, Cambodia on June 12-13, 2017. Their active participation during this workshop provided not only the case study included in the present document but helped in designing it.

We also would like to thank Channe Suy (iLab..... for contributing their valuable contribution and comments to the document.

## Authors

Steve Ebener - AeHIN GIS Lab, Philippines

Martin Verzilli - InSTEDD, Argentina

Jonathan Payne - Open Concept Lab / Regenstrief Institute, United States of America

Scott Teesdale - InSTEDD, United States of America

# Preface

To be provided by Susann - Regional and Dr Kiry - National

## Target audience

The present guidance targets individuals or organizations who are encountering problems with the simultaneous hosting, management, update and sharing of the master lists for the geographic objects core to public health. The audience may include officials from different ministries, implementing partners or donors involved in the planning and/or implementation phase as well as those interested in understanding the process and requirements for solving these issues and making such data and information accessible and used across the whole Health Information System (HIS).

## Background

Geography and time are intrinsically linked to public health and therefore present across its three main functions<sup>1</sup>: (1) assessing and monitoring the health of communities and populations at risk; (2) assuring that all populations have access to appropriate and cost-effective care; and (3) formulating public policies designed to solve identified health problems and priorities.

From a geographic perspective this means:

- The most critical health risks are observed in areas where the hazards and vulnerability levels are at their highest and the health capacity at its lowest level.
- Physical accessibility to health services is greatly influenced by the geographic location of the health services, the spatial distribution of the population, and the environment between them.
- Maps represent a powerful media to visualize and analyze the spatial distribution of public health related issues at all levels and therefore to support effective and informed decision making and policy formulation.

The dynamic nature of the environmental and social ecosystems in which we are living, and therefore the changes this dynamic generates on our geography, requires for time to be taken into account across these three functions as well. Specifically:

- New health facilities are being built while others are being closed.

---

<sup>1</sup> <http://www.who.int/topics/en/>

- Administrative divisions are being created, split, or merged and this often has an impact on the management of facilities as well as the geographic relationships that exist between these objects in a given country.

Despite the foundational importance of the above, geography and time are generally among the most poorly captured and therefore analysed dimensions in the country's Health Information System (HIS) and this simply because capturing and regularly updating the above mentioned dynamic and relationships is difficult.

Part of this issue can be addressed through the development, maintenance, update, and use of master lists for the geographic objects that are core to public health. A master list is an authoritative, standardized, complete, up-to-date, and uniquely coded list of all active records. The geographic objects core to public health include health facilities, reporting divisions, and administrative divisions down to the village level. However, having up-to-date master lists populated with these core geographic objects is not alone sufficient. There is still a need for a technology solution -- a common geo-registry -- to simultaneously host, maintain, update and openly share these master lists and relationships, together with their associated geography stored in a GIS readable format.

Unfortunately, the master lists in question are often either lacking or, when existing, incomplete, out-of-date. In addition, the technology solutions currently available to maintain them do not allow for the simple and easy management of each of the above mentioned geographic objects.

To fill the gap, the Asian eHealth Information Network (AeHIN) GIS Lab, with the support of the the Asian Development Bank (ADB), has included the need for such master lists and a common geo-registry to be established and maintained into the framework they have developed and are now implementing across Asia and the Pacific in collaboration with partners including but not limited to WHO, UNICEF and InSTEDD for the health sector to fully benefit from the power of geography, geospatial data and technologies. This framework is referred to as the HIS Geo-Enabling Framework.

## Objectives of the document

In this context, the primary objective of the present guidance is to provide the data flow and requirements that any platform being used as a common geo-registry should comply with in order to effectively and simultaneously host, maintain, update and openly share the master lists, geographic hierarchies and associated data/information for the geographic objects core to public health. It is therefore primarily meant to be used as an assessment and planning tool to:

- Help evaluate if existing platforms can already fulfill these requirements and support implementers in the process of filling any gaps if they exist.
- Supports the design of new platforms meant to serve as common geo-registry if no solution already exists.

It is also hoped that this document will contribute to a more systematic and comprehensive integration of geography, and indirectly time, in the different IT architectures or solutions being currently promoted in countries by the developmental partners when strengthening the Health Information System (HIS).

The document has been divided into the following sections in order to help achieving the above:

- Use Cases, Benefits, Challenges and Opportunities: to make the case for establishing, maintaining, updating and sharing master lists for the geographic objects core to public health through the use of a common geo-registry;
- Core Geographic Objects and Relationships: to describe the objects and relationships core to public health that should be covered by the common geo-registry;
- Specifications for the Master Lists and Associated Geographies: to provide recommendations on the format, structure and content that the master lists and other data describing their associated geographies should comply to in order to facilitate the management of the all data flow in the common geo-registry;
- Dataflow and Requirements of a Common Geo-registry: to support not only the assessment aimed at identifying if the platform currently being used, or planned to be used, as a common geo-registry supports the main requirements you would expect to see in such a platform but also its update or (re-)designing when needed;

This guidance intentionally does not provide any recommendation on how these requirements could be converted into functionalities into the platform and this for the simple reason that these could actually be implemented in different ways.

## Use Cases, Benefits, Challenges and Opportunities

Geography plays an important role in almost every health program and master lists together with their associated geographies could help these programs effectively assess, plan, implement and monitor their activities and this within classic vertical programs as well as cross-cutting and more interdisciplinary activities.

Despite this recognized need, Ministries of Health and implementing partners have reported facing a repeating set of problems. These are caused by a lack of a common location or simple methods to both maintain and access the master lists. This results in the following set of problems, which create a recurring tax or even blocker among many different partners:

- Key stakeholders do not have important geographic data available or easily accessible, and are often unable to answer simple questions like, “How many health centers are in this district?”
- It is difficult and expensive for each system to separately manage and keep these core pieces of data current, harmonized, and complete.
- Systems remain closed-off and siloed as:

- They are not interoperable; and remain closed due to a lack of common identifiers and ability to reference core objects.
- Groups of stakeholders and the related systems that they govern are competing for funding; the result is they are largely disincentivized to share or normalize data.
- Overall, the health architecture and set of systems remain brittle, monolithic, and are difficult and expensive to evolve over time.

The following sections make the case for establishing, maintaining, updating and sharing master lists for the geographic objects core to public public through the use of a common geo-registry. We provide a set of use cases that illustrate the above needs and then describe the benefits, challenges, and current opportunities.

## Use Cases

The following three programs in particular illustrate the critical need for geography to be better managed and integrated with their respective information systems through the use of master lists and associated geographies:

**A) Communicable Disease Prevention, Control and Elimination** - These programs which oftentimes revolve around seeking to prevent and reduce morbidity and mortality of diseases such as TB, HIV or Malaria rely heavily on geography. For example:

- ***Malaria elimination activities*** - When a positive malaria case is diagnosed in a health facility; there is a need for the case investigation team to not only being able to contact and visit the facility in question but also the place of residence of the patient (village) for further investigation, intervention and monitoring activities through time. Ideally, the information needed by the team (geographic coordinates, contact information) would be captured in the Malaria Management Information System based on master lists.

**B) Emergency and Outbreak Management** - Geography provides the common operational picture across all the phases of the emergency cycle (Preparedness, response and recovery) or during an outbreak. For example:

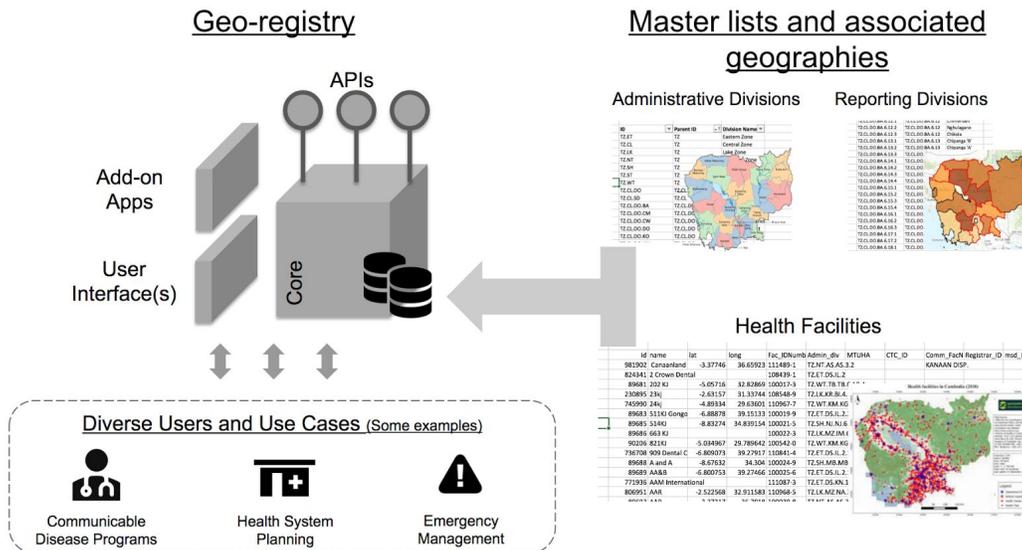
- ***Typhoon preparedness and response*** - As soon as a typhoon is confirmed to make landfall, the master lists and their associated geography allows completing the pre-disaster assessments and warning actions by: 1) evaluating the communities and key infrastructures, including health facilities most likely to be affected by being located within its path; 2) Contacting the concerned infrastructures. Right after landfall, the lists and associated geographies support the Rapid Impact Assessment (RIA) and therefore the response phases by coordinating and directing resources in the areas where they are needed the most while leveraging the infrastructures that remained operational.
- ***Disease outbreak*** - The need to map health related infrastructure and disease incidence, which has been critical ever since John Snow's famous work during the London cholera outbreak of 1854, continues to be a critical tool in combatting disease

outbreaks and Up-to-date master lists and related maps are more needed than ever to be able to contain large scale events such as the recent ebola outbreak in Western Africa.

**C) Health Planning** - Critical not only for developing health plans and programmes that answer the population's needs but also for budgeting and resource allocation in a systematic and equitable way, effective health planning requires to know where the population in need and the existing services are located in order to perform its functions. For example:

- **Health coverage assessment** - Improving health coverage throughout a country requires having a good picture of the population distribution down to the village level in order to make sure that the whole population is being covered by the catchment areas of the existing health facilities. A complete, up-to-date list of villages with the indication of their location and associated population figures is critical in this regard.
- **Health service delivery** - Once the health coverage plan has been developed, the appropriate resources (human, equipment, financial) need to be attached to each health facility in order to ensure a cost effective delivery of health services.

The above use cases directly justify the diverse needs for master lists to be established, maintained, updated and shared together with their associated geographies through a common geo-registry. As presented in Figure 1, master lists contain the actual geographic data that could be used to support the above use cases, while the common geo-registry is the software that stakeholders use to host, maintain, update and share these master lists and associated geographies



**Figure 1. Geo-registry, master lists, and users**

## Benefits

There are a number of expected benefits from implementing and using a common geo-registry to manage master lists and their associated geographies for the geographic objects core to public health. Some of these gains have been realized<sup>2</sup> among different implementations with master facility lists and facility registries, although a more comprehensive geo-registry is a novel concept.

The expected benefits of such an approach include, but are not limited to:

- ✓ Allowing more comprehensive integration of geography and time across the HIS as well as the whole data lifecycle, therefore allowing for a more systemic approach to solving public health problems.
- ✓ Supporting the implementation of programs that are heavily dependent on geography and time for their implementation and success (i.e. Malaria elimination).
- ✓ Improving the ability for governing bodies to simultaneously and more easily manage maintain, update and share master lists as a foundational element of the Health Information System (HIS).
- ✓ Facilitating the capture of the source and data collection date associated to each information stored in the master list as well as the tracking of changes being implemented and this for better accountability
- ✓ Improving data quality across the six dimensions this concept covers (completeness, validity, timeliness, uniqueness, accuracy, consistency).
- ✓ Reducing duplication of efforts and therefore cost by maintaining only one authoritative set of lists instead of several ones.
- ✓ Supporting interoperability and collaboration across partners, ability to further promote innovation and data use.
- ✓ Enhancing ability for systems to evolve and grow over time by leveraging collaborative, shared, and interoperable services around geospatial data.
- ✓ Supporting a more effective use of the visualization and analytical power offered by Geographic Information Systems (GISs).

## Challenges and how they can be overcome

Several challenges needs to be addressed in order to fully benefit from a common registry hosting the master lists for the geographic objects core to public health.

Among these challenges we can mention:

- **Governance** is one common challenge for managing the multiple master lists, especially given that the responsibility to support systems and keep data current is typically split

---

<sup>2</sup> <https://www.mendeley.com/community/facility-registry-research/documents/>

across multiple different actors. Workflows and solutions need to account for and empower collaboration between different actors.

- **Data sharing and open access** policies vary among partners. Any solution should enable administrators to set the level of data sharing that is appropriate for their environment, and, if data is not provided openly, to simplify the request for access. The combination of clear policies and suitable technologies are important tools here.
- **Sustainability** is another challenge area, in that resources need to be made available to support these foundational datasets and technologies.
- **Policy** to incentivize the use and collaboration around the master lists is important as well. Governing bodies should work with stakeholders, include NGOs, private entities, donors and other partners, to be aware of and encourage the use of the geo-registry and master list.

These challenges need to be addressed to ensure not only the quality of the content that will be shared through the common registry (master lists, hierarchies and flows) but also the long term sustainability of both the common registry and its content. The objective of this section is therefore to provide the list of the most important elements to be taken into account.

Establishing, maintaining, regularly updating and sharing master lists for the geographic objects core to public health through a common registry is a long term and time consuming activity but the benefits, not only from a data management but also a programmatic perspective, are such that the return on the original investment is worth the effort. This is especially true if some critical elements are being taken into account since the beginning of the process.

The elements in question, which are independent from the master list being considered, are being covered by:

- Establishing a proper governance structure: While such governance structure can take different form (e.g. specific division/unit within a ministry, technical working group, etc.) its main purpose is two folds:
  - Provide leadership and accountability across the development, implementation and sustainability phases
  - Ensure for the process to be as inclusive and collaborative as possible (see further down for more details)
- Clearly defining the:
  - Needs: Aiming at the highest data quality right from the beginning of the process allows for the content of the master lists to serve the largest number of needs. This approach is cost-effective as the cost for collecting high quality data is the same as for collecting poor quality ones.
  - Terminology: Defining the terminology behind the concepts and the data elements included in the different master lists is critical to ensure for all the stakeholders and the users to have the same understanding of the content of the different master lists. For example, defining the concepts of health facility,

administrative and reporting divisions is key to determine the content of each master list.

- Data specifications: These specifications are meant to capture the level of data quality expressed during the definition of the needs. As such, they should cover the 6 dimensions of data quality (completeness, validity, timeliness, uniqueness, accuracy, consistency) and this across master lists.
- Standard Operating Procedures (SOPs): Such SOPs will be needed to ensure data quality as well as the transfer of methods across institutions and individuals involved in the development, maintenance, update and sharing of the master lists through the use of the common geo-registry. As such, these SOPs, should not only cover each of the functionalities reported in the data flow (Figure...) but also all the steps linked to the collection and cleaning of the content of each master list and this respecting the data specifications mentioned here above.
- Ensuring that the necessary resources (human and financial) are available: The establishment of the master lists, their integration in a common registry, as well as the provision of technical support to entities within the concerned sector for the integration of the master lists in their respective information system and/or practices will require for at least one full time person to be dedicated to this task over the first year of implementation of the process. While the volume of work might decrease (mainly depending on the size of the country and the number of changes occurring over a year), at least one person, and ideally a replacement, would have to continue working on the maintenance and regular update of the master lists within the registry during the subsequent years. Financial resources will therefore be needed to cover the salary of the staff(s) in question on the long term. Depending on the situation, additional resources might also be necessary to perform activities in the field.
- Establishing collaboration across sectors and stakeholders: The mandate over the geographic objects core or of importance to public health is spread among different governmental entities among which we can mention the Ministry of Health (health facilities and health reporting divisions), the Ministry of Interior or the Ministry of Home Affairs (administrative divisions down to the village level) as well as Ministries such as the Ministry of Education (schools). In addition to that, costs are significantly reduced if all the stakeholders are contributing to the completion and regular update of the master lists through their respective data collection opportunities. These two points directly call for the establishment of collaborations among the sectors and stakeholders
- Releasing policies: The value of the master lists is only fully accomplished if their content is of quality, if they are accessible and if their use is enforced across the all information system. Reaching all the three requires for policies to be in place and enforced across the health system. Such policies should at least:
  - Specify the governance structure and therefore indicate which governmental entity(ies) has/have the mandate over the development, maintenance, update and sharing of the master lists.
  - Ensure the long term sustainability of the master lists by securing the necessary human and financial resources.

- Promote the open sharing and use of the master lists by all the stakeholders in the health sector.

## Opportunities

Other cross-cutting or derivative use cases of common geo-registries begin to emerge as more collaborative and integrated healthcare delivery and ecosystems of data use grow. Some of these types of use cases include the following:

**A) Open Data Initiatives** where the governing body for each of the datasets decides to make it simple and easy for all actors (government, private industry, NGOs, and the general public) to access current and trustworthy data. For example:

- **Support an ecosystem of data use and innovation** - by considering these core datasets as a public good; that many different types of actors could use; driving efficiency and innovation by making them easily accessible.
- **Decrease need for answering simple data requests** - Openly accessible data does not require for the institution in charge to answer data requests, therefore reducing the load on its staff by encouraging self-service among partners.

**B) System Integration and Interoperability**, for example:

- **M&E Indicators** - Multiple systems want to contribute indicators into a common HMIS or disease surveillance system; they require an interoperable facility ID to easily bring this data together.
- **Health Information Exchanges and System Integration** - That enable more complex workflows and integrations amongst systems. These may include things like referrals, medical record management, or combining health worker and facility data effectively. A service oriented approach also helps to add flexibility and reliability to the eHealth ecosystem by creating reusable services.

## Geographic objects and relationships core to public health

Each public health program will have its own set of geographic objects that will need to be taken into account to implement a geo-enabled registry<sup>3</sup>, <sup>4</sup>. However, there is a subset of specific geographic objects that are common to the implementation of any health program. We consider these common geographic objects as core to public health in general and describe their characteristics and use in detail in this chapter.

---

<sup>3</sup> [http://www.aehingislab.net/DOCUMENTS/Guide\\_GIS\\_Lab\\_Part2\\_1.pdf](http://www.aehingislab.net/DOCUMENTS/Guide_GIS_Lab_Part2_1.pdf)

<sup>4</sup> <https://drive.google.com/file/d/0B0iJNe5CNLsHaVFsTTYyTmlaX3M/view>

## Core geographic objects

The geographic objects core to public health that we consider here are:

- *Health facilities*, which represent points of health care services delivery;
- *Administrative divisions* down to the village level, as they represent the objects to which population data (e.g. demographic, socio-economic) are attached and the levels at which political and financial decisions are made; and
- *Reporting divisions*, which represent geographic entities used by the health system for planning, reporting and monitoring.

Table 1 further describes each of these objects.

**Table 1.** Summary of geographic objects core to public health

Geographic Object	Object Type	Examples
Health facilities	Point	<ul style="list-style-type: none"> <li>• Tertiary Hospital</li> <li>• Secondary Hospital</li> <li>• Primary Hospital</li> <li>• Health Center</li> <li>• Health Post</li> <li>• Other types of infrastructure such as offices, warehouses, labs, or pharmacies</li> </ul>
Administrative divisions <i>* Also referred to as political divisions</i>	Polygon <i>* In many countries villages may be represented by points due to the absence of boundaries in digital format.</i>	<ul style="list-style-type: none"> <li>• Country</li> <li>• Province/Region</li> <li>• District/Zone</li> <li>• Commune</li> <li>• Village</li> </ul>
Reporting divisions	Polygon	<ul style="list-style-type: none"> <li>• Health districts</li> <li>• Catchment Area</li> </ul>

The core nature of these geographic objects is confirmed by examining the three use cases presented above. For malaria elimination, it is critical to record the health facility where a case was identified and to know the place of residence of the patient (village) to deploy a case investigation team and to properly allocate resources. In the event of a typhoon or other emergencies, health facilities and administrative offices involved in mobilizing preparedness and response activities are identified by proximity and relationships to the at-risk geographic areas. Health coverage planning requires a mapping of all villages and health facilities in the country with each of their relationships to an administrative and a health division to ensure appropriate allocation.

The core nature of these objects is also confirmed by country example. The conceptual data model developed in Myanmar's HIS geo-enabling process and included in Annex 2 clearly

depicts that health facilities, administrative divisions and reporting divisions (indicated by a red box) are central to the model due to the large number of relationships (arrows) that either originate from or point to them.

### Core geographic relationships

The geographic objects above are significantly more valuable if they are logically organized by their relationships to each other. The two most important relationships are the administrative division hierarchy and the reporting division hierarchy (refer to Table 2). These hierarchies make it possible, for example, to aggregate data from the village level up to a region or to query all the catchment areas within a health district.

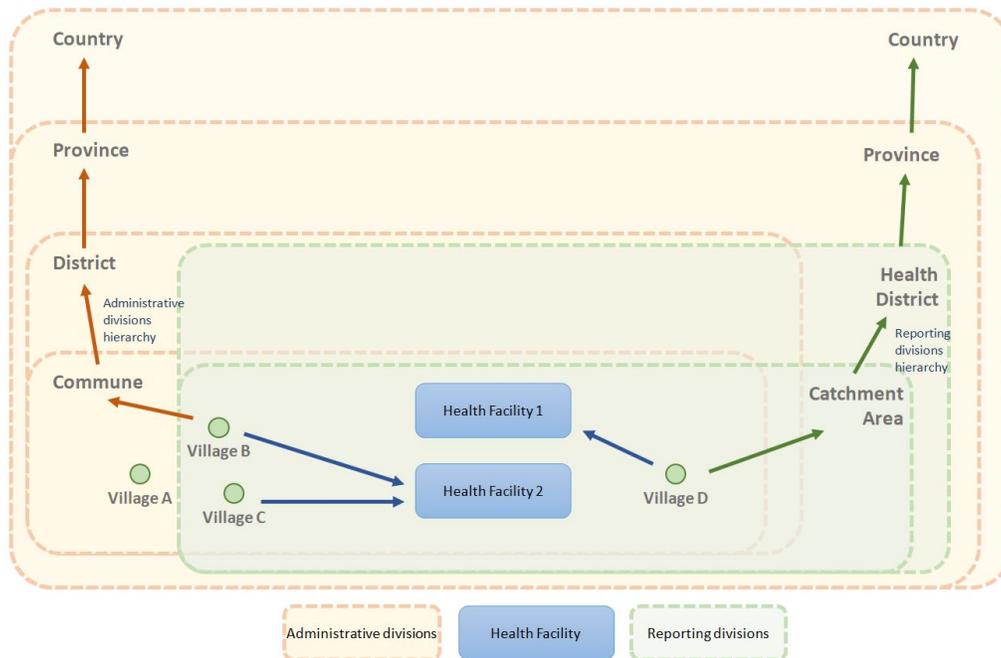
**Table 2.** Summary of geographic relationships core to public health

Relationship	Relationship Type	Examples
Administrative divisions hierarchy	Hierarchy	Village A is geographically located within Commune B
Reporting divisions hierarchy	Hierarchy	Village B is part of the catchment area attached to health facility F  Health facility G is geographically located within Health District H

Many other types of relationships between geographic objects exist and could be modeled as part of a common geo-registry. Examples include the patient referral pathway between health facilities or the reporting pathway from health facilities to administrative offices. Whether these additional relationships are included should be determined by the planned use of the registry.

### Unified view of core geographic objects and relationships

By visualizing geographic objects and relationships together, as shown in Figure 2, we can begin to see how these elements connect to each other as part of the health system. Please note that in this example it is considered that the health districts are made by the sum of a group of catchment areas.



**Figure 2.** Generic representation of the geographic objects and relationships core to public health

## The example of Cambodia

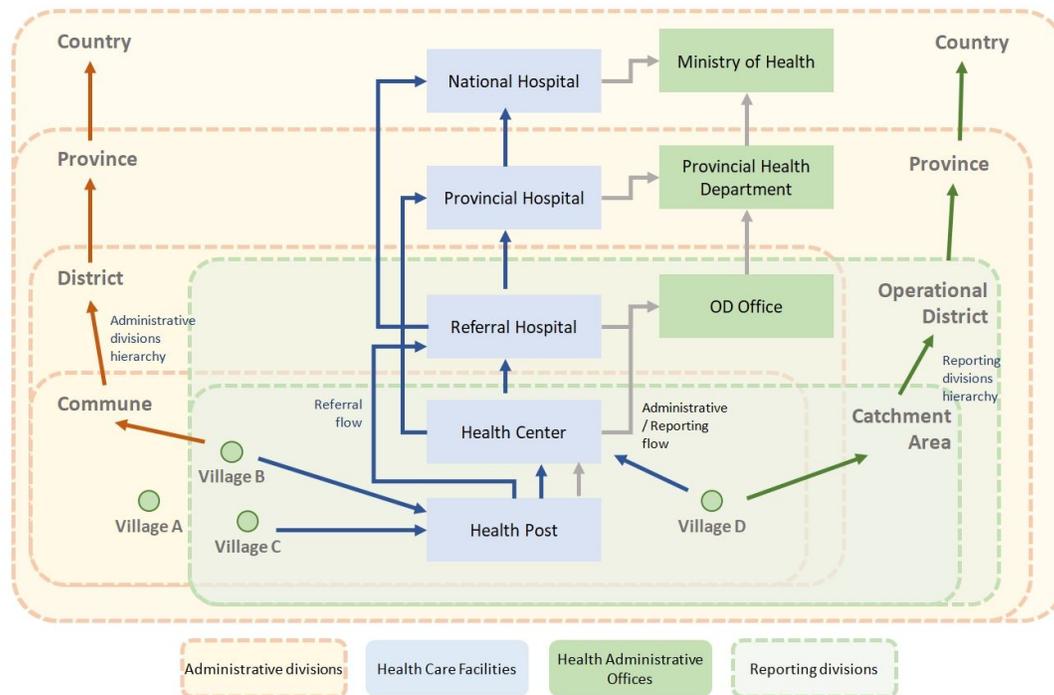
The geographic objects and relationships core to public health in Cambodia were captured during the Core Geo-registries for HIE design workshop, which took place in Phnom Penh on June 12-13, 2017<sup>5</sup>.

Workshop participants determined that in Cambodia it was important to capture certain non-core objects and relationships, in addition to the core ones (Figure 3). These included the referral pathway, the administrative or reporting pathway, and the health administrative offices. The administrative offices were included specifically to enable the accurate modeling of the reporting pathway.

While capturing these non-core objects and relationships within a common geo-registry is certainly possible, and may even be necessary for certain use cases, they are not required components and are therefore considered beyond the scope of this document. However, it is important to note that non-geographic components, like the referral and reporting pathways, could be modeled and stored alongside core geographic objects and relationships within a common geo-registry.

5

[http://www.aehingislab.net/MEETINGS/CGR\\_2017/Geo-Registriy\\_Workshop\\_Final\\_Exec\\_Sum\\_010717.pdf](http://www.aehingislab.net/MEETINGS/CGR_2017/Geo-Registriy_Workshop_Final_Exec_Sum_010717.pdf)



**Figure 3.** Geographic objects and relationships in the Cambodian context as defined during the June 2017 workshop.

## Recommended specifications for the master lists and associated data

While different countries may format their master lists and associated data slightly differently before uploading them in the common geo-registry, there are best practices for how to encode these datasets and this to not only simplify the process but also guide and facilitate the work of the data manager in charge of establishing the masterlist and indirectly the experience of the user experience once the resulting data is being downloaded from the registry. as well as reduce associated costs.

The objective of this section is to provide recommendations and specifications for:

1. The following master lists:
  - Health facilities - often referred to as master facility list (MFL)
  - Administrative divisions
  - Reporting divisions
2. Other important data associated to:
  - a. Geographic hierarchies (administrative and reporting divisions)
  - b. Historic changes (all master lists)

- c. The availability of boundaries layers in GIS format (administrative and reporting divisions)

As mentioned here above, such specifications should apply not only to the format of the data being uploaded into the common geo-registry but also to the format of the data being downloaded from it.

## Specifications for master lists

### *Health facility master list*

Several efforts by MEASURE Evaluation<sup>6,7</sup>, WHO<sup>8</sup>, OpenHIE<sup>9</sup>, and various other partners have already defined the suggested fields to consider in a health facility master list.

These fields are generally grouped into two components:

1. The signature domain: the set of fields that allow to uniquely identifying, locating and contacting each health facility in the master list.
2. The service domain: that captures the available services and capacity of each health facility, information that is essential for health system planning and management, and resource allocation.

The signature domain component is the one containing the critical set of fields to support the integration of geography through time into the health information system. Among those fields, the following should be considered as mandatory:

1. Health facility identifier: Official unique identifier attached to the health facility
2. Health facility name: Official name of the health facility in both English and local language
3. Address: street name and number in which the health facility is located (when applicable)
4. Location in the administrative structure: Official name and code of the administrative divisions in which the health facility is located and this down to the lowest existing administrative level (a separated set of fields for each level)
5. Location in the reporting structure: Official name and code of the reporting divisions in which the health facility is located and this down to the lowest existing reporting level (when applicable and with a separated set of fields for each level)
6. Geographic coordinates: latitude and longitude of the health facility expressed in decimal degrees together with the indication of the source, method and accuracy level attached to the coordinates (a separate field for each information)

---

<sup>6</sup> <https://www.measureevaluation.org/resources/publications/wp-07-91>

<sup>7</sup> <https://www.measureevaluation.org/resources/publications/ms-15-99>

<sup>8</sup> [http://www.who.int/healthinfo/systems/WHO\\_CreatingMFL\\_draft.pdf](http://www.who.int/healthinfo/systems/WHO_CreatingMFL_draft.pdf)

<sup>9</sup> <https://ohie.org/facility-registry/>

When it comes to the unique identifier, each country might use a different coding scheme but it is important to remember that there are best practices that should be considered, such as:

- Using a check digit, such as the luhn algorithm<sup>10</sup> (e.g., 1000001-3) or 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.
- Ensuring for the code schemes to be meaningless, meaning not to embed codes or other semantic meaning in an identifier (e.g., admin level, health facility type, etc.).
- Using a sequence as short as possible but taking into account the number of changes that could take place over the coming decades (function of the current number of health facilities observed in the country)
- Having the ID generated sequentially. (e.g., 100001, 100002, 100003,... or HF00001, HF00002, HF00003,...)

In addition to that, adding legacy identifiers from others systems can significantly improve interoperability and reduce the effort required during the first phase of development of common geo-registry, but efforts should focus towards the use of a common unique identifier across the entire health sector.

Apart from the above, and while not directly linked to the geographic dimension of the health facilities, the following fields (partly included in the signature domain) are also key to uniquely identify and contact health facilities and therefore recommended for inclusion in the master list:

1. Health facility type: Type of the health facility based on the official classification used in the country
2. Health facility ownership or managing authority: Major (government, private) and sub classification (Ministry of Health, army, police,...) describing the entity that owns or manage the health facility. (may require for the use of separated fields)
3. Past health facility names: Official names previously attached to the health facility (may require for the use of several fields)
4. Health facility contact information: Information needed to contact the health facility through different media including but not limited to the full name and position of the health facility head as well as the phone numbers (mobile, landline) and email address (to be captured in separated fields)
5. Health facility status: to be able capturing in the common geo-registry if the health facility is active or inactive and the reason for inactivity if the later applies (to be captured in separated fields)

Any other fields, including those being part of the service domain component, are considered as optional in the context of the presence guidance. This being said, the use of common geo-registry allows for the capture and management of such fields if needed.

---

<sup>10</sup> [https://en.wikipedia.org/wiki/Luhn\\_algorithm](https://en.wikipedia.org/wiki/Luhn_algorithm)

The final list of identified fields should be captured in a data dictionary<sup>11</sup> similar to the one reported in Appendix 3. In order to be useful, such model should at least contain the following information for each of the final fields:

1. Field identifier as is used within the geo-registry (FIELD ID)
2. Field name as it will be implemented in the master list (FIELD NAME)
3. Description of the field content (FIELD DESCRIPTION)
4. Type of character that will be captured (CHAR TYPE)
5. Size of the field expressed in number of characters (FIELD SIZE)
6. The description of the specifications attached to the field (SPECIFICATIONS)
7. The indication of the fields that are mandatory when adding a new health facility in the master list (MANDATORY FOR NEW)

Please note that:

- It is strongly recommended for the source of the information and information collection date attached to each field to be captured in the master list.
- While the Ministry of Health would most likely govern the establishment, maintenance, update and sharing of the health facility master list,, the information it contains might come from different sources. For example:
  - The official codes and names for the administrative divisions should come from the governmental agency in charge of them (generally the Ministry of Interior or the Ministry of Home Affairs) through the content of the administrative division master list (see next section)
  - The geographic coordinates can be collected by different stakeholders
- the time stamp for the Unique identifier corresponds to the data when the health facility was opened while the same timestamp corresponds to the date of last update for all the other fields;
- The official codes and names for the administrative divisions are meant to come from the administrative divisions master list (see next section)

The master list can then represent all the fields reported in the data dictionary, allowing for a source (“\_S” at the end of the column label) and a data collection timestamp (“\_T”) field to be attached to all of them . Table 3 provides an extract of the master facility list based on the data dictionary reported in Appendix 3.

---

<sup>11</sup> A set of information describing the contents, format, and structure of a database

**Table 3.** Extract of the resulting database after implementing the data dictionary reported in Appendix 3

HF_ID	HF_ID_S	HF_ID_T	HF_NAME_EN	HF_NAME_EN_S	HF_NAME_EN_T
HF000062	MOHS	7-27-2001	Ywar Thar	MOHS	7-27-2001
HF002135	MOHS	3-27-1998	Zee Pin Hla	MOHS	3-27-1998
HF004754	MOHS	4-29-1972	Myo Thit	MOHS	4-29-1972
HF003336	MOHS	7-1-1983	Ya Thit	MOHS	7-1-1983
HF003302	MOHS	6-17-2005	Let Wea	MOHS	6-17-2005
HF000008	MOHS	6-15-1934	Let Pan Taw	MOHS	6-15-1934
HF002212	MOHS	6-6-1969	Mei Ni Ma Kone	MOHS	6-6-1969
HF004708	MOHS	8-31-2009	Zee Kyun	MOHS	8-31-2009
HF000469	MOHS	6-15-1999	Myay Ni Kone	MOHS	6-15-1999
HF004090	MOHS	7-12-2002	Let Pan Kyun	MOHS	7-12-2002
HF002204	MOHS	6-17-2009	Let Pan Taw	MOHS	6-17-2009
HF002021	MOHS	9-12-1943	Kan Su	MOHS	9-12-1943
HF002790	MOHS	7-29-1985	Na Nwin Taw Boe	MOHS	7-29-1985

Two additional important pieces of information need to be attached to the master list, namely:

1. A simple **data catalog** for the user to understand the content of each field included in the master list. Such catalogue could for example consist in the content of the “FIELD NAME” and “FIELD DESCRIPTION” fields from the data dictionary reported in Appendix 3
2. **Metadata** for the user to understand the content of the health facility master list and have a way to contact the person or institution in charge if needed. Ideally such metadata should at least contain:
  - a. A Dataset title
  - b. The name of the Originator
  - c. The Publication date and version number
  - d. A short abstract
  - e. Any data licensing for access and use constraints, licensing
  - f. Any potential disclaimer
  - g. A Point of contact details - name, organization and contact details
  - h. A Web URL for data repository location when applicable

The above therefore requires for the use of a file format that allows for multiple worksheets (e.g, MS Excel) in order for both of these pieces of information to remain attached to the master list when being shared.

### ***Administrative divisions master list***

In order to fulfill its central role once uploaded in the common geo-registry as well as being useful from a data management perspective, the administrative divisions master list needs to be structured in such a way that it captures:

1. The official and complete list of administrative divisions down to the village level;
2. The official unique identifier (code) attached to each division and this for all the level
3. The geographic coordinates of each village in case village boundaries are not available;
4. The administrative division hierarchy. As such the master list should be able capturing the unique identifier of the parent division (upper level in the hierarchy):
5. The relationship that exists between villages and health facilities to build catchment areas as part of the reporting divisions hierarchy.

In addition to the above, the master list should preferably be organized in a way than its content would directly correspond to the content of the attribute table of the GIS format layer containing the administrative boundaries divisions.

The easiest way to reach the above is to organize the content of the master list as per the examples reported in:

- Table 4 when village boundaries are available in GIS format;
- Table 5 when village boundaries are not available and villages therefore have to be represented by points (geographic coordinates).

The data dictionaries for both options are themselves reported in Appendix 4.

**Table 4.** Example of administrative divisions master list when village boundaries are available in GIS format (Philippines)

Reg Code	Reg Name	Prov Code	Prov name	Mun Code	Mun name	Bgy Code	Bgy Name	HF_Cat_Code
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280100	ADAMS	PH012801001	Adams (Pop.)	HR0001
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802001	Bani	HR0002
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802002	Bayon	HR0003
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802003	Cabanuan	HR0004
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802004	Cabulisan	HR0005
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802005	Cabusigan	HR0006
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802006	Cadaranan	HR0007
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802007	Calook-Libong	HR0008
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802008	Casilan	HR0009
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802009	Coracor	HR0010
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802011	Duripes	HR0011
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802012	Genapan	HR0012
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802013	Liblong	HR0013
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802014	Macupit	HR0014
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802015	Nambaran	HR0015
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802016	Naba	HR0016
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802017	Paninan	HR0017
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802018	Pasiocan	HR0018
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802019	Pasogai	HR0019
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802020	Pipias	HR0020
PH01000000	REGION I (ILOCOS REGION)	PH01280000	ILOCOS NORTE	PH01280200	BACARRA	PH012802021	Pulang	HR0021

**Table 5.** Example of administrative divisions master list when village boundaries are not available in GIS format (Cambodia)

PRO C MO	PRO N MO	DIS C MO	DIS N MO	COM C MO	COM N MO	VIL C MO	VIL N MO	Lat	Long	Source Lat Long	Acc Lat long	HF CAT CODE
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020301	Ou Thum	13.51366	103.03564	MOI (GPS)	High	HF0001
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020302	Phnum	13.51352	103.03809	MOI (GPS)	High	HF0002
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020303	Banteay Neang	13.51291	103.03822	MOI (GPS)	High	HF0003
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020304	Kouk Phoy	13.50461	103.00895	MOI (GPS)	High	HF0004
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020305	Trang	13.51004	103.00983	MOI (GPS)	High	HF0005
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020306	Pongro	13.50158	103.00564	MOI (GPS)	High	HF0006
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020307	Kouk Tonloab	13.49329	103.02461	MOI (GPS)	High	HF0007
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020308	Timbaek	13.49029	103.02754	MOI (GPS)	High	HF0008
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020309	Khleik	13.48756	103.02479	MOI (GPS)	High	HF0009
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020310	Samrong Pon	13.49214	103.03214	MOI (GPS)	High	HF0010
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020311	Dang Run Lech	13.49348	103.03541	MOI (GPS)	High	HF0011
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020312	Dang Run Kaout	13.49765	103.04410	MOI (GPS)	High	HF0012
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020313	Ou Snguot	13.46671	103.02665	MOI (GPS)	High	HF0013
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020314	Prey Chingha Lech	13.46738	103.02795	MOI (GPS)	High	HF0014
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020315	Prey Chingha Kaout	13.46762	103.03315	MOI (GPS)	High	HF0015
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020316	Ou Andoung Lech	13.46794	103.03994	MOI (GPS)	High	HF0016
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020317	Ou Andoung Kandsi	13.46612	103.04059	MOI (GPS)	High	HF0017
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020318	Ou Andoung Kaout	13.46901	103.05462	MOI (GPS)	High	HF0018
1000000	Banteay Meanchey	3020000	Mongkol Borei	1020000	Banteay Neang	3020319	Kouk Kduoch	13.47076	103.07202	MOI (GPS)	High	HF0019

Please note that:

1. While the official code and names of the administrative divisions should come from the government agency in charge of them (generally the Ministry of Interior or the Ministry of Home Affairs) the official code of the health facility allowing to build the catchment area (HF\_CAT\_CODE in Tables 4 and 5) should be included by the Ministry of Health;
2. The comments mentioned for the health facility unique identifier (see previous version) are also applicable in the case of the administrative divisions.
3. A data catalog and metadata worksheet should be attached to the administrative divisions master list like for the health facility master list
4. The label for each field as well as the content of the data catalog allows to know the type of divisions observed at each level of the administrative structure (Region, Province, District, Commune,...)
5. When possible, the spelling of each division in the local language(s) should also be added to the master list

### **Reporting divisions master list**

Two main types of reporting divisions need to be distinguished:

- The catchment areas which corresponds to the area served by a healthcare facility or health plan. The other types of reporting divisions that can sometime be encountered in countries (e.g. health districts, operational districts,...)
- The other types of reporting divisions that can sometime be encountered in countries (e.g. health districts, operational districts,...)

Catchment areas are present in all countries. Their geographic extent correspond to sum of the the villages being attached to a particular health facilities. As such, the recommendation is to capture the link between each village and the health facility it is attached to directly in the administrative division master list (see previous section).

In the majority of the cases, the other types of reporting divisions do not present several levels like administrative divisions do but can be related to such administrative divisions and this in different ways. Namely:

- They can be composed of administrative divisions from a particular level (e.g. health districts composed of communes)
- They can sum up to match the boundaries of a particular administrative division level (see example in Figure 2)

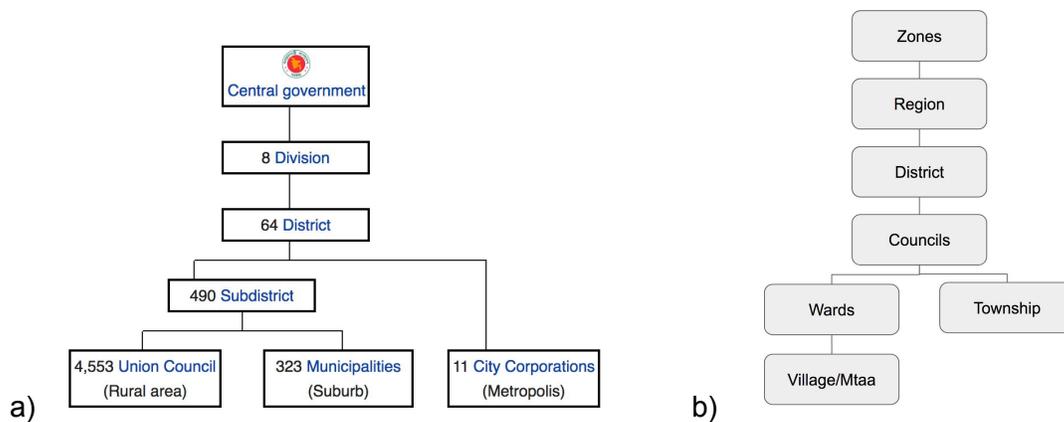
The reporting divisions master list to be uploaded in the common geo-registry should therefore not only contain the unique code and names of all the reporting divisions observed in the country but also, when applicable, the existing links with the administrative division structure as indicated here above.

The master list in question should therefore be organized in the same way as the administrative divisions master list (Table 4 or 5) and a data catalog and metadata attached to it as well.

## Specifications for the associated data

### *Geographic hierarchies*

For either the administrative or the reporting divisions it is important to capture the hierarchy that allows understanding how they are organized across the different levels of the administrative or reporting structure. Such hierarchy can simply be captured through a schema such as those presented in Figure 4.



**Figure 4.** Examples of administrative divisions hierarchies for a) Bangladesh<sup>12</sup> and b) Tanzania.

<sup>12</sup> [https://en.wikipedia.org/wiki/Administrative\\_geography\\_of\\_Bangladesh](https://en.wikipedia.org/wiki/Administrative_geography_of_Bangladesh)

## ***Historic Changes***

A core function of the common geo-registry is to track changes made over time.

Changes happening at the health facility level are limited and easy to capture through the use of a simple form/table containing the following set of fields depending on the type of change to be implemented in the health facility master list:

- Opening of a new health facility:
  - Minimum set of fields for adding a new health facility in the master list (See Appendix 3 for an example)
  - Opening date for the facility
- Closing/inactivating a health facility:
  - Official code and name of the health facility in the master list
  - Reason for closing/inactivating the facility in the master list
  - Date at which the facility has been closed/made inactive
- Modifying/updating attributes:
  - Official code and name of the health facility in the master list
  - Field for which the information should be modified/updated
  - Information currently contained in the field in question
  - New information that should be entered in the field in question
  - Time stamp for the change of information
- Removal of a duplicate:
  - Official name of the facility in question
  - Official code for the first record in the database
  - Official code for the record considered as duplicate

Please note that the following information should also be captured and stored in the common geo-registry logbook for each change implemented in the master list:

- Full name of the requestor
- Request date
- Source of the information on which the request is based

Rebuilding the information for a given health facility at a given moment in time is then possible thanks to the logbook and the data collection time stamp attached to each field.

Capturing, visualizing and keeping track of the changes occurring for the administrative and reporting divisions is a more complex exercise.

The historic changes for these divisions are usually captured in a list similar to the one reported in Table 6 for the Philippines<sup>13</sup>. Such list generally capture the following information:

- The date of the change

---

<sup>13</sup> <http://nap.psa.gov.ph/activestats/psgc/>

- The type of change (split, merge, transfer, change of name,...)
- The unit that is concerned by the change
- Any potential change in unique identifier/code
- The mention of the legal document in which the change is recorded

**Table 6.** Example of list capturing administrative divisions historic changes (PSGC, Philippines)

January - March 2001 Updates					
Region/Province/Municipal/Barangay Name	Unit Type	New Code	Mother Unit/Old Name	Old Code	Description/Legal basis
Province of Zamboanga Sibugay, Region IX	Newly created province	098300000			R.A. No. 8973
Municipality of Alicia, Zamboanga Sibugay	Transferred Municipality	098301000	Zamboanga del Sur	097301000	R.A. No. 8973
Municipality of Buug, Zamboanga Sibugay	Transferred Municipality	098302000	Zamboanga del Sur	097304000	R.A. No. 8973
Municipality of Diplahan, Zamboanga Sibugay	Transferred Municipality	098303000	Zamboanga del Sur	097339000	R.A. No. 8973
Municipality of Imelda, Zamboanga Sibugay	Transferred Municipality	098304000	Zamboanga del Sur	097336000	R.A. No. 8973
Municipality of Ipil, Zamboanga Sibugay	Transferred Municipality	098305000	Zamboanga del Sur	097309000	R.A. No. 8973
Municipality of Kabasalan, Zamboanga Sibugay	Transferred Municipality	098306000	Zamboanga del Sur	097310000	R.A. No. 8973
Municipality of Mabuhay, Zamboanga Sibugay	Transferred Municipality	098307000	Zamboanga del Sur	097314000	R.A. No. 8973

While the above list is useful for implementing the changes in the corresponding master list, it is not easy for the user to visualize how a particular division has evolved over a given period of time. To address this issue, it is recommended to also capture the changes under the form of a historic changes tables similar to the one reported in Table 7.

**Table 7.** Example of table for capturing administrative divisions historic changes (Philippines)

Period	01.01.2000-21.02.2001		22.02.2001-18.09.2001	
Unit type	Provinces			
	Basilan	090700000	Basilan	090700000
	Zamboanga del Norte	097200000	Zamboanga del Norte	097200000
	Zamboanga del Sur	097300000	Zamboanga del Sur	097300000
			Zamboanga Sibugay	098300000

According to the National Statistical Coordination Board, the province of Zamboanga Sibugay was created from Zamboanga del Sur on 22.02.2001.

Please note that:

- A separated historic changes table need to be developed for each of the level in the administrative structure (province level in Table 7)
- Each time period is captured in a separated set of columns (e.g. 01.01.2000-21.02.2001 in Table 7)
- Change between periods are highlighted (in blue)
- The legal document in which the change is recorded is mentioned in a comment

Both of the approaches mentioned here are complementary and should ideally not only be maintained by the institution having the mandate of capturing them but also at the disposal of the person in charge of integrating these changes into the common geo-registry and their users.

In addition to the above, we can't talk about historic changes without talking about the type of coding scheme being used to uniquely identify administrative or reporting divisions and how this code evolves when a change occur.

We can distinguish two types of coding schemes:

1. Those that follow a postal approach when a change occur. In this case the practice is that a division that keeps the same name after the change also keeps its unique

identifier code. An example of such practice in the case of a series of splits is reported in Table 6. This kind of scheme is the one being mostly used by governments as it simplifies their administrative work, for budget allocation follow up for example. The major disadvantage of this scheme from a data management point of view, is that it is difficult to identify to which time period the code refers to without having a time stamp associated to it.

**Table 6.** Example of implementation of the postal approach

Period 1		Period 2		Period 3	
Name	Codes	Name	Codes	Name	Codes
Blue	121008	Blue	121008	Blue	121008
				Yellow	121010
		Red	121009	Red	121009

- Those that follow a data management friendly approach in the case of a change. Here, the practice is that a new code is attributed to the division each time that its geography changes significantly. This practice is illustrated in Table 7 using the same set of splits than those used in Table 6. The difference between the two approaches can be observed with the “Blue” division. The main advantage here is therefore that the code does directly carry the time period, making it easier to identify which representation of a particular division we are referring to. This scheme is generally been used in projects that have a strong data management focus like the Second Administrative Level Boundaries (SALB) data set project which documented how such type of scheme is meant to evolve for any changes that can be observed in countries <sup>14</sup>.

**Table 7.** Example of data management based approach

Period 1		Period 2		Period 3	
Name	Codes	Name	Codes	Name	Codes
Blue	A	Blue	A1	Blue	A2
				Yellow	C
		Red	B	Red	B

In view of the above, the choice of the coding scheme being used will have an important on the information that will need to be captured in the common geo-registry in order to be able rebuilding the historic changes over a given period of time, more specifically:

<sup>14</sup> [www.gaiageosystems.org/PROJECTS/SALB/Update\\_SALB\\_codes\\_250609.pdf](http://www.gaiageosystems.org/PROJECTS/SALB/Update_SALB_codes_250609.pdf)

1. The use of the postal approach will require not only to capture the official name and code of division after each change but also (Table 8):
  - a. An internal code that follows the data management friendly approach (in green). This code will have to be created and maintained in addition to the official one
  - b. The indication of the mother unit (division) based on the data management friendly code
2. The use of the data management friendly approach on the other hand will only require for the indication of the mother unit to be captured in addition to the official code and name (Table 9).

In conclusion, the use of a data management friendly approach makes it easier to capture administrative or reporting divisions historic changes in a common geo-registry.

**Table 8.** Information to be captured in the common geo-registry when following the postal approach

Period 1		Period 2		Period 3		
Name	Codes	Name	Codes	Name	Codes	
Blue	Official code: 121008	Blue	Official code: 121008	Blue	Official code: 121008	
					Internal code = A2	
					Mother unit = A1	
	Internal code = A	Red	Red	Official code: 121009	Red	Official code: 121010
						Internal code = A1
						Mother unit = A
Internal code = A	Red	Red	Official code: 121009	Red	Internal code = C	
					Mother unit = A	
					Mother unit = A	
Internal code = A	Red	Red	Official code: 121009	Red	Internal code = B	
					Mother unit = A	
					Mother unit = A	

**Table 9.** Information to be captured in the common geo-registry when following the data management friendly approach

Period 1		Period 2		Period 3			
Name	Codes	Name	Codes	Name	Codes		
Blue	Official code: A	Blue	Official code: A1	Blue	Official code: A2		
					Mother unit = A1		
					Mother unit = A		
		Red	Red	Red	Official code: B	Red	Official code: C
							Mother unit = A1
							Mother unit = A
Red	Red	Red	Official code: B	Red	Official code: B		
					Mother unit = A		
					Mother unit = A		

### ***Available boundary layers in GIS format***

The geographic location of health facilities and of villages (when villages boundaries are not available in GIS format) is directly captured in the respective master list through their geographic coordinates (latitude and longitude)..

The boundaries of the administrative and reporting divisions can't themselves be captured in the master list and therefore have to be stored in a separated GIS readable format file, a layer.

In addition to that, the changes that occurs through time for these divisions (see previous section) requires having access to a different boundary layer for each time period.

As such, the common geo-registry should be able to redirect the user to any available GIS format layer for each given time period.

The most effective way to provide the user with such information is to maintain a table containing the following minimum set of fields for the layers in question:

- Time period start date
- Time period end date
- Number and types of levels included
- Source
- Access and/or use restrictions
- Layer format (shape file, json)
- Link to the layer

Any other information of importance to the user, such as scale and accuracy, projection information or contact details, should themselves be included in the metadata associated to each layer.

## **Data flow and requirements for the Common geo-registry**

In previous sections, we identified the geographic objects core to public health as well as the attributes that should be included in the different master lists in order to geo-enable the HIS.

In this section, we will discuss the data flow, features and requirements that a common geo-registry must respectively provide and comply with in order to effectively host, manage, update and share these master lists for the benefit of the whole community.

## Data Flow

Let's start by analyzing the data lifecycle of a common geo-registry for the geographic objects core to public health identified earlier in the document, namely: health facilities, administrative divisions, and reporting divisions.

At the beginning of the flow we have the following datasets that have been developed based on the guidelines reported in the previous sections of the present document when it comes to their format, structure and content:

- An initial health facility master list. This list might not be perfect but it mature enough for being uploaded in the common geo-registry and therefore benefit from its functionalities.
- A master list of administrative divisions. For the geo-registry to be effective, this list should be complete, up-to-date and uniquely coded down to the village level.
- A master list of reporting divisions. Like for the administrative divisions, this list should be complete, up-to-date and uniquely coded to benefit from all the functionalities offered by the registry.
- A set of schemas describing the geographic hierarchies observed in the country.

From this point, the following is expected to happen inside the common geo-registry:

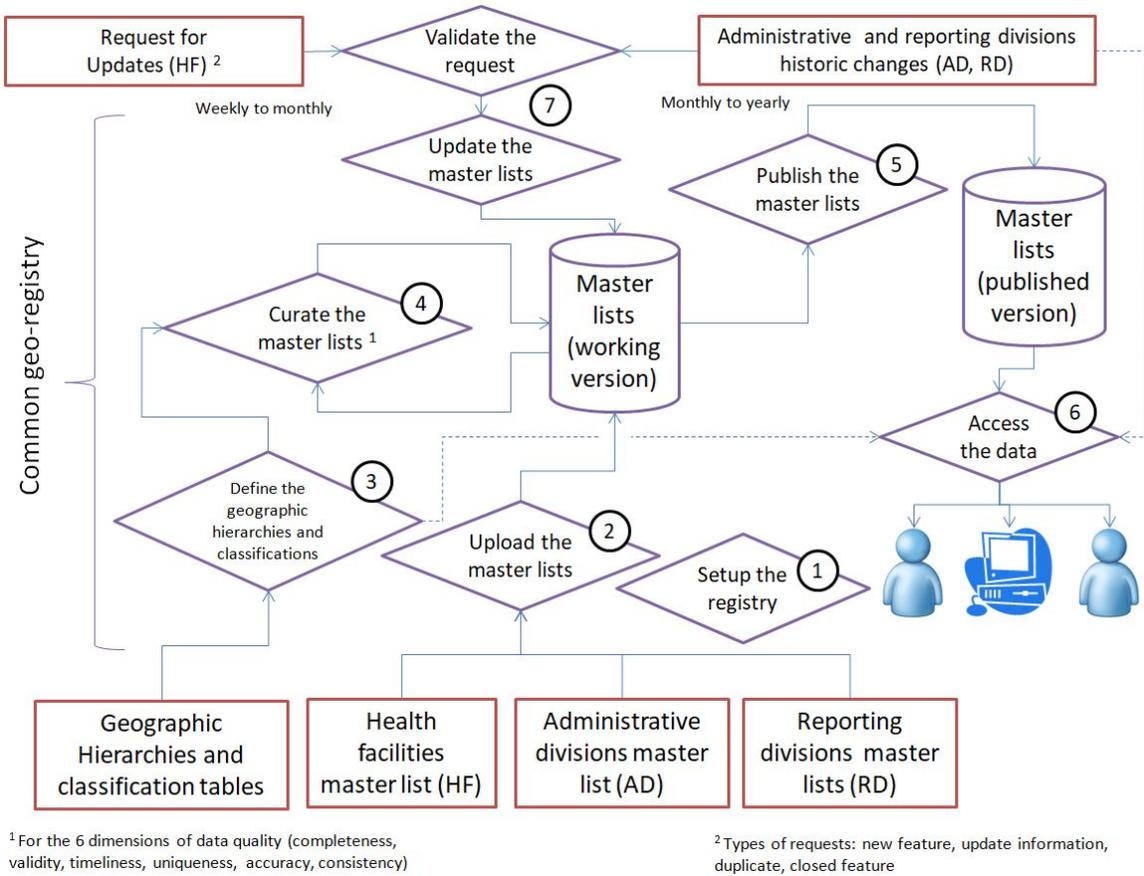
1. **Initial Setup:** This step consists of attributing the core roles and corresponding responsibilities to users who will have the mandate over the main functions of the common geo-registry. The roles in question are as follow:
  - a. Registry Maintainer (RM): a user who has the permission to upload master lists, define geographic hierarchies, curate the master lists, validate and integrate updates in the master lists based on the request that have been submitted.
  - b. Registry Administrator (RA): a user who has all the privileges of a RM plus the permission to manage other users' permissions and to publish new versions of the master lists.

It is important to mention here that due to the fact that the master lists are under the mandate of different Ministries, several users might have the role of registry maintainer (RM) and therefore the editing privilege over only one master list.

2. **Upload:** The master lists are being uploaded into the geo-registry and the registry is able to understand how these lists are structured and the content of each of field that composes them. This can be done in two ways:
  - a. The user provide information about the structure and content of each master list before their upload and then each list is uploaded through a bulk upload.
  - b. The list are uploaded first and then the user specifies which fields contains the different informations for the registry to effectively function through the use of a wizard.

3. **Definition of the geographic hierarchies:** before being able to curate the different master lists it is necessary for the registry to understand the geographic hierarchies that exist in the country and how these are captured in the different master lists that were uploaded.
4. **Curation:** now that the first set of master lists are uploaded to the common geo-registry, the registry understands the structure and content of these master lists and how the geographic hierarchies are being captured, it is possible to perform different checks, first internally to each list and then between lists in order to identify gaps, potential duplicates, inconsistencies across master lists, etc. So at this stage we need the platform to support our goal of transforming the working version of the master lists into something that could be published.
5. **Publication:** once curated, the master lists and other associated information (geographic hierarchies, historic changes) can be **published** in a separated database to make it accessible to different groups of users.
6. **Access:** Once published, the master lists and the associated information can be made publicly and easily accessible by users through different media including but not limited to a Graphical User Interface (GUI) or one or more Application program interface (APIs). In other words, we want the platform to make it easy for users to **access the master lists and other associated information**. This step in the data flow will require for two more roles to be managed by the common registry, namely:
  - a. Anonymous User (AU): any user who hasn't been identified but has the possibility to consume the public data/information shared through the registry.
  - b. API Consumer (AC): a user with permission to consume the registry's APIs.
7. **Update:** while the **Curation** process is managed by the organization or organizations having the mandate over master lists, the process of updating may be a shared responsibility among the stakeholders mandated to or interested in ensuring the continuous quality of these masters lists. The platform serving as common geo-registry should therefore allow selected users and organizations to issue requests for update, detailing what needs to be changed, whether they detected duplicates, etc. It should also let those managing the geo-registry **validate and implement the updates**, making it easy to follow up on them and to mark them as accepted or rejected. This step leads to the definition of one more role to be managed by the common registry:
  - a. Registry Contributor (RC): a user who can issue update requests to the registry.

To help you visualize how all these elements interact, we put laid them out in the diagram below (Figure 5), the number corresponding to the bullet points reported here above.



**Figure 5.** Generic data flow for the common geo-registry

Table 10 complements the above figure by indicating in which step of the data flow each user role intervenes. Please note that, depending on how the solution is implemented, the same user might be assigned one or more of these roles.

**Table 10.** User’s intervention by step in the data flow

	RA	RM	RC	AU	AC
1. Initial setup	X				
2. Upload	X	X			
3. Hierarchies	X	X			
4. Curation	X	X			

5. Publication	X				
6. Access	X	X	X	X	X
7. Update	X	X	X		

## Requirements from the perspective of the different user roles

Let's now look at the same data flow within the common geo-registry but from the perspective of the user roles that have been defined previously and this in order to identify the requirements that the common geo-registry should provide to each of them.

We describe these requirements as a list of *user stories*. A user story is an informal, natural description of one or more features of a software system. User stories are often written from the perspective of an end user.<sup>15</sup> Each user story takes the form:

“As a [user role], I want/need to [requirement] in order to [higher level need].”

### **Registry Administrator**

As a registry administrator I want to:

1. Have access to the same functionalities than a Registry Maintainer (RM).
2. Manage users:
  - a. Collect their contact details: name and last name, email address, phone number, organization.
  - b. Activate and deactivate them.
  - c. Assign and modify their role.
  - d. Attribute the management of each list to different users to empower the right governing authorities.
  - e. Specify authority on the different master lists to avoid for non-authorized users, including the Registry Maintainers (RM), to perform any changes on those lists.
3. Consume master lists that are managed in other platforms. If an existing list is managed elsewhere, I should be able to easily connect to that other system to avoid disruption.
4. Publish the current working version of any of the lists once it has passed all the registry checks.
5. See every significant activity in the platform to be logged, so that I can reconstruct what happened in case something goes wrong.

### **Registry Maintainer**

As a registry maintainer I want to:

---

<sup>15</sup>[http://en.wikipedia.org/wiki/User\\_story](http://en.wikipedia.org/wiki/User_story)

1. Define the structure and content (schema) of each master list in a way that the common geo-registry understand them. This would include:
  - a. Specify a type for each field (integer, string, date,...) so that I can check the validity of the master list being uploaded in the registry.
  - b. Indicate which fields are supposed to contain unique values.
  - c. Indicate which field contains the information necessary for the registry to build the hierarchies and curate the master lists, in particular:
    - i. Administrative divisions master list:
      1. Unique identifier and names of each level in the administrative division hierarchy.
      2. Health facility unique ID to build the catchment areas.
      3. Village geographic coordinates (lat/long) if captured by a point.
      4. Source and method for the geographic coordinates.
      5. Accuracy level of the geographic coordinate.
    - ii. Reporting divisions master list:
      1. Unique identifier and names of each level in the reporting division hierarchy.
    - iii. Health facilities master list:
      1. Unique identifier and names of each health facility.
      2. Unique identifier and names of each level in the administrative division hierarchy.
      3. Unique identifier and names of each level in the reporting division hierarchy.
      4. Geographic coordinates (lat/long).
      5. Source and method for the geographic coordinates.
      6. Accuracy level of the geographic coordinate.
2. Specify whether:
  - a. The villages are being represented by points or polygons in the country.
  - b. The coding scheme used for administrative and reporting divisions follow a postal or data management approach.
3. Define the geographic hierarchies for the administrative and reporting divisions based on the master lists contents, so that I can use them to enforce data quality when new data gets into the Registry. This would include:
  - a. The specification of the fields containing each of the “parent-child” pairs.
  - b. The indication of the division type for each level observed in the hierarchy (Region, Province, District,...).
4. Upload master lists stored in Excel files to an empty Registry.
  - a. Specify how the columns in the Excel files map to the Schemas I previously defined.
  - b. Integrate the metadata and data catalogue associated to the master list into the registry.
  - c. Have the Registry:
    - i. Use the Schemas to validate data quality.

- ii. store the data as is so I can curate it in multiple sessions.
  - iii. prevent me from publishing the working version until it has been completely curated and the checks pass.
- 5. Upload and maintain a table providing the availability of administrative and/or reporting divisions boundaries layers.
- 6. Have the Registry continuously perform data quality validations to the current working versions of lists, so that I can only publish them when they comply with all the data quality requirements (link matrix).
  - a. Be notified when upstream master lists<sup>16</sup> cause breaking changes, so I can resolve the issues as soon as possible.
  - b. Be able to work on the working version even while it is not passing validations.
  - c. Be assisted in resolving validation issues: I want the Registry to inform me of each concrete validation issue and provide me with shortcuts to solve each one.
- 7. See a list of pending and accepted requests for Update.
  - a. I can accept a request for update and provide any follow up details.
  - b. I can reject a request for update and provide reasons for the rejection.
  - c. I can easily contact the RC who issued the request to get complementary information.
  - d. Have each request stored in a logbook that I can consult.
- 8. Implement the historic changes in the administrative and/or reporting division master list:
  - a. Split a division.
  - b. Merge divisions.
  - c. Promote divisions to a higher level in the hierarchy.
  - d. Downgrade divisions to a lower level in the hierarchy.
  - e. Transfer divisions from a parent to another.
  - f. Modify the hierarchy (to add a new administrative level for example).
  - g. Specify when these changes start being valid.

### ***Registry Contributor***

As a registry contributor I want to

1. Issue requests for Update as follow:
  - a. Health facilities: new health facility, update attribute information, duplicate, closed facility.
  - b. Administrative and reporting division: Missing divisions, historic change, spelling issue, duplicate.
  - c. Geographic hierarchies: error in parent-child attribution.
2. See a list of my Requests for Update and how they were resolved by whom.

---

<sup>16</sup>A master list is considered as being upstream to another one when a change in the geography of the object contain in the former has an impact on the later one (for example, a change in the administrative divisions master list might impact the distribution of the health facilities across these divisions => the administrative division master list is upstream to the health facility one.

## ***API Consumer***

As an API Consumer I want to:

1. Access the master lists in the format recommended in the present guidance through:
  - a. REST APIs:
    - i. These APIs allow me to paginate the lists, specifying a page size.
    - ii. I can specify a moment in time so I can get a historical view of the master lists. I want it to default to the current version of all the lists if I don't specify a moment in time.
    - iii. There are specific endpoints to download the lists and hierarchies stored in the Registry.
  - b. Streaming API, so I can directly get changes since a given time and adjust my system incrementally.
2. Access the associated data through:
  - a. A schema presenting the geographic hierarchy followed for the administrative and reporting divisions and this for a given point in time
  - b. The list of historic changes that have been implemented in the administrative and reporting divisions master list for a given period. Such information should be presented in a format that would allow me to easily rebuild these changes .
  - c. The list of available shapefiles, if any, for the administrative and/or reporting divisions for a given period in time.

## ***Anonymous User***

As an Anonymous User I want to:

1. Download:
  - a. The master lists as self-contained Excel files presenting the structure and content recommended in the present guidance.
  - b. A schema presenting the geographic hierarchy followed for the administrative and reporting divisions and this for a given point in time.
  - c. The list of historic changes that have been implemented in the administrative and reporting divisions master list for a given period. Such information should be presented in a format that would allow me to easily rebuild these changes.
  - d. The list of available shapefiles, if any, for the administrative and/or reporting divisions for a given time period.
2. Specify a moment in time so I can get a historical view of the master lists. I want it to default to the current version of all the lists if I don't specify a moment in time.
3. Navigate all elements in the published version of the Registry on a map widget.
4. Use the geographic coordinates stored in the administrative boundaries (if villages are captured as points) and/or health facility master list(s) to see their location on a map in the common registry.

5. Filter all elements in the lists by a given time or a specific level in the administrative or reporting division hierarchy
6. Search all elements at once by entering text to match in a textbox. The search should return any list item that matches the text.

## Requirements specific to each node in the data flow

This section has for objective to provide a more exhaustive and detailed overview of the requirement that the common geo-registry must (mandatory), should (recommended) or may (good idea)<sup>17</sup> provide in order to enable organizations to effectively manage the data flows and user’s needs described in the previous sections of the document.

### *Initial setup of the Registry*

The common geo-registry:

<b>Must</b>	<b>Should</b>	<b>May</b>
<ul style="list-style-type: none"> <li>• Implement user stories number 2 and 5 for the Registry Administrator (RA).</li> <li>• Implement user stories number 1, 2 and 3 for the Registry Maintainer (RM).</li> </ul>	<ul style="list-style-type: none"> <li>• Implement user story 4 for the Registry Administrator (RA)</li> </ul>	<ul style="list-style-type: none"> <li>• Be flexible enough to accommodate for other master list formats, structures and contents</li> </ul>

### *Upload the Master Lists and define the geographic hierarchies*

The common geo-registry:

<b>Must</b>	<b>Should</b>	<b>May</b>
<ul style="list-style-type: none"> <li>• Implement user story 4 for the Registry Maintainer (RM)</li> </ul>	<ul style="list-style-type: none"> <li>• Implement user story 5 for the Registry Maintainer (RM).</li> <li>• Log who uploaded a master list and when.</li> <li>• Log who made a change to the hierarchy definition when.</li> </ul>	

---

<sup>17</sup> The “must”, “should” and “may” classification used here is based on the keywords for use in RFCs to indicate requirement levels (RFC2119) <https://www.ietf.org/rfc/rfc2119.txt>

## Curate the Master Lists

As we discussed in previous sections, one of the key requirements of a geo-registry is to guarantee the quality of the master lists it publishes across the 6 dimensions of data quality<sup>18</sup>:

1. **Validity:** data is valid if it conforms to the syntax (format, type, range) of its definition.
2. **Completeness:** the proportion of stored data against the potential of "100% complete".
3. **Uniqueness:** nothing will be recorded more than once based upon how that thing is identified.
4. **Timeliness:** the degree to which data represent reality from the required point in time.
5. **Accuracy:** the degree to which data correctly describes the "real world" object or event being described.
6. **Consistency:** the absence of difference, when comparing two or more representations of a thing against a definition.

In order to provide those guarantees, the registry must be able to automatically check whether the current working version of a master list fulfills all the these requirements.

This first requires for the common geo-registry to be able implementing user stories 6 and 8 for the Registry Maintainer (RM).

In addition to that, the following contains the other requirements according to the 6 dimensions for each of the master lists considered here Health facilities, administrative and reporting divisions).

<b>Administrative divisions master list</b>			
	<b>Must check</b>	<b>Should check</b>	<b>May check</b>
Within each record	<p><b>Validity:</b> that the format for each field matches the one stored in the schema</p> <p><b>Completeness:</b> for 100% completeness Regarding the unique identifier and name at each level in the hierarchy</p> <p><b>Timeliness:</b> that the master list is up-to-date</p>	<p><b>Completeness:</b> for 100% completeness for:</p> <ol style="list-style-type: none"> <li>1. The geographic coordinates (lat/long) for villages in case they are captured as points</li> <li>2. The fields that allow building catchment areas if this information is available</li> <li>3. If all the levels indicated in the</li> </ol>	

<sup>18</sup> <https://www.scribd.com/document/262094353/DAMA-UK-DQ-Dimensions-White-Paper-R37>

	at the time of upload  <b>Consistency:</b> that the parent unit exists; parent-child relationship conforms to hierarchy definition	hierarchy are included in the master list  <b>Timeliness:</b> Division creation date; name/code of the mother unit(s)	
Between records	<b>Uniqueness:</b> that there are no duplicates of Unique identifiers at each level	<b>Uniqueness:</b> indicate duplicates of geographic coordinates for the RM to make sure that these are indeed attached to different facilities located in the same building.	<b>Consistency:</b> that there are no cycles in parent-child relationships; that there is only one root element if a file format different than the one recommended here is being used.
Other		<b>Completeness:</b> 100% completeness in the metadata profile and data catalog.	
<b>Reporting Divisions master list</b>			
	<b>Must check</b>	<b>Should check</b>	<b>May check</b>
Individual record	<b>Validity:</b> that the format for each field match the one stored in the schema  <b>Completeness:</b> for 100% completeness regarding of the unique identifier and name at each level in the hierarchy  <b>Timeliness:</b> that it is up-to-date at the time of upload  <b>Consistency:</b> that the administrative division exists	<b>Timeliness:</b> that has a creation date	
Between records	<b>Uniqueness:</b> that the identifiers are unique		<b>Uniqueness:</b> that name is unique

Other		<b>Completeness:</b> 100% completeness in the metadata profile and data catalog.	
<b>Health Facilities master list</b>			
Within each record	<p><b>Validity:</b> that the format for each field matches the one stored in the schema (data dictionary);</p> <p><b>Completeness:</b> for 100% of the fields required to build hierarchies</p> <p><b>Timeliness:</b> up-to-date at the time of upload</p> <p><b>Consistency:</b> that the administrative division exists, that the health administrative office exists</p>	<p><b>Completeness:</b> 100% for lat/long</p> <p><b>Timeliness:</b> that the record includes a source and/or creation date</p>	
Between records	<b>Uniqueness:</b> that the identifiers are unique	<b>Uniqueness:</b> Indicate duplicates of geographic coordinates for the RM to make sure that these are indeed attached to different facilities located in the same building.	<b>Uniqueness:</b> that the name is unique
Other		<b>Completeness:</b> 100% completeness in the metadata profile and data catalog	

***Publish the Master Lists and associated information***

The common registry:

<b>Must</b>	<b>Should</b>	<b>May</b>
<ul style="list-style-type: none"> <li>Implement user stories number 5 for the</li> </ul>	<ul style="list-style-type: none"> <li>Allow consumers to read a subset of the lists</li> </ul>	

<p>Registry Administrator (RA).</p> <ul style="list-style-type: none"> <li>• Keep previous versions of its Master Lists available.</li> <li>• Include a list of changes each time a version is published.</li> </ul>	<p>(e.g.: all facilities in a given province).</p> <ul style="list-style-type: none"> <li>• Log who published the list when.</li> <li>• Implement user story 7 for Registry Maintainers (RM).</li> </ul>	
--	--	--

***Access the Master Lists and associated information***

The common registry:

<b>Must</b>	<b>Should</b>	<b>May</b>
<ul style="list-style-type: none"> <li>• Implement user stories 1 and 2 for the API consumer (AC).</li> <li>• Implement user story 1 for the Anonymous User (AU).</li> </ul>	<ul style="list-style-type: none"> <li>• Implement user stories 2, 3, 4, 5 and 6 for the Anonymous User (AU).</li> </ul>	

***Update the master list after curation of the request***

The common registry:

<b>Must</b>	<b>Should</b>	<b>May</b>
<ul style="list-style-type: none"> <li>• Implement user stories 7 and 8 for the Registry Maintainer.</li> <li>• Implement user stories 1 and 2 for the Registry Contributor (RC).</li> </ul>	<ul style="list-style-type: none"> <li>• Notify the RC user that her request was received, accepted and/or rejected.</li> <li>• Log every activity around the request: who made it, who accepted or rejected it, and what are the changes applied (if any).</li> <li>• Allow entering a change older than the date attached to the initial master list entered in the system.</li> </ul>	

## Concluding words

The primary objective of the present guidance is to provide the data flow and, requirements any platform being used as a common geo-registry should comply to in order to effectively and simultaneously host, maintain, update and openly share the master lists, geographic hierarchies and associated data/ information for the geographic objects core to public health.

The document is organized as follow in order to support the institutions having the mandate over these master lists and the common geo-registry

- Use Cases, Benefits, Challenges and Opportunities: to help make the case for establishing, maintaining, updating and sharing master lists for the geographic objects core to public public through the use of a common geo-registry;
- Core Geographic objects and relationships: to help users in the identification of the objects and relationships core to public health that should be covered by the common geo-registry;
- Specifications for the master lists and associated geographies: to provide recommended on the format, structure and content that the master lists and other data describing their associated geographies should comply to in order to facilitate the management of the all data flow in the common geo-registry;
- Data flow and requirements for the common geo-registry: to support not only the assessment aimed at identifying if the platform currently being used, or planned to be used, as common geo-registry complies to the main requirements you would expect to see in such a platform but also its update or (re-)designing when needed;

While recognizing that this guidance does not cover all the issues that needs to be addressed, it is hoped that it will significantly contribute to a better integration of geography and time in the country's Health Information System (HIS) and therefore allow them to benefit from what geography, geospatial and technologies have to offer in order to improve geographically based decision making and therefore a more systemic approach to solving public health problems.

This document is also meant to be expanded or improved based on the lessons learned from the establishment of common geo-registries in countries.

## Appendix 1 - Glossary of Terms

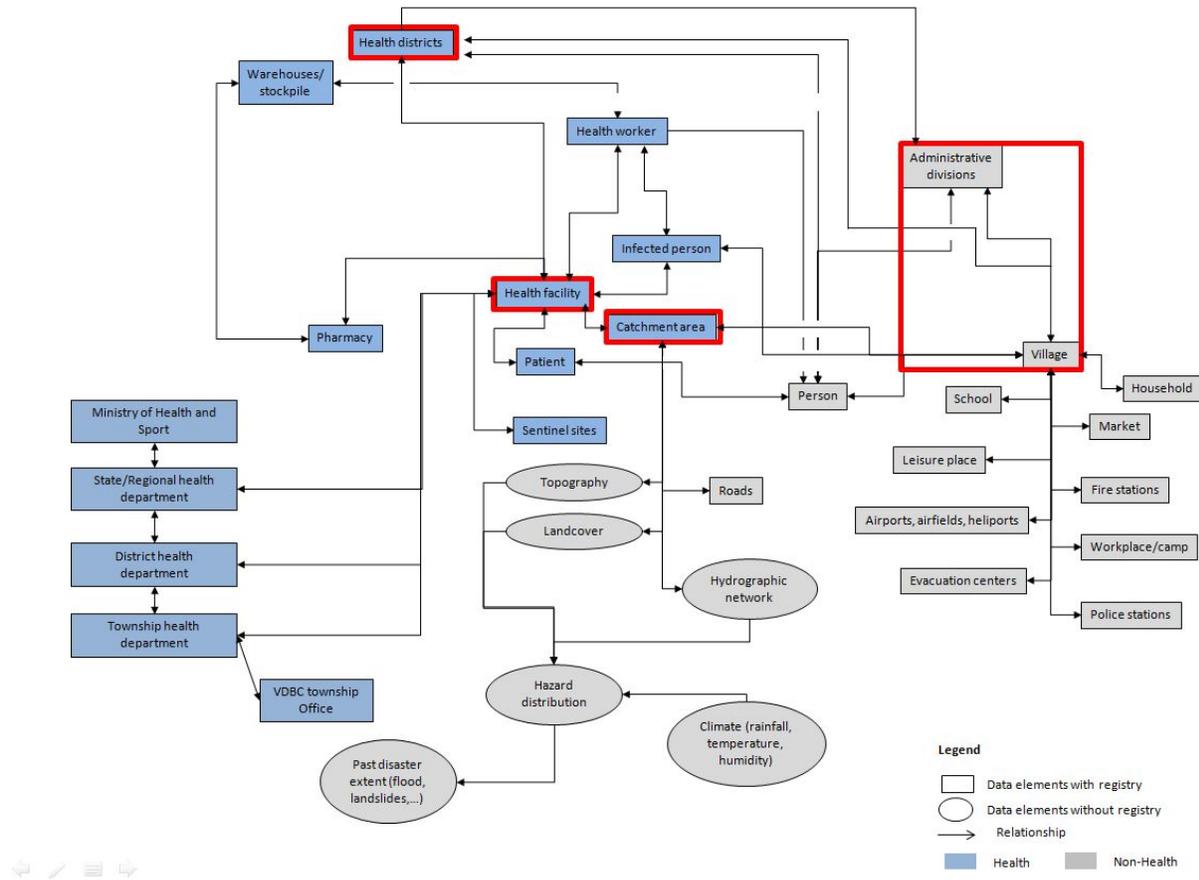
The final version of this appendix will contain the source for each definition. This information can be consulted here:

<https://docs.google.com/spreadsheets/d/1xHQOTGRpOSx-nTSeEfC3XTICcft7CjDI5kEeHcyBYYg/edit#gid=0>

<b>Administrative Division</b>	Portion of a country or other region delineated for the purpose of administration
<b>Authoritative</b>	Having, marked by, or proceeding from authority
<b>Catchment area</b>	A region served by a healthcare facility or health plan, and delineated by population distribution, geography, or transportation patterns
<b>Community</b>	A particular area or place considered together with its inhabitants.
<b>Data model</b>	An abstract model that organizes elements of data and standardizes how they relate to one another and to properties of the real world entities
<b>Functionality</b>	The range of operations that can be run on a computer or other electronic system.
<b>Geographic Information System (GIS)</b>	An integrated collection of computer software and data used to view and manage information about geographic places, analyze spatial relationships, and model spatial processes.
<b>Health Facility</b>	Fixed physical structure where healthcare is provided
<b>Health Information System (HIS)</b>	A system that integrates data collection, processing, reporting, and use of the information necessary for improving health service effectiveness and efficiency through better management at all levels of health services
<b>Hierarchy</b>	Any group of objects ranked so that every one but the topmost is subordinate to a specified one above it.
<b>Master list</b>	Authoritative, standardized, complete, up-to-date and uniquely coded list of all active records for a given object
<b>Metadata</b>	Data that provides information about other data
<b>Registry</b>	IT solution that allows storing, managing, validating, updating and sharing a master list
<b>Reporting divisions</b>	Subnational divisions that are different from administrative divisions and that are used for decision making in the health sector and/or the reporting of health related statistics
<b>Requirement</b>	That which is required; a thing demanded or obligatory:
<b>Rule</b>	One of a set of explicit or understood regulations or principles governing conduct within a particular activity or sphere
<b>Sequence diagram</b>	Interaction diagram showing how objects operate with one another and in what order
<b>Settlement</b>	A place, typically one that has hitherto been uninhabited, where people establish a

	community.
<b>Specification</b>	An act of describing or identifying something precisely or of stating a precise requirement.
<b>User story</b>	Tool used in Agile software development to capture a description of a software feature from an end-user perspective
<b>Workflow</b>	The sequence of industrial, administrative, or other processes through which a piece of work passes from initiation to completion.

## Appendix 2 - Conceptual data model for public health in Myanmar *(Will be in landscape format in the final version)*



## Appendix 3 - Example data dictionary for the health facility master list (Myanmar)

The complete dictionary will be added in the final version. The table in question can be downloaded from here:

[https://www.dropbox.com/s/4fuvr6tbq4avt4l/Appendix\\_3\\_150717.xlsx?dl=0](https://www.dropbox.com/s/4fuvr6tbq4avt4l/Appendix_3_150717.xlsx?dl=0)

#	FIELD NAME	FIELD DESCRIPTION	CHAR TYPE	FIELD SIZE	SPECIFICATIONS	MANDATORY FOR NEW
1	HF_ID	National identifier that uniquely identifies a health facility	Varchar	8	The code is generated and maintained by MOHS/DOPH	X
2	HF_NAME_EN	Official or business name of the health facility in English	Varchar	150	- Follows the official name as stated in the License to Operate or the name used for legal purposes. - The health facility name is free of any abbreviations. - The health facility name uses proper capitalization and avoids the use of special characters or punctuation symbols.	X
3	HF_NAME_MM	Official or business name of the health facility in Myanmar language	Varchar	150	- Follows the official name as stated in the License to Operate or the name used for legal purposes. - The health facility name is free of any abbreviations. - The health facility name uses proper capitalization and avoids the use of special characters or punctuation symbols.	
4	HF_N_OLD_1	First former official or business name of the health facility in English	Varchar	150	- Follows the official name as stated in the License to Operate or the name used for legal purposes. - The health facility name is free of any abbreviations. - The health facility name uses proper capitalization and avoids the use of special characters or punctuation symbols. - Not applicable by default	
5	HF_N_OLD_2	Second former official or business name of the health facility in English	Varchar	150	- The format for the health facility follows the official name as stated in the License to operate or the name used for legal purposes. - The health facility name is free of any abbreviations. - The health facility name uses proper capitalization and avoids the use of special characters or punctuation symbols. - Not applicable by default	
6	HF_N_OLD_3	Third former official or business name of the health facility in English	Varchar	150	- The format for the health facility follows the official name as stated in the License to operate or the name used for legal purposes. - The health facility name is free of any abbreviations. - The health facility name uses proper capitalization and avoids the use of special characters or punctuation symbols. - Not applicable by default	
7	HF_TYPE	Type of the health facility spelt in full	Varchar	25	Based on the agreed upon classification of health facilities	X
8	HF_OWN	Classification according to major ownership spelt in full	Varchar	10	Government = Created by law Private= Owned, established, and operated with funds through donation, principal, investment or other means by any individual, corporation, association or organization.	X
9	HF_OWN_G	Sub-classification of the government facility according to ownership spelt in full	Varchar	50	MOHS, MOD, MOHA, MOL - "Not applicable" if private facility	
10	HF_OWN_P	Sub-classification of the private facility according to ownership spelt in full	Varchar	50	-Single Proprietorship   Partnership   Corporation   Corporation   Cooperative   Foundation   Religious   Non-government Organization - "Not applicable" if government facility	
11	HF_ADDRESS	Street number and name	Varchar	100	- Not applicable when the street has no name	

## Appendix 4 - Example data dictionary for the administrative divisions master list

### A. When the village boundaries are available (Philippines)

#	FIELD NAME	FIELD DESCRIPTION	CHAR TYPE	FIELD SIZE	SPECIFICATIONS
1	Reg_Code	Official code of the Region as per the PSGC	Varchar	11	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
2	Reg_Name	Official name of the Region as per the PSGC	Varchar	100	
3	Prov_Code	Official code of the Province as per the PSGC	Varchar	11	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
4	Prov_Name	Official name of the Province as per the PSGC	Varchar	100	
5	Mun_Code	Official code of the Municipality as per the PSGC	Varchar	11	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
6	Mun_Name	Official name of the Municipality as per the PSGC	Varchar	100	
7	Bgy_Code	Official code of the Barangay as per the PSGC	Varchar	11	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
8	Bgy_Name	Official name of the Barangay as per the PSGC	Varchar	100	
9	LAT	Latitude of the health facility	Num Floating	8	- Numeric value with (5 digits after the period) - Use 0 for default value.
10	LONG	Longitude of the health facility	Num Floating	9	- Numeric value with (5 digits after the period) - Use 0 for default value.
11	HF_Cat_Code	Official code of the health facility to which the village is attached to (catchment area)	Varchar	10	Follows the code from the health facility master list

### B. When the village boundaries are not available (Cambodia)

#	FIELD NAME	FIELD DESCRIPTION	CHAR TYPE	FIELD SIZE	SPECIFICATIONS
1	PRO_C_MO	Official code of the Province as per the MOI/NCDD gazetteer	Varchar	7	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
2	PRO_N_MO	Official name of the Province as per the MOI/NCDD gazetteer	Varchar	100	
3	DIS_C_MO	Official code of the District as per the MOI/NCDD gazetteer	Varchar	7	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
4	DIS_N_MO	Official name of the District as per the MOI/NCDD gazetteer	Varchar	100	
5	COM_C_MO	Official code of the Commune as per the MOI/NCDD gazetteer	Varchar	7	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
6	COM_N_MO	Official name of the Commune as per the MOI/NCDD gazetteer	Varchar	100	
7	VIL_C_MO	Official code of the Village as per the MOI/NCDD gazetteer	Varchar	7	"PH" has been added in front of the sequence to avoid the issues linked to the "0"
8	VIL_N_MO	Official name of the Village as per the MOI/NCDD gazetteer	Varchar	100	
9	Lat	Latitude of the village	Num Floating	8	- Numeric value with (5 digits after the period) - Use 0 for default value.
10	Long	Longitude of the village	Num Floating	9	- Numeric value with (5 digits after the period) - Use 0 for default value.
11	Source_Lat_Long	Source and method used to obtain the geographic coordinates of the health facility	Varchar	50	
12	Acc_Lat_long	Qualitative measure of the accuracy level for the geographic coordinate	Varchar	6	- values: low, medium, high
13	HF_CAT_CODE	Official code of the health facility to which the village is attached to (catchment area)	Varchar	10	Follows the code from the health facility master list