HIS geo-enabling: Guidance on the establishment of a common geo-registry for the simultaneous hosting, maintenance, update and sharing of lists as well as associated hierarchies and spatial data

Version 2.0

With the support of:

Revision History

Version	Release Date	Comment	By
1.0	18 August 2017	First version of the guidance	Steeve Ebener, Martin Verzilli, Jonathan Payne, Scott Teesdale
2.0	18 January 2022	Revision of the original guidance to capture the lessons learned through the development and deployment of the first IT solution aiming at serving as Common Geo-Registry	Steeve Ebener

Table of content

Acknowledgements

First and foremost, we would like to express our gratitude to the Bill & Melinda Gates Foundation for the financial support provided to the writing of the present version of the guidance.

Our thanks then go to the Clinton Health Access Initiative (CHAI), Anne Liu, Rica Duchateau and Varshana Rajasekaran in particular, for having accepted to share their experience and precious insight to the document.

Our gratitude finally goes to the Asian Development Bank (ADB), 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, Martin Verzilli (InSTEDD), Jonathan Payne (Open Concept Lab/Regenstrief Institute), and Scott Teesdale (InSTEDD) as well as to Channé Suy (iLab), Niamh Darcy (RTI), Nicholas Oliphant (Global Fund), Rocco Panciera (UNICEF), and Sam Libby (Esri) who contributed to the release of the first version of the guidance.

This guidance was prepared with the support of the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

Author

Steeve Ebener – Health GeoLab, Philippines (steeve@tropmedres.ac)

Acronyms

AU	Anonymous user
AC	API Consumer
BMGF	Bill and Melinda Gates Foundation
CGR	Common Geo-Registry
CHAI	Clinton Health Access Initiative
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
HGLC	Health GeoLab Collaborative
HIS	Health Information System
IGIF	Integrated Geographic Information Framework
NSDI	National Spatial Data Infrastructure
NDP	National Development Plan
RA	Registry Administrator
RC	Registry Contributor
RM	Registry Maintainer
RS	Remote Sensing
SA	System Administrator
SDG	Sustainable Development Goals
SOP	Standard Operating Procedure
UN-GGIM	United Nations Committee of Experts on Global Geospatial Information Management

Objectives and target audience

The primary objective of the present guidance is to provide the necessary information and tools for countries and development partners to assess if the Common Geo-Registry (CGR) as a concept and as platform addresses a need. If this is the case, the guidance can then help them in identifying if the IT solution they are considering using, or developing, to serve as a CGR is appropriate and, if not, identify gaps to be addressed.

This guidance also aims to serve as a reference for a cost-effective and sustainable deployment of a CGR and to promote a systemic and systematic integration of the geographic and time dimensions in information system's architecture.

In view of the above, this guidance targets officials from different ministries as well as implementing partners or donors who aim to strengthen geographic information management through the simultaneous hosting, management, updating, and sharing of the master lists as well as associated hierarchies and spatial data. This may be inclusive of any geographic objects core to sustainable development, though this guidance particularly focuses on public health.

How to use this guidance

The guidance is organized in such a way that readers can easily access the information that would be the most useful for them depending on where they currently stand with their understanding of the CGR concept and/or the deployment of the IT solution to serve as a CGR.

The document has been divided into the following sections:

- Introduction: Describes the CGR concept and its origin, provides use cases as an illustration of the business needs it aims at addressing, presents the benefits brought by a CGR, discusses the challenges and opportunities of deployment
- CGR content: Describes the type of data that a CGR should be able to host, manage, regularly update, and share to fulfil its objectives
- CGR Requirements: Describes data flow, organizations, user roles and rights as well as minimum set of functionalities for an IT solution to function as a CGR
- CGR supporting environment: Describes key elements that should be in place to ensure the long-term sustainability and cost-effectiveness of a CGR

The technical terms used across this guidance together with their respective definitions are presented in Annex 1.

1. Introduction

Geography and time are intrinsically linked to any activity or event taking place on our planet - everything happens at a given place and time. As such, these two dimensions are common to all the sectors and therefore foundational to reaching a more systemic, and systematic, approach to solving developmental problems in general and public health ones in particular.

Geography and time are present across the three main functions of public health: (1) assessing and monitoring the health of communities and populations at risk; (2) assuring that all populations have access to quality, timely, and cost-effective care; and (3) formulating public health policies designed to solve identified health problems and priorities.

From a geographic perspective:

- 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.
- Physical accessibility to health services is greatly influenced by their geographic location, the spatial distribution of the population, and the environment surrounding them.
- Maps represent a powerful medium 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 we live in, and therefore the changes this caused in our geography, also require time to be considered across these three functions. Some examples:

- New health facilities are being built while others are being closed.
- Administrative units are being created or modified.
- People are moving around, finding themselves in different places at different points in time

Despite the foundational importance of the above, geography and time are generally among the most poorly captured and, therefore, analyzed dimensions in the country's Health Information System (HIS). Often this is due to limitations of capturing and regularly updating the above mentioned dynamic and relationships within available information system platforms.

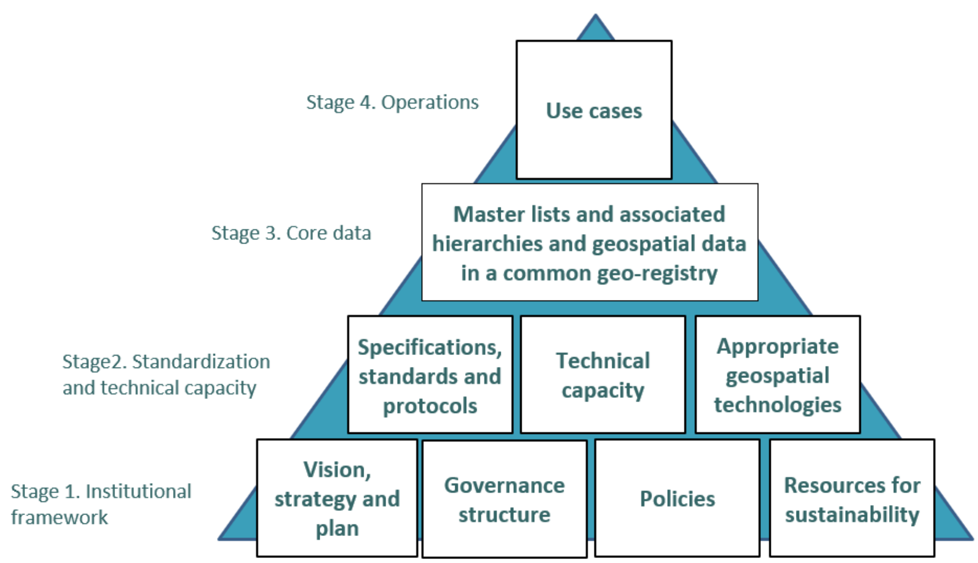
The Health GeoLab Collaborative has developed and has been implementing a framework – the Health Information System (HIS) geo-enabling framework (Figure 1) – to address these challenges and to support the health sector to fully benefit from the power of geospatial data and technologies. The HIS geo-enabling framework is composed of nine (9) elements that need to be in place for a Health Information System (HIS) to be considered as geo-enabled, namely:

- A clear **vision, strategy(ies), and action plan** for the management and use of geospatial data and technologies have been defined.

- A **governance structure** supporting the vision, strategy(ies), and action plan has been established.
- Enough **technical capacity** has been developed.
- Geospatial data **specifications, standards, and protocols** have been defined and are being implemented to ensure the availability and quality (completeness, uniqueness, timeliness, validity, accuracy, and consistency) of geographic information across the whole data lifecycle.
- The **master lists** [1] for the core geographic objects (health facilities, administrative units and villages and reporting divisions) and their associated hierarchies and geospatial data have been developed, made accessible, and an updating mechanism put in place for each of them through the use of a **common geo-registry**.
- The appropriate **geospatial technologies** have been identified and are being used in accordance with good geospatial **data management practices.**
- **Use cases** (applications) supporting health programs (communicable diseases surveillance, malaria elimination, health service coverage, disaster management, etc.) towards reaching Sustainable Development Goals (SDG) 3 are being implemented and documented
- **Policies** supporting and enforcing all of the above as well as geospatial data accessibility have been released.
- The necessary **resources** to ensure long term sustainability have been identified and secured.

The concept of a Common Geo-Registry therefore plays a pivotal role in this framework by supporting the operationalization of use cases through the provision of the necessary content allowing to contextualize any piece of information or statistics geographically and temporarily [1].

Figure 1 – The HIS geo-enabling framework [1]



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

The first version of this guidance[2] resulted from the consultation conducted during the Common Geo-registries for Health Information Exchanges Design Workshop that took place in Phnom Penh, Cambodia from June 12-13, 2017[3].

Since then, lessons learned during the development and deployment of the first IT solution aiming at operationalizing the CGR concept confirmed the relevance of the concept beyond the health sector and demonstrated that the technology was already at a stage where the functionalities in a CGR solution could be expanded to benefit of a larger group of users.

At the same time, the CGR concept has started to make its way to becoming an important component of the data infrastructure to support different health programs, particularly routine immunization [2, 3].

The above, together with the need to further clarify some of the concepts included in the first version of the guidance, and to provide more details on specific topics led to the development of the present document.

1.1 The CGR concept

Part of the challenge in managing geography through time can be addressed through the development, maintenance, regular update, sharing, and use of master lists as well as associated hierarchies and geospatial data. However, to do so requires an appropriate technology solution that facilitates this in a cost-effective and synchronized way.

While some IT solutions provide such capabilities for a single geographic feature (e.g., health facilities), there was a need for an IT solution able to manage and connect multiple geographic features the relationships that exists between them, as well as their respective geographic representation in parallel and through time, while respecting the curation mandate of each organization. The Common Geo-Registry (CGR) concept was born to fill this gap. The CGR allows to provide a container accessible by any information system to obtain reliable information on geography through time. Subsequently, information system users can then improve geographically based decision making and support a more systemic approach to solving developmental problems in general and public health ones in particular.

A CGR is not meant to serve as warehouse hosting program specific information or statistics (e.g., population, disease prevalence, etc.) but as an enabler that will allow any information system to:

1. Contextualize data from different sources in both space and time
2. Use geographic objects as the common link between these sources
3. Aggregate data according to different hierarchies
4. Support the creation of maps based on the same geography
5. Facilitate spatio-temporal analyses including but not limited to trend analysis

[2] HGLC (2017): 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: https://healthgeolab.net/DOCUMENTS/Guidance_Common_Geo-registry_Ve1.pdf

[3] https://www.healthgeolab.net/MEETINGS/CGR_2017/Geo-Registriy_Workshop_Final_Exec_Sum_010717.pdf

While this guidance builds on the implementation of the CGR concept within the health sector, it can be applied to any sector and should be seen as one of the innovative geospatial technologies to support the availability, quality, and accessibility of data and information about the evolution of geography through time. As such, the CGR as a concept and as an IT solution contributes to the National Spatial Data Infrastructure (NSDI) under the guidance of the Integrated Geographic Information Framework (IGIF)[4] promoted by the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM).

1.2 Public health use cases

The following public health use cases illustrate the critical need for the geographic and time dimensions to be better integrated within program-specific information systems due to the importance that geography plays in each of them:

1. **Communicable disease prevention, control, and elimination** - Programs seeking to prevent and reduce morbidity and mortality of diseases such as tuberculosis, HIV, or malaria rely heavily on geography. For example:
    - ***Malaria elimination activities*** - In reactive case detection, when a positive malaria case is diagnosed in a health facility, there is a need for the case investigation team to not only be able to contact and visit the facility in question but also the place of residence and place of work of the patient (full address, geographic coordinates) for further investigation (e.g., foci investigation), intervention, and monitoring activities through time.

2. **Emergency and outbreak management** - Geography provides the common operational picture across all phases of the emergency cycle (mitigation, preparedness, response, and recovery) or during a disease outbreak. For example:
    - ***Typhoon preparedness and response -*** As soon as a typhoon is confirmed to make landfall, the master lists and associated spatial data allow to complete 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 to warn them; and (3) supporting the Rapid Impact Assessment (RIA) and therefore the response phases right after landfall, 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 since John Snow's famous work during the 1854 cholera outbreak in London[5], continues to be a critical tool in rapidly identifying and combating disease outbreaks. Up-to-date master lists and related maps are needed more than ever to be able to contain large scale events such as the Ebola outbreak in Western Africa between 2013 and 2016 or the COVID-19 pandemic since December 2019.

---

[4] https://ggim.un.org/IGIF/

[5] https://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak

3. **Health planning -** Critical not only for developing health plans and programs that answer the population's needs but also for budgeting and allocating resources in a systematic and equitable way, effective health planning requires knowing 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 overview of the spatial distribution of health services, including community health workers [8], and of the population distribution down to the human settlement level in order to make sure that the whole population can access health services within an acceptable travel time.
   - ***Macro- and microplanning:*** The lists and associated geospatial data needed to assess health coverage is also key to the development of macro- and microplans that will not only ensure an equitable access to the concerned fixed or outreach health services (e.g., vaccination points) but also the optimization of session plans for outreach activities based on physical accessibility to human settlements and the organization of the territory in contiguous health areas.
   - ***Health service delivery*** **-** Once a plan has been developed, the appropriate resources (medical staff including community health workers, equipment, financial) need to be attached to each point of service (health facility, vaccination post…) in order to ensure a comprehensive and cost-effective delivery of health services.

The above use cases illustrate the importance of geography within the health sector and the need for master lists, associated hierarchies and geospatial data for the concerned geographic objects (health facilities, health areas, human settlements, health workforce….) to be established, maintained, updated, shared, and used as reference across the HIS.

As the CGR concept aims at covering sustainable development in general, the above-mentioned use cases should be extended to cover the entire public sector (education, environment, utilities, etc.). Doing so would increase the list of geographic features to be hosted in a CGR.

1.3 Benefits of a CGR

Benefits to deploying a CGR include:

- Allowing a more comprehensive integration of the geographic and time dimensions in any information system across the whole data lifecycle, therefore allowing a more systematic and systemic approach to solving development issues.
- Supporting the successful implementation of programs that are heavily dependent on geography (e.g., communicable disease control and elimination, planning, emergency management, planning, immunization…).
- Improving the ability of the governmental entities in charge of curating this information to simultaneously manage and share lists, hierarchies, and geospatial data as foundational elements of any information system.
- Facilitating the capture of the metadata associated with each piece of information stored in a master list, as well as the tracking of changes being implemented over time for better accountability.

- Improving data quality across its six dimensions (completeness, validity, timeliness, uniqueness, accuracy, consistency) [4] by supporting good data management practices.
- Reducing the duplication of efforts and therefore cost by maintaining only one authoritative list together with the associated hierarchies and geospatial data for each geographic feature, instead of several non-authoritative sources.
- Supporting collaboration and coordination across partners as well as the ability to further promote innovation and data utilization through the use of interoperable services using geography as the common integrating dimension.
- Enabling the following functionalities in the information systems connected to the CGR:
    - Contextualizing data from different sources in both time and space
    - Using geographic objects as the common link between data sources
    - Facilitating trend analysis
    - Aggregating data according to different hierarchies
    - Supporting a more effective use of the visualization and analytical power offered by Geographic Information Systems (GISs) and other geo-enabled technologies

All the above contributes directly to the development, integration, strengthening, and maximization of geospatial information and, as such, improves geographically based decision making and a more systemic approach to solving nowadays developmental challenges. In view of this, a CGR should be seen as one of the pillars supporting the NSDI.

2. CGR content

As per its definition, a CGR is meant to host, manage, regularly update **lists** as well as associated **hierarchies** and **geospatial data** for the **geographic objects** core to development in general and public health in particular.

The following sections describe in more detail these different elements, as well as those associated with them (geographic features, geographic object types, data elements, classification tables), before providing recommendations to consider when preparing such content.

2.1 Geographic features and objects

Geographic features are features of the Earth, natural or artificial.

Each sector has its own set of geographic features (fixed and/or mobile) that are core to the implementation of its programs and activities. In public health, for example, these geographic features would include but not be limited to health facilities, vaccination points, administrative or health divisions, and human settlements.

Identifying these geographic features and having a clear definition for each of them is an important step towards the deployment of an effective CGR as this will directly influence the content to be uploaded, hosted, managed in, and shared from the CGR.

A CGR is only meant to host, manage, regularly update, and share information associated with geographic features for which it makes sense to establish a master list (e.g., health facilities, administrative units, villages, vaccination post, vehicle, people). A master list for other geographic features would either generally not be necessary (e.g., roads, rivers) or not applicable (e.g., continuous features like topography or population distribution) [5].

A geographic object is the computer representation of a geographic feature[6]. As such geographic objects can be categorized based on how they would be represented on a map [5]:

- Fixed geographic objects that can be simplified by a point (examples: household, health facility, human settlement).
- Fixed geographic objects that should be represented by polygons due to their much larger extent (examples: administrative units, health districts...) or by a line due to their shape (example: road, river...).
- Mobile geographic objects (examples: individuals, patients, vehicles...): the geography of these objects would either be obtained by considering them attached to a fixed geographic object or by simplifying them as a point that would be located through their geographic coordinates (latitude and longitude) taken at a given point in time.

The identification of the types of geographic objects and of their respective mode of representation is an important exercise to be conducted as part of the CGR content preparation process (Section 7.8).

In view of the above, a CGR should be able to perform the following for geographic objects:

- Handle as many types of geographic objects as needed to cover the defined conceptual data model
- Attach each type of geographic object to a specific organization
- Manage geographic object types (add, edit, delete)
- Maintain metadata for each geographic object type
- Differentiate between the geographic object types managed by the organization having the official curation mandate over them (master) and those managed by other organizations (non-master) (see Section 2.7.3)
- Specify access rights for each geographic object type (public, restricted, or private)
- Define the geometry for each geographic object type (point, line, polygon)
- Handle changes over time down to a specific date (temporal validity of a given geographic object type)
- Group geographic object types for the purpose of generating lists at the group level (e.g., health facilities grouping hospitals, health centers, and health posts).

[6] https://www.igi-global.com/dictionary/nature-of-geographic-knowledge-bases/57190

2.2 Data elements and classification tables

A CGR focuses on all the information that allows to contextualize any piece information in both space and time. As such, a CGR only handles the data elements that allow to uniquely identify, classify, locate, and when applicable, contact each geographic object it contains (the set of data elements referred to as the signature domain[7]). Such a list will be specific to each type of geographic object hosted in the CGR.

Other types of data elements, including programmatic ones (e.g., statistics), are not meant to be managed in a CGR but in program-, or even sector-, specific information systems. Some of the reasons behind this are the sensitivities often linked to programmatic data, the difference in data flow between this type of data and the geographic objects they are attached to as well as difference in the management of the time dimension (e.g., the temporal validity of a statistic might be different from the temporal validity of the geographic object it is attached to). In addition, managing these data elements in these other systems allows for programs to benefit from functionalities that are not available within a registry, including, but not limited to, data collection, statistical analysis, or visualization (graphs, thematic maps).

The identification of the data elements to be associated with each geographic object and stored in a CGR should take place as part of a process involving all the concerned stakeholders. The result is a data dictionary containing at least the following information for each of the selected data elements (an example for a health facility master list is presented in Annex 2):

- The data element code as implemented in the CGR
- A short description of the data element
- The data element type
- The size of the data element (number of storage units)
- Whether the data element is considered as mandatory when adding a new record in the master list

Examples, notes, and, when applicable, the source (reference) of each data element can be added to the above information if it can help those in charge of managing the CGR content and its users to have a clear understanding of each of these data elements.

In addition to the above, a classification table needs to be defined for each of the data elements where the values are limited to a few options to ensure consistency across records (example: type, ownership, etc.). Such tables should at least contain the following for each option: a unique code, a label (English and local language) and/or acronym, a description (English and local language), and, when applicable, the source of the definition. Table 1 provides an example of classification table for health facility types in Cambodia following this structure.

Developing the above-mentioned data dictionary and associated classification tables for each geographic object is another important exercise to be conducted as part of the CGR content preparation process (Section 4.8). Ideally, this exercise would be conducted after the definition

[7] https://www.measureevaluation.org/resources/publications/wp-07-91

and documentation of the relationships existing between the different types of geographic objects (Section 2.2) as this exercise might identify additional data elements to be captured in the data dictionary to be able to rebuild these relationships.

Please refer to Section 3.3 for more details regarding the functionalities that are required to support the above.

Table 1 – Example of classification table (health facility types in Cambodia)

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

2.3 Hierarchies and conceptual data model

Knowing how geographic objects relate to each other over time as part of different hierarchies is important across use cases, including those mentioned in Section 1.2, not only to be able to aggregate information to support decision making when applicable but also to ensure relational consistency between geographic objects.

While the primary role of the CGR is to capture the geographic relationships that exist between geographic objects (is within, lives in), other types of relationships such as administrative (is reporting to), health-related (covers, provides services to, refers to) or associative (is part of) ones should also be considered as they also change through time.

While you might want to start by capturing these relationships in distinct hierarchies (Figure 2), it is always good to then combine them into a conceptual data model like the one in Figure 3. This may then allow the identification of additional relationships not being captured in each separate hierarchy.

Developing these hierarches and associated conceptual data model is another important exercise to be conducted as part of the CGR content preparation process (Section 4.8). Ideally, this exercise would be conducted after the identification of the geographic features (Section 2.1.1).

A CGR should then be able to capture, manage through time, and export the hierarchies that have been defined together with metadata explaining what these are about (Section 3.3).

Figure 2 – Example of hierarchies

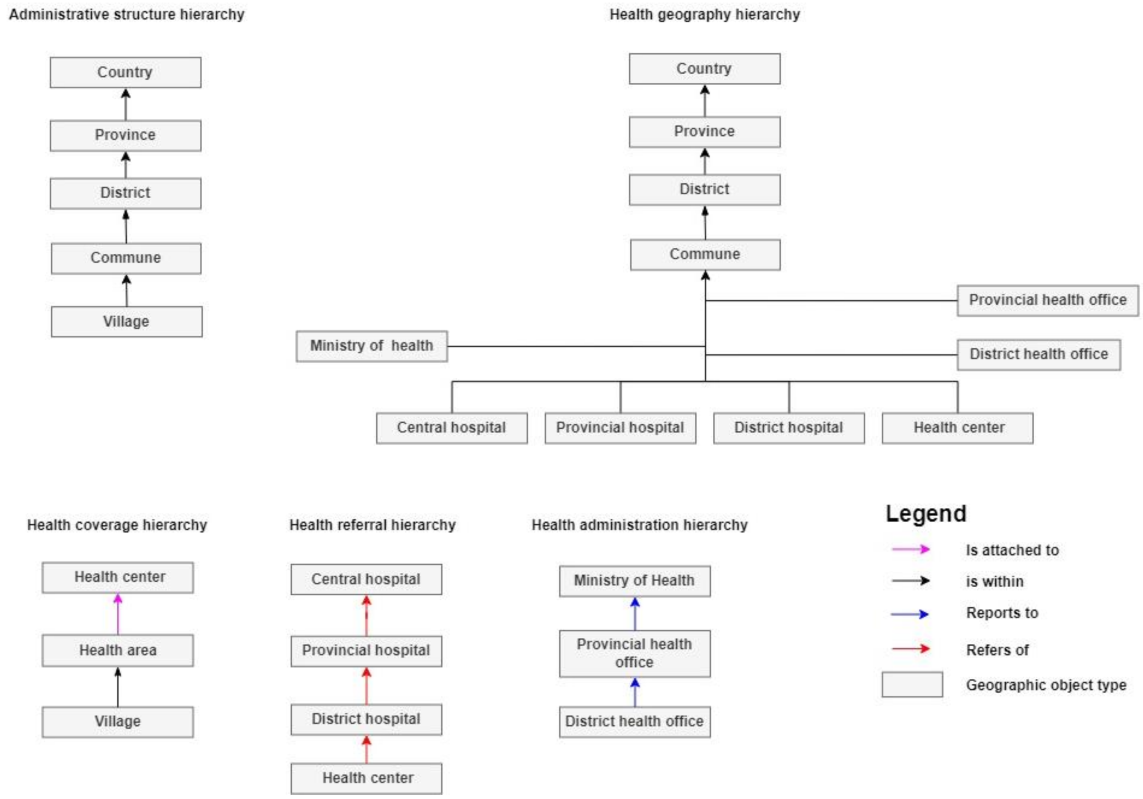
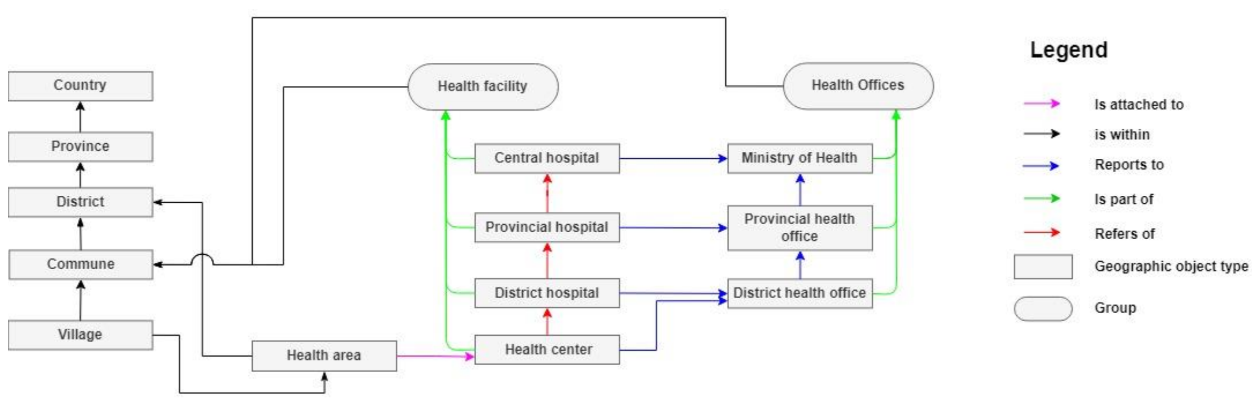


Figure 3 – Example of a conceptual data model resulting from the grouping of the hierarchies reported in Figure 2



2.4 Lists

A list is a tabular representation of the data elements associated with all the active, and past active, geographic objects (records) of a given type at a given point in time. Table 2 contains an extract of the master list of health facilities based on the data dictionary in Annex 2.

Table 2 – Extract of the list corresponding to the data dictionary reported in Annex 2

HF_ID	HF_N_EN	HF_N_MM	OP_DATE	HF_T_EN	HF_OWN
HF000062	Ywar Thar SRHC	ရွာသာ	2001-07-27	Sub Rural Health Center	Public (MOH)
HF002135	Zee Pin Hla SRHC	ဇီးပင်လှ	1998-03-27	Sub Rural Health Center	Public (MOH)
HF004754	Myo Thit SRHC	မြို့သစ်	1972-04-29	Sub Rural Health Center	Public (MOH)
HF003336	Ya Thit SRHC	ရသစ်	1983-07-01	Sub Rural Health Center	Public (MOH)
HF003302	Let Wea SRHC	လက်ဝဲ	2005-06-17	Sub Rural Health Center	Public (MOH)
HF000008	Let Pan Taw RHC	လက်ပံတော	1934-06-15	Rural Health Center	Public (MOH)
HF002212	Mei Ni Ma Kone SRHC	မယ်နီမကုန်း	1969-06-06	Sub Rural Health Center	Public (MOH)
HF004708	Zee Kyun SRHC	ဇီးကျွန်း	2009-08-31	Sub Rural Health Center	Public (MOH)
HF000469	Myay Ni Kone SRHC	မြေနီကုန်း	1999-06-15	Sub Rural Health Center	Public (MOH)
HF004090	Let Pan Kyun SRHC	လက်ပန်ကျွန်း	2002-07-12	Sub Rural Health Center	Public (MOH)
HF002204	Let Pan Taw SRHC	လက်ပံတော	2009-06-17	Sub Rural Health Center	Public (MOH)
HF002021	Kan Su SRHC	ကန်စု	1943-09-12	Sub Rural Health Center	Public (MOH)
HF002790	Na Nwin Taw Boe RHC	နနွင်းတော်ဘိုး	1985-07-29	Rural Health Center	Public (MOH)
HF002194	Myo Taung SRHC	မြို့တောင်	2009-07-29	Sub Rural Health Center	Public (MOH)
HF002563	Let Pan Pyar Station Hospital	လက်ပံပြားတိုက်နယ်ဆေးရုံ	2009-06-17	Station Hospital	Public (MOH)
HF002072	Let Pan Pu SRHC	လက်ပံပု	2005-12-15	Sub Rural Health Center	Public (MOH)

Lists represent the most familiar way for users to look at the information linked to any given geographic object. When they are authoritative and curated by the mandated governmental entity, these lists, referred as master lists in this case, also serve as a ground reference to assess the completeness, uniqueness, timeliness, validity, and consistency of the geospatial data also hosted in a CGR [5], as well as a denominator for the implementation of any public health program.

Given this, lists are central to the CGR concept and have been the topic of different documents aiming at supporting the establishment of a master list for given geographic objects core to public health [7, 8].

To this end, a CGR should be able to import, host, manage, regularly update, and share content under the form of a list together with its associated metadata and data dictionary (see Section 3.3 for more details on the requirements associated to these functionalities).

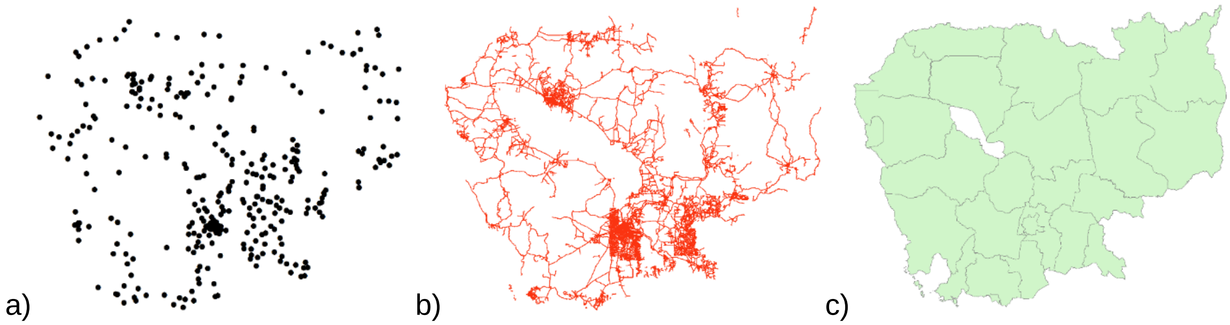
2.5 Geospatial data

Also referred to as spatial data, geospatial data corresponds to information about the locations and shapes of geographic features and the relationships between them, usually stored as coordinates and topology (e.g., geographic location of health facilities, boundaries of administrative units…).

As per Section 2.1.1, a CGR is meant to only handle types of geographic objects for which it is important to host, manage, regularly update, and share a list. This corresponds to geographic objects that are stored in vector format GIS layers. Raster format GIS layers are therefore not meant to be managed in a CGR.

In the vector format, geographic features are represented by points, lines, or polygons (Figure 4). As such, a CGR should be able to handle these three modes of representation.

Figure 4 – Modes of representation in the vector format (points (a), lines (b), and polygons (c))



A CGR should be able to import, host, manage, regularly update, and share content under the form of geospatial data together with its associated metadata and data dictionary (see Section 3.3 for more details on the requirements associated with these functionalities).

2.6 Changes over time

Geography evolves over time. A core function of a CGR is therefore the capability to not only capture the changes that are occurring over time for any information it hosts (creation of a new district, closing of a health facility, change of phone number for a community health worker) but also to conserve the information as it was before the change and, when applicable, the information necessary to rebuild the change (e.g., the unique identifier and name of the district from which a new district has been carved). This information is for example critical to perform trend analysis, especially in countries where the administrative structure changes frequently (e.g. to reconstruct disease prevalence trends across district that have changed).

This capability should allow to retrieve the list, the hierarchies as well as the geographic location or extent for any geographic object at any point in time over the entire period for which data has been uploaded in the CGR.

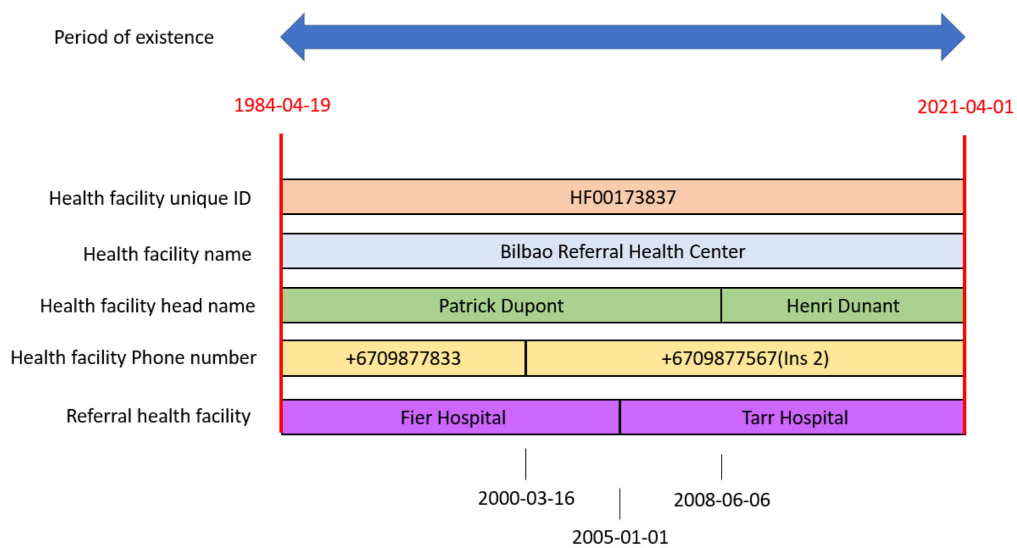
There are three main types of changes to be captured in a CGR for any given geographic object (Table 3).

Table 3 – Type of changes to be captured in a CGR

Type	Description	Example
Start of existence	The start of existence of the geographic object	Opening a new health facility, creation of a new district
Value change	A change in value for a given data element	Change of health facility name, phone number, or status or the geographic location/extent of the geographic object
End of existence	The end of existence of the geographic object	Closure of a health facility, end of existence of a province being merged with another one

Figure 5 provides an illustration of these main types of changes for a health facility that opened on April 19, 1984, and closed definitively on April 1, 2021. Over that period, the health facility's name, head, and phone number each changed once at different points in time.

Figure 5 – Example of changes occurring over time for some data elements associated with a health facility



In this case, the CGR will need to have the capability to:

- Set the period of existence of the geographic object (1984-04-19 to 2021-04-01)
- Capture and retain the value for each instance of the data elements that have changed through time together with the period over which each value is valid (e.g., Patrick Dupont over the 1984-04-19 to 2008-06-06 period and then Henri Dunant over the 2008-06-06 to 2021-04-01 period).

Capturing the above information is sufficient to understand changes occurring to any independently created geographic objects, i.e., geographic objects that are not being created out of already existing ones (e.g., infrastructures such as health facilities or schools).

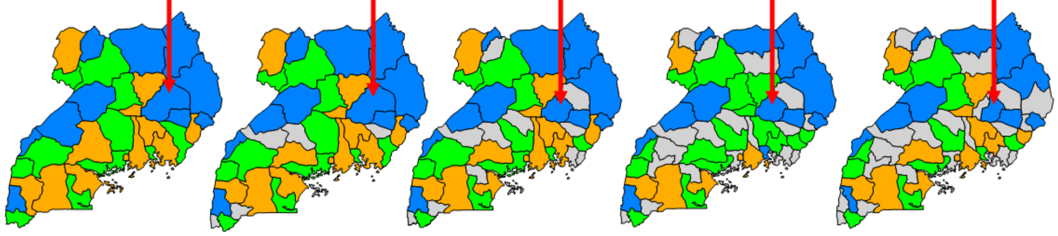
The situation is different when looking at the changes occurring through time for geographic objects that are being carved out from existing one to start their existence. For example, this occurs when boundaries change (e.g., split or merge) across administrative, health, statistical, or electoral units covering the entire territory of a given country.

In this case, capturing the period over which each unit has been in existence is not sufficient to rebuild these changes. To illustrate this, Figure 6 captures the changes that occurred through time for part of the districts of Uganda over the January 1, 1990, to June 30, 2005, period.

While the district of Soroti, for example, has been in existence over the complete period reported in Figure 6, its geographic extent has significantly changed twice over that same period. In this case, a CGR should be able to capture the period of existence of the district of Soroti as well as the two changes in question including the indication of the instance (single occurrence) of the district of Soroti from which the district of Katakwi in March 1997 and then the district of Kaberamaido in July 2007 have been carved.

Figure 6 – Example of changes occurring over time for a subset of the districts of Uganda[8]

1990-01-01 to 1991-03-14	1991-03-15 to 1997-03-20	1997-03-21 to 2000-11-27	2000-11-28 to 2001-06-30	2001-07-01 to 2005-06-30
Soroti	Soroti	Katakwi	Katakwi	Katakwi
		Soroti	Soroti	Soroti
				Kaberamaido
Mubende	Kiboga	Kiboga	Kiboga	Kiboga
	Mubende	Mubende	Mubende	Mubende
Kitgum	Kitgum	Kitgum	Kitgum	Kitgum
			Pader	Pader



One way to capture this information is under the form of the table included in Figure 6. This approach is, for example, followed by the Second Administrative Level Boundaries program (SALB) to capture the historical changes of administrative units down to the second subnational level for all the United Nations Member States.[9]

Another possible way is capture this is to record the following information for each of the changes in a separate table:

- The date of the change
- The type of change (split, merge, transfer, upgrade, downgrade)
- The type of geographic object before and after the change (same type in the case of a split, merge, or transfer or different type in the case of an upgrade or downgrade)
- The unique identifier and name of the geographic object before and after the change
- Notes allowing to understand and/or track the change (e.g., the legal document in which the change is recorded)

This type of approach is, for example, used by the Philippines Statistical Agency (PSA) to capture the changes that have been occurring through time for the administrative units of the Philippines[10].

Table 4 provides an example of such a separate table capturing the changes illustrated in Figure 6.

[8] https://www.unsalb.org/
[9] https://www.unsalb.org/sites/default/files/wysiwyg_uploads/docs_uploads/SALB_DataSpecifications.pdf
[10] https://psa.gov.ph/classification/psgc/

Table 4 - Example of table capturing the changes reported in Figure 6

Event Date	Event Type	Type of geographic object before the change	Unique identifer of the geographic object before the change	Name of the geographic object before the change	Type of geographic object after the change	Unique identifer of the geographic object after the change	Name of the geographic object after the change	Notes
1991-03-15	Split	District	UGA992	Mubende	District	UGA021	Kiboga	Decree 1991/03
1991-03-15	Split	District	UGA992	Mubende	District	UGA035	Mubende	Decree 1991/03
1997-03-21	Split	District	UGA998	Soroti	District	UGA019	Katakwi	Decree 1997/03
1997-03-21	Split	District	UGA998	Soroti	District	UGA044	Soroti	Decree 1997/03
2000-11-28	Split	District	UGA023	Kitgum	District	UGA051	Kitgum	Decree 2000/11
2000-11-28	Split	District	UGA023	Kitgum	District	UGA057	Pader	Decree 2000/11
2001-07-01	Split	District	UGA044	Soroti	District	UGA066	Soroti	Decree 2001/7
2001-07-01	Split	District	UGA044	Soroti	District	UGA061	Kaberamaido	Decree 2001/7

A CGR should then be able to provide the necessary functionalities to not only capture the changes mentioned here above but also a way to render them in a way that the user can use this information to support its work (Section 3.3).

2.7 Content-related considerations

This section provides a series of important considerations regarding the content to be hosted in a CGR.

2.7.1 Quality

The content hosted in a CGR is meant to serve as the source of truth (ground reference) for any information system to properly contextualize, visualize, and analyze business data in both space and time. As such, this content is meant to present the highest level of quality possible across the 6 dimensions of data quality (uniqueness, completeness, timeliness, accuracy, validity, and consistency) [6].

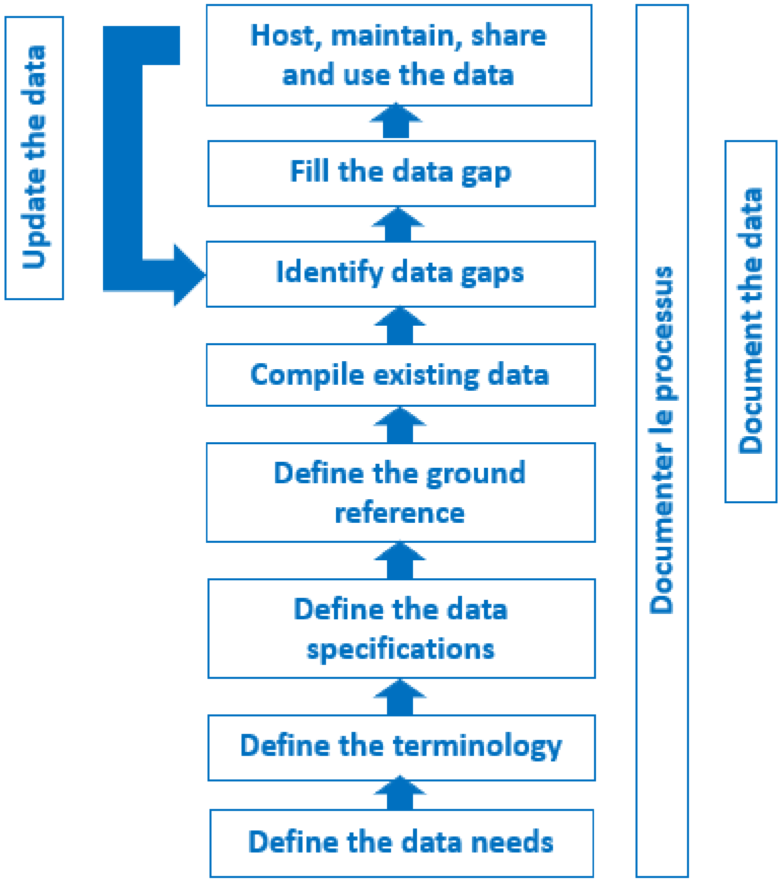
While a CGR is meant to help improve and maintain the quality of this content across these dimensions, the higher the quality of this content at the time of uploading it in the CGR, the lesser the work afterwards and the faster positive impacts can be seen.

The quality of the content before its upload in the CGR can be ensured through the implementation of good data management practices throughout the overall geospatial data management cycle in Figure 7 (extracted from [9]). The implementation of this cycle, and associated good practices, should ideally take place as part of the activities of the NSDI.

If an NSDI is not in place, the reference material generated by the Health GeoLab Collaborative[11] can be used as a reference to ensure the quality of the data and information to be uploaded in a CGR.

[11] https://healthgeolab.net/resources/reference-materials/

Figure 7 – Geospatial data management cycle



2.7.2 Unique identifier

Critical to data interoperability is the ability to uniquely, and therefore unequivocally, identify each of the geographic objects in a CGR and ensuring that this unique identifier is used across information systems. Therefore, it may be beneficial to use geography as a neutral platform for the integration of information and data coming from different sources.

The choice of an appropriate coding scheme to be associated with each type of geographic object is therefore critical as modifying the scheme once implemented in different systems can not only be costly but also have a significant impact on data compatibility as well as on the trust in the CGR content.

Geographic object types to be hosted in a CGR need to be separated into two categories when it comes to deciding on a proper coding scheme for each of them:

- Independent geographic object types, meaning geographic objects that are not being created out of already existing ones and therefore not covering the whole territory. These can be fixed (e.g., infrastructures such as health facilities or schools) or mobile (e.g., person, vehicle).

- Contiguous geographic object types (areas) covering the whole territory and that are being created out of already existing geographic objects. These types of geographic objects are parts of a geographically based hierarchy, meaning that the geographic extent of the objects at any given level in the hierarchy is obtained by grouping geographic objects from the lower level (e.g., administrative, statistical, electoral, or health units).

The following rules should be applied when deciding on the coding scheme for geographic objects part of the first category (modified from [8]):

1. The scheme should be meaningless and therefore not embed any information that could change through time (e.g., unique identifier of the administrative unit in which it is located, type, etc.).
2. The unique identifier attached to each geographic object should remain the same from its first day of existence (e.g., opening date of a health facility, birth of an individual) until its last day of existence (e.g., definitive closing of a health facility, death of an individual).
3. The unique identifier attributed to a geographic object that does not exist anymore is not attributed to any other geographic objects.
4. The structure of the unique identifier should be as short as possible but considering the number of new geographic objects that could be created in the future (function of the current number of geographic objects in the country).
5. Do not start the sequence with zeros ("0") as they might be automatically removed depending on the software being used. If the purpose is to use a coding scheme that includes the same number of digits/characters across all the records then include a set of characters not meant to change at the beginning of the sequence (e.g., HF000001; CHW000001).
6. Generate unique identifiers sequentially or randomly depending on the process used to generate them keeping in mind that the objective is to ensure that no duplicates are being generated on the long term.

The reason why some coding schemes tend not to respect the first rule for this category of geographic objects and find themselves in conflict with the second rule as soon as any information embedded in the unique identifier changes, is the intent for the unique identifier to provide a set of information that could help the user to rapidly contextualize the geographic object, for example by indicating in which district it is located.

If having access to such a contextualizing summary is important to the users, an easy way to do this, while respecting the first and second rules, is by generating a "tag" that would concatenate the information in question on the fly using the content of the list and store it in a separate data element. Such a tag could, for example, concatenate the unique identifier of the district in which a health facility is located (BGD001002), the type of health facility (HC for health center), and the unique identifier of the health facility in question (HF00232) to obtain a tag that could look like this: BGD001002-HC-HF00232. It is very important to remember that these kinds of tags cannot be used as unique identifier – they are just a complement to it.

The following two kinds of schemes can be considered when it comes to the second category of geographic objects (administrative, statistical, electoral, or health units):

1. The same set of rules as with the first category of geographic object are applied.
2. Rules 1 and 2 are not applied, meaning that the scheme includes information that can change through time resulting in unique identifiers that are also evolving through time.

The first kind of coding scheme is generally used when the concerned geographic objects are part of a complex hierarchy (high number of levels, different types of units at any given level, relationships between certain units overpassing a given level…). This is, for example, the case with the administrative hierarchy of Myanmar (Figure 8) and the reason behind the structure of the National Coding System (NCS) that has been decided upon by the government to uniquely identify each administrative unit in the country. The structure of the NCS is as follows: thematic prefix corresponding to the department in charge of the coding scheme, the General Administration Department (GA) + code indicating the administrative level (example: 10, 11, 20, 30…) + independent code with 6 digits (example of resulting unique identifier for a township: GA40000001).

The second kind is applicable when the geographic objects in question are part of a simple hierarchy (limited number of levels, single type of units at each level, no by-passing). In this case, the information that is included in the coding scheme structure of the geographic objects at a given level in the hierarchy is the unique identifier of the geographic object within which each of them is located at the upper level. This is, for example, the case with the statistical units for the Philippines. The Philippine Statistics Authority (PSA) uses the Philippine Standard Geographic Code (PSGC)[12] to uniquely identify each of these units across the hierarchy reported in Figure 9. The structure of the PSGC is as follows: Region code (2 digits); Province code (Region code + 2 digits); City/Municipality code (Region code + Province code + 2 digits); Barangay code (Region code + Province code + City/Municipality code + 3 digits). Here is an example of the PSGC for a Barangay: 103515012.

Managing the geographic objects of the second category through time in a CGR requires more information to be captured to be able to rebuild any changes as described in Section 2.6.

[12] https://psa.gov.ph/classification/psgc/

Figure 8 – Administrative hierarchy of Myanmar

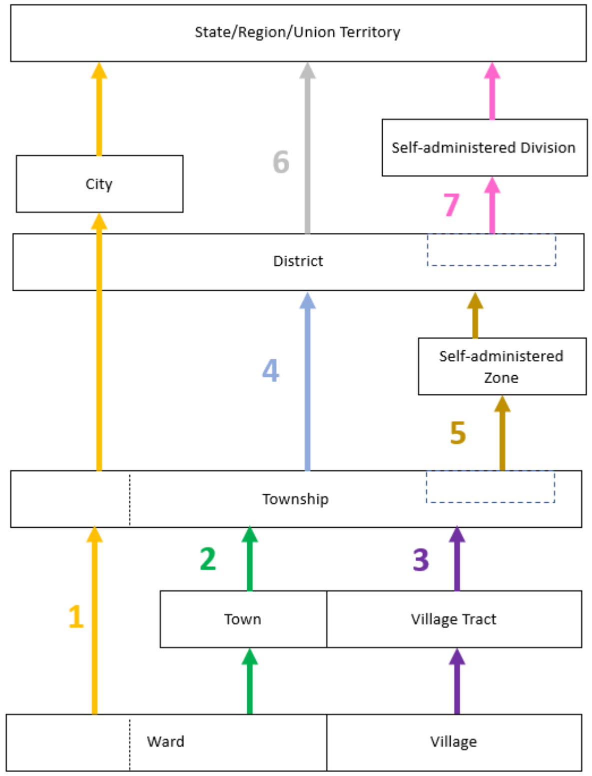
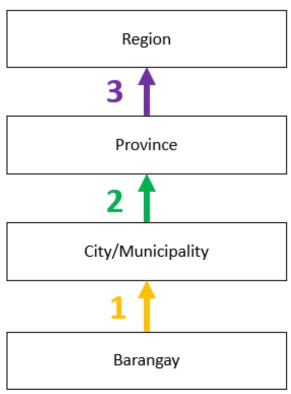


Figure 9 – Statistical hierarchy of the Philippines



It is also important to emphasize that the second kind of coding scheme can be implemented by following two different approaches when changes occur through time:

1. **The "postal" approach:** In this case, the practice is that the geographic object that keeps the same name after the change also keeps its unique identifier. An example of such practice in the case of a series of splits is reported in Figure 10. This approach is very often used by governments as it simplifies administration (budget allocation for example). The major disadvantage from a data management point of view, is that it is difficult to

identify to which instance of the geographic object the unique identifier refers to without having a time stamp associated with it.

Figure 10 – 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

2. **The "data management friendly" approach**: In this case, a new unique identifier is attributed to each new instance of geographic object as long as this is the result of a significant change in its geographic extent (example in Figure 11). The main advantage here is that the unique identifier indirectly carries the time stamp (period of validity), making it easier to identify to which instance of a particular geographic object it refers. This scheme has generally been used in projects that have a strong data management focus like the Second Administrative Level Boundaries (SALB) program which documents how changes in the unique identifier are being implemented through time in its data specification document[13].

Figure 11 – Example of implementation of the "data management friendly" approach

Period 1		Period 2		Period 3	
Name	**Codes**	**Name**	**Codes**	**Name**	**Codes**
Blue	121008	Blue	121009	Blue	121011
				Yellow	121012
		Red	121010	Red	121010

Decisions on which coding scheme to use for each type of geographic object should ideally take place as part of the NSDI activities to ensure that the above listed practices are followed and to ensure consistency between coding schemes across sectors.

A CGR is meant to not only provide the ability to manage unique identifiers for each of the geographic object types it hosts but also to support the above-mentioned practices.

[13] https://www.unsalb.org/sites/default/files/wysiwyg_uploads/docs_uploads/SALB_DataSpecifications.pdf

2.7.3 Master vs non-master

The primary objective of a CGR is to share the authoritative data coming from the governmental entity having the official curation mandate over it. This is what is being referred to as the master data and information [7, 8].

While this is indeed the type of information that should be hosted as a priority in a CGR to serve as the source of truth, it is possible that either the master data/information is not available, not of quality (incomplete, out-of-date…), and/or not accessible.

In these cases, and to ensure the continuity of operations, it is important for a CGR to not only provide the possibility to upload, host, manage, and share “non-master” data/information in parallel or in complement to the master one (meaning that both types should be clearly differentiated) but also to promote for the convergence between the two once the master version is available, of quality, and/or accessible.

2.7.4 Accessibility

The full value of a CGR is expressed once its content is accessible to the largest number of users possible.

It is possible that not all content can necessarily be made publicly accessible, at least not until there are appropriate government approvals, it has reached a certain level of quality, and/or the necessary policy and legal framework is in place.

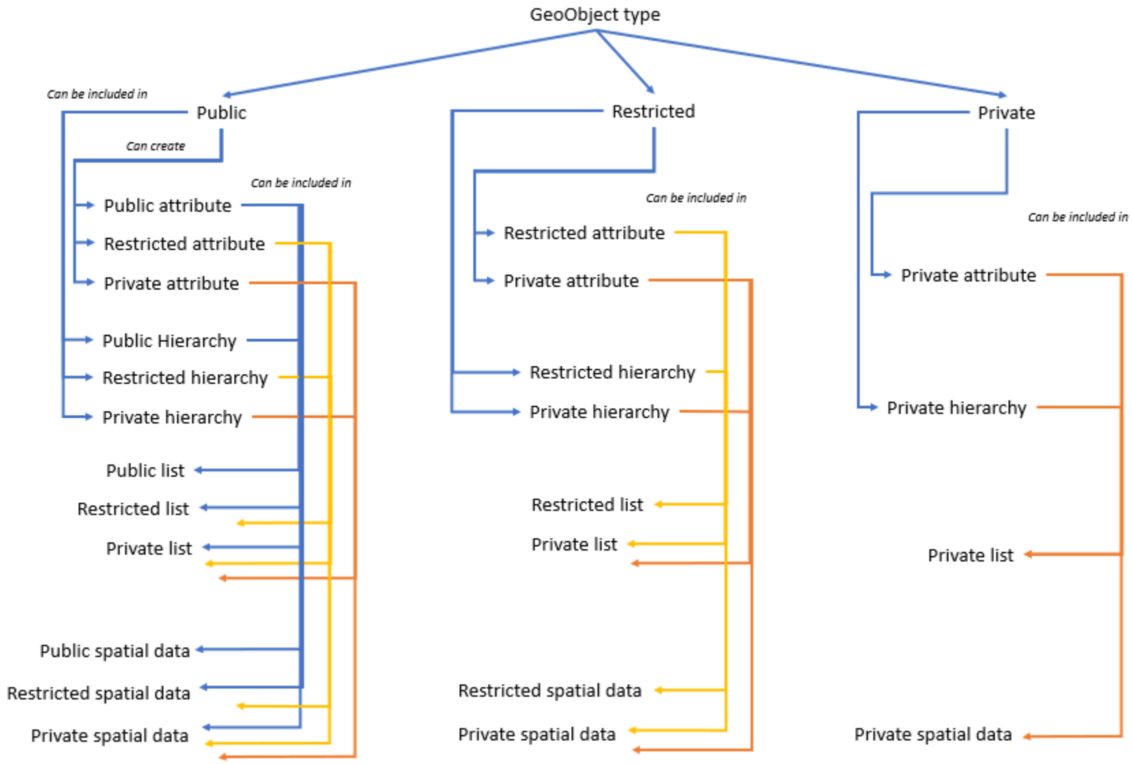
Therefore, it is important to ensure that a CGR can provide the capability to handle at least three levels of access for all its content:

- Public: content accessible by any user
- Restricted: content accessible to specific organizations registered in the CGR
- Private: content only accessible to the organization having the curation mandate over it

Doing so not only allows to respect the current accessibility policy that each organization might want to apply to the content they are hosting in a CGR, while still benefitting from the functionalities it provides, but also to change such a level as the data sharing policy associated with this content evolves.

The important thing, though, is to ensure that the accessibility set at the geographic object type level is transferred properly to the rest of the content stored in a CGR as presented in Figure 12.

Figure 12 – Impact of the accessibility right set at the geographic object type level to the rest of the CGR content



2.7.5 Documentation

Documenting the content of the CGR in a way that informs the users if such content corresponds to their needs can be considered as being part of data quality.

Such documentation should be captured in a metadata profile specific to each key element in the CGR (geographic object type, data element, hierarchy, list, spatial data) and accessible to the users both when consulting the information or data in the CGR or once exported.

The HGL guidance on documenting the data using a metadata profile [10] provides recommendations for the content of the metadata profiles to be considered for lists and geospatial data.

The following attributes should be considered for inclusion in the metadata profile to be associated with geographic object types, data elements, and hierarchies:

- Label/Title
- Description/Definition
- Validity period (for geo-object types and hierarchies)
- Accessibility rights
- Organization with the curation mandate
- Contact information of the person in charge

2.7.6 Languages and character encoding

The ability to store information in the country's local language(s) is another important capability that facilitate the uptake and use of a CGR in most countries.

This capability does not only concern the user interface but also some of the content at different levels. More specifically, it is important for the following to be captured in the country's local language in a CGR:

- The content of the metadata associated with geographic objects, data elements, hierarchies, lists, and geospatial data
- The values for some data elements (i.e., name of the geographic object, contact information)
- The description associated with each category in a classification table

Handling multiple languages does also require a CGR to provide the capability to handle multiple character encoding and fonts. While the UTF-8 Unicode standard is largely spread and should be the first one to be considered for inclusion (Section 4.5), some countries like Myanmar are still using other types of character encoding and different types of fonts.

In view of the above, a CGR should at least handle any language and, indirectly, multiple character encoding and, if possible, fonts from both a user interface and content perspective.

3. CGR functional requirements

This section provides the functional requirements that an IT solution should support to be considered a CGR.

These requirements are organized as follows and covered in the next sections:

- Data flow: Workflow that the IT solution should support for its content to be uploaded, hosted, managed, regularly updated, and shared
- Organizations and users: Minimum set of roles and associated rights that the IT solution should support for each organization
- Functionalities: Minimum set of functionalities that the IT solution should provide to support the data flow as well as organizations and users

3.1 Data and information flow

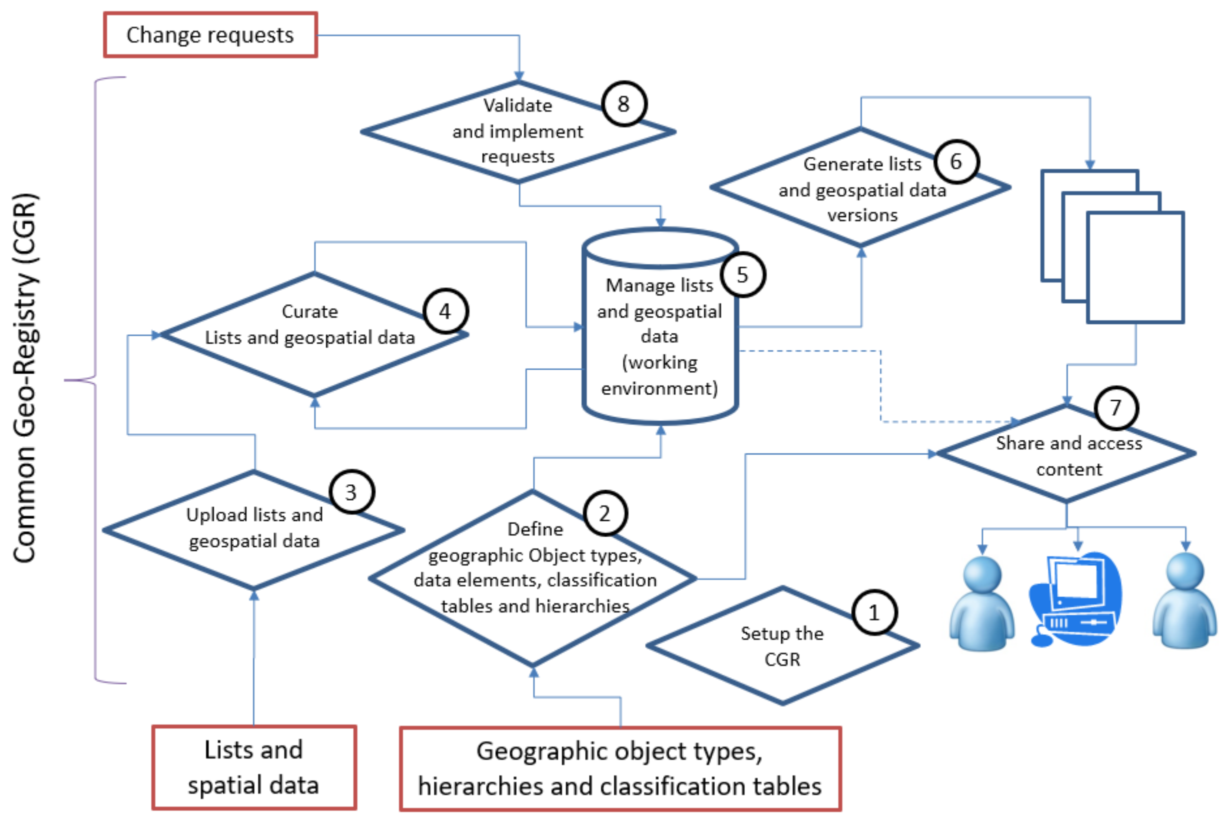
Once a CGR has been set up at both the cross-sectoral governance structure and organizational levels (Step 1 in Figure 13), data and information are expected to flow as follows:

- Step 2: Geographic object types together with associated data elements, classification tables, and hierarchies are defined in the CGR based on the exercise that identified them (Section 2)

- Step 3: Existing lists and geospatial data for the geographic objects identified in Step 2 are uploaded
- Step 4: Curation of lists and geospatial data takes place at the time of the upload as well as on demand when through its management (Step 5) or to address change requests (Step 8)
- Step 5: The content is managed in the working environment of the platform[14] to maintain its quality across the six dimensions of data quality
- Step 6: Versions are generated to capture multiple releases of lists and geospatial data
- Step 7: Once considered of sufficient quality, and in accordance with the data sharing policy, the content (geographic objects definitions, data elements, classification tables, hierarchies, version of lists and geospatial data, and eventually the lists and geospatial data from the working environment) is made accessible to users external to the organization.
- Step 8: Change requests submitted by authorized users are being validated and the corresponding change implemented in the lists and/or spatial data in the working environment if validated

Each of the above-mentioned steps are repeated as new content is added or modified in the CGR.

Figure 13 – Expected CGR data and information flow diagram



[14] Component of the backend used by the authorized users to work on the CGR content

3.2 Organizations and users

A CGR is a cross-program and cross-sectoral concept because it aims at covering geographic features under the curation mandate of different organizations (e.g., Ministry of Health for health facilities, health offices and health areas, Ministry of Interior for the list of administrative units and the National Mapping Agency for the geographic extent and boundaries of these administrative units). As such, the IT solution to serve as a CGR needs to be able to handle as many organizations as needed to reflect the data governance model established for each of the geographic object types it hosts.

The data governance model in question can be of three main types:

1. Centralized: The geographic object type is under the curation mandate of a single organization and the curation takes place at the central level.
2. Decentralized: The geographic object type is under the curation mandate of a single organization and the curation takes place in a decentralized manner (for example at the province level).
3. Federated: The curation mandate over a given geographic object type is spread between multiple self-governing organizations with each of these organization overseeing the curation over a single sub-geographic object type or geographic area for example.

While the concept of organization mainly corresponds to a governmental entity such as a ministry in the CGR context, some of these organizations might have the curation mandate over several geographic object types.

To facilitate shared usage of a CGR across multiple organizations, CGR user roles need to be defined at four different levels to operationalize the CGR data flow reported in Figure 13[15]:

- System Administrator(s) (SA) ideally attached to the CGR cross-sectoral governance structure (Section 4.1)
- Registry Administrator(s) (RA) for each organization defined in the CGR
- Registry Maintainer(s) (RM) and Registry Contributor(s) (RC) for each geographic object type under the curation mandate of each specific organization. The number of RMs would depend on the number of geographic object types.
- Anonymous users (AU) having access to the public content hosted in the CGR

Table 5 provides a description as well as a short profile for each of these roles.

Each user role is meant to be involved in specific steps of the data flow (Table 6) and meant to have specific rights across the different actions associated with each key component of a CGR (Annex 3).

[15] The first version of this guidance included a fifth type of user - the API consumer – but this is more a functionality providing an information system access to the CGR content through an API with the right of a specific user role (RA, RM, RC, AU) than a physical person.

Table 5 – CGR main user roles

Implementation level	User role	Role description	Profile
Cross-sectoral governance structure	System Administrator (SA)	A user having all the privileges of a Registry Administrator (RA) across all organization plus the capability to setup the CGR as well as manage organizations and RA	IT professional with a background in system administration and trained on the programs and tools composing the CGR backend
Organization	Registry Administrator (RA)	A user having all the privileges of a Registry Maintainer (RM) plus the possibility to manage users, geographic object types, hierarchies, lists, and geospatial data under the curation mandate of his organization	Head of the unit in charge of data management in the organization or his representative
Geographic object type (within each organization)	Registry Maintainer (RM)	A user having all the privileges of a Registry Contributor (RC) plus the possibility to manage lists and geospatial data for geographic object types over which he has the curation mandate	Data manager professional with experience in the management and use of geospatial data and technologies
	Registry Contributor (RC)	A user who has the possibility to access lists and geospatial data as well as submit change requests for the geographic object types over which he has the curation mandate	Field-based professional overseeing the regular update of the information associated with specific types of geographic object
Data/information	Anonymous User (AU)	A user who has the possibility to access the public data shared through the CGR without necessary having to login	Any individual interested in having access to the CGR public content

Table 6 - User's involvement in the CGR data flow

	SA	RA	RM	RC	AU
1.1 Set up the CGR (cross-sectoral governance)	X				
1.1 Set up the CGR (organization)	X	X			
2. Define geographic object types, data elements, classification tables, and hierarchies	X	X			
3. Upload lists and geospatial data	X	X	X		
4. Curate lists and geospatial data	X	X	X		
5. Manage lists and geospatial data in the working environment	X	X	X		
6. Generate lists and geospatial data versions	X	X	X		
7. Share and access content	X	X	X	X	X
8. Submit, validate and implement change requests	X	X	X	X	

3.3 Required functionalities

93 functionalities have been identified as being required for an IT solution to be considered as being suitable to serve as a CGR and, as such, support the management of its content through time (Section 2), the defined data and information flow (Section 3.1) as well as the multi-organization and multi-geographic object type nature of the CGR concept. These required functionalities are organized in Annex 4 according to the steps of the CGR data and information flow (Figure 13).

This list, together with the data and information flow as well as the requirements set at the organization and user level (Section 3.2), are meant to guide countries when assessing if the IT solution they are considering using, or developing, to serve as CGR is appropriate and, if not, identify gaps to be addressed.

In order to help during this process, a CGR self-assessment matrix has been developed in the form of a Microsoft Excel file and is available at the following URL:
https://healthgeolab.net/DOCUMENTS/Common_geo-registry_assessment_matrix_Ve3.xlsx.

This matrix is meant to be filled by the provider of the considered IT solution(s) before being validated by the group of experts in charge of deciding which IT solution should be deployed as CGR.

4. CGR supporting environment

Deploying and maintaining a CGR is a long-term investment that can only be sustained if the appropriate supporting environment is in place.

The supporting environment in question is composed of the 7 elements forming the first and second stage of the HIS geo-enabling framework pyramid (Figure 1). The following sections provide more details for each of these components. Please consult the implementation guide of the IGIF for additional information and benchmarks regarding the cross-sectoral dimension of this environment[16] and the Health Information System (HIS) geo-enabling framework toolkit[17] for the health sector dimension.

4.1 Governance

Establishing a good CGR governance structure is key to ensure an inclusive and cost-effective deployment of the CGR concept, IT solution, and content. Additionally, it is critical towards promoting the long-term sustainability of a CGR as a mechanism to operationalize SOPs and other elements for a supportive enabling environment.

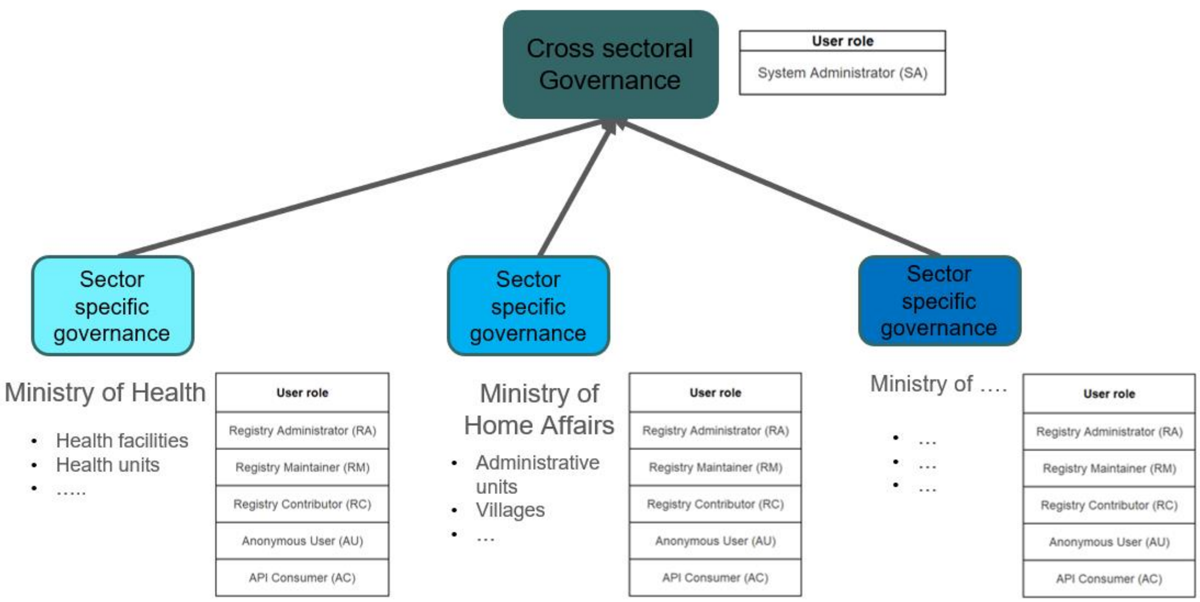
When possible, the following two-tier CGR governance structure, together with their respective functions, should be considered to leverage the cross-sectoral benefits provided by the CGR while respecting each sector's mandate (see Figure 14 for an illustration and the list of potential scenarios for both types of structures just after that figure):

1. Cross-sectoral governance structure:
    - Provides the cross-sectoral vision, strategy, and plan under which the CGR as concept and platform operates.
    - Hosts, manages, and maintains the CGR platform. As such represents the natural host for the System Administrator role (SA)
    - Encourages and facilitates coordination and collaboration between sectors, including the identification of the ministry having the curation mandate on each type of geographic object.
    - Defines the specifications, standards, and protocols that will ensure the quality of the CGR content across all 6 dimensions (uniqueness, completeness, timeliness, accuracy, validity, and consistency).
    - Develops and implements cross-sectoral policies supporting the availability, quality, accessibility, and use of the CGR content.
    - Trains and provides technical support to the ministries in charge of developing, maintaining, updating, and sharing content through the CGR.
    - Raises resources for the sustainability of cross-sectoral activities including the hosting and maintenance of the CGR platform.
2. Sector-specific governance structure:

[16] https://ggim.un.org/IGIF/part2.cshtml
[17] https://www.healthgeolab.net/DOCUMENTS/HIS_geo-enabling_toolkit.pdf

- Provides the vision, strategy, and plan under which the sector-specific CGR content is maintained, updated, shared, and used in alignment with the cross-sectoral one.
- Agrees upon the government entity having the curation mandate over each of the geographic features covered by the sector.
- Facilitates the maintenance, regular update, sharing, and use of the master lists as well as associated hierarchies and geospatial data for the geographic object types under the sector's curation mandate (e.g., health facilities for the health sector), including coordination and collaboration among concerned stakeholders, as well as the use of the content shared by other sectors.
- Applies the specifications, standards, and protocols defined at the cross-sectoral level to ensure the quality of the sector-specific CGR content.
- Develops and implements sector-specific policies supporting the availability, quality, accessibility, and use of the CGR content
- Raises resources for the sustainability of sector-specific activities including the maintenance of the sector-specific CGR content.

Figure 14 - Two-tier CGR governance structure and associated CGR user roles



Different scenarios can be considered for both types of governance depending on the country and this in order to be inclusive of all the skillsets necessary to deploy and then manage the CGR platform and its content. To give some examples (by decreasing order of long-term sustainability):

- For cross-sectoral governance:
  - National Spatial Data Infrastructure (NSDI)
  - Other type of formal forum led by one of the ministries (example: national agency for spatial development, working committee under the Ministry of Interior…)
  - Informal forum led by one of the ministries (Ministry of Health, for example)

- For the sector-specific governance:
    - Official established technical working group lead by the ministry (examples: Health Information System (HIS) working group, working group for the management and use of geospatial data and technologies).
    - Informally recognized leadership by one of the entities in the ministry (department, unit…)

In the absence of any cross-sectoral governance structure, or in the case when there is insufficient readiness or interest to deploy a CGR within the existing cross-sectoral structure, the deployment of a CGR could also take place within a given sector and then expand to cover other sectors when appropriate. When this is the case, it is important to ensure that the deployment involves both the management of the CGR platform (e.g., IT department) and of its content (e.g., department in charge of data management).

It is important for the governance structure established at both levels to be as inclusive as possible for the following two main reasons:

1. The curation mandate over the geographic objects core to sustainable development in general and public health in particular is distributed between different governmental entities among which we can mention the Ministry of Health (health facilities and health reporting divisions, health areas, etc.), the Ministry of Interior, the Ministry of Home Affairs, Ministry of Local Government, or the Ministry of Land Management (administrative units and human settlements) as well as any other ministries in charge of maintaining a master list for specific types of geographic objects such as the Ministry of Education (schools).
2. The deployment and maintenance costs of a CGR platform and content are significantly reduced if all the stakeholders are contributing to it instead of developing parallel platforms and processes leading to content that is not compatible

4.2 Vision, strategy, and plan

Vision, strategies, and plans are instruments providing the short- and long-term direction for the adoption of the CGR concept and the deployment of a CGR platform.

Due to its cross-sectoral nature, the CGR would ideally first benefit from the availability of a vision, strategy, and plan regarding the management and use of geospatial data and technologies for sustainable development in general.

These could, for example, be provided by the NSDI if in place in the country and/or the National Development Plan (NDP) if it considers geospatial data and technologies as important tools for its implementation.

At the same time, it is also important for the sectors to benefit from the deployment of a CGR to also have their own sector specific vision, strategy, and plan regarding the management, use, and sharing of geospatial data and technologies (e.g., Myanmar National Health Plan 2017-

2021[18]). These would ideally be aligned to the cross-sectoral ones provided by the NSDI and/or NDP.

If there is no NSDI in place in the country or the geospatial data and technologies are not part of any cross-sectoral or sector specific plans, then the deployment of a CGR can contribute to the establishment of such an NSDI in line with global framework such as the IGIF[19] and promote for geospatial data and technologies to be included in the next set of national plans.

4.3 Policies

The value of the CGR content is only fully accomplished if it is of quality, accessible, and if its use is enforced as the unique source of truth across information systems. Reaching all the three requires policies to be in place and enforced in a cross-sectoral and sector-specific way. Such policies should at least:

- Specify the governance structure and therefore indicate which governmental entity/ies has/have the mandate over the management and maintenance of the CGR platform as well as the development, maintenance, update and, sharing of the CGR content (master lists, hierarchies, and spatial data).
- Promote the open sharing and use of the CGR content by all the stakeholders. 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.
- Promote and incentivize a collaborative and coordinated approach to maintain and regularly update the CGR content.

4.4 Resources for sustainability

Human and financial resources are needed not only to support the deployment of the CGR platform and of its content but also to sustain their maintenance and regular update in the long term. In view of the ideal two-tier governance structure to be established, or leveraged if already existing, resources should be secured at the cross-sectoral level as well as by any sector sharing through the CGR.

For cross-sectoral activities, resources will be needed for the following:

- Facilitation of all cross-sectoral activities (planning and strategies, coordination, collaboration, standards, policies, meetings…).
- Setup, maintenance, hosting, regular update, and troubleshooting of the IT solution that will function as the CGR (hosting fees, maintenance fees, system administration staff).
- Technical support to the ministries in charge of developing, maintaining, updating, and sharing content through the CGR (training, hotline…).

[18] https://www.healthgeolab.net/KNOW_REP/myanmar_national_health_plan_2017-2021_eng_.pdf
[19] https://ggim.un.org/IGIF/

In each sector, resources will be needed to facilitate the following:

- Sector-specific activities (planning and strategies, coordination, collaboration, policies, meetings…).
- Maintenance, regular update, sharing, and use of the master lists as well as hierarchies and geospatial data for the geographic object types under the sector's curation mandate.
- Training, field data collection, staff (registry administrator, maintainers, and contributors) and
- Use of the content shared by other sectors.

4.5 Data specifications, standards, and standards operating procedures

Specifications, standards, and Standard Operating Procedures (SOPs) are key to ensure the quality of the CGR content across its six dimensions (uniqueness, completeness, timeliness, accuracy, validity, and consistency). The following should be covered as part of this element:

- *Terminology*: Defining the terminology behind the concepts and the data elements included in the CGR content is critical to ensure that all the stakeholders and the users have the same understanding of such content. For example, defining the concepts of health facility and administrative and reporting divisions is key to determine the content of each master list as well as associated hierarchies and spatial data.
- *Standards*: The use of documented and agreed upon standards contributes not only to ensuring consistency across the entire CGR content but also interoperability between the CGR platform and the information systems that will plug into it. As much as possible, globally recognized standards should be used to facilitate the use of CGR content by the international community. Among these standards, we can mention:
    - Character encoding: UTF-8.
    - Dates: ISO 8601 (YYYY-MM-DD).
    - Phone numbers: ITU-E.164 (Country Code (CC) + GSN (Global Subscriber Number).
    - Geographic coordinates: ISO 6709:2008 (Decimal degrees (EPSG: 4326); latitude: +/- DD.DDDDD; longitude: +/- DDD.DDDDD). This is the format suggested for computer data interchange and the most universally used to limit potential data entry errors. It is also recommended to capture the source and method used to collect the geographic coordinates as well as a qualitative measure of the coordinates' accuracy.
    - Metadata: ISO 19115 metadata standard as the basis for creating a metadata profile for geospatial data. For the CGR maintaining and sharing lists, it is also important to have a metadata profile for tabular information [10].
    - Data exchange protocols: To exchange content with other information systems over the internet, it is important that the CGR supports commonly used exchange protocols. The most fundamental exchange protocols are REST API for serving machine interoperable content, Web Map Service (WMS) for serving georeferenced map images, and Vector Tile for serving highly customizable map data. Thematic specific protocols can be important for

specific operations and should also be considered for inclusion. The Fast Healthcare Interoperability Resources (FHIR), which is a standard healthcare data exchange that includes an extension for sharing geographic location, is an increasingly important thematic protocol for interoperability of health systems.

- *Data specifications*: These specifications are meant to provide the characteristics and benchmark that the CGR content should comply with or reach to be considered as being of quality across the six dimensions of data quality. Some of these specifications can be based on standards. Annex 1 of the following guidelines gives an example of this kind of specifications: https://www.healthgeolab.net/KNOW_REP/KHM_MOH_Guidelines_2018.pdf.
- *Standard Operating Procedures (SOPs)*: These are needed to ensure data quality as well as the transfer of methods across institutions and individuals involved in the generation, maintenance, update, and sharing of the master lists using the CGR. As such, these SOPs should not only cover the flows being supported by the CGR but also all the processes linked to the collection, cleaning and updating of the content of each master list in line with the data specifications mentioned here above. As an example, the following guidance provides examples of SOPs for the collection of geographic coordinates in the field: https://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_4_2.pdf.

Aiming at the highest data quality right from the beginning of the process allows the content of the CGR to serve the largest number of use cases. This approach is cost-effective as the cost for collecting high quality data is not much higher, if not the same, compared to collecting poor quality ones.

If an NSDI is in place in the country, the above should be discussed and agreed upon as part of its activities and implemented by the different ministries including the Ministry of Health.

4.6 Technical capacity

Sufficient technical capacity needs to be present at different organizations and different levels to ensure the maintenance, regular update, and use of the CGR platform as well as the quality of its content.

When it comes to the platform, there is a need for the CGR System Administrator (SA) role to be established and sustained. This capacity is ideally either located within the entity providing the governance structure that has been established (Section 1.4.1) or closely related to it (e.g., in a cross-sectoral IT department).

The CGR content is itself meant to be managed by each organization having the curation mandate over it. As such, it is anticipated for each of these organizations to have the necessary technical capacity to fulfill and sustain the role of Registry Administrator (RA), Registry Maintainer(s) (RMs), and Registry contributors (RCs).

In each of these organizations:

- The RA role would ideally sit within a central-level entity having the mandate to provide technical support to the entire organization when it comes to the management and use of geospatial data and technologies (e.g., a Center for Health Information serving all departments and centers within a Ministry of Health).
- The RM(s) would sit within each of the entities having the direct curation mandate over the different geographic object types. This could either be at the central or at the subnational level depending on the size of the country (e.g., a Department of Planning for health facilities, a Department of Primary Health Care for community health workers).
- The RC(s) would ideally be part of the same entity as the RM but located at a lower level in the organization structure to be as close as possible to the field to provide timely updates (e.g., a district or provincial level health office).

Each role should have an alternate to ensure the continuity of service especially when it comes to the SA and the RA in each organization. The need for alternate RM and RC will depend on the updating mechanism and frequency.

Please refer to Section 3.2 for more details about the tasks, and therefore anticipated profile, for each of the above-mentioned user roles.

4.7 Appropriate geospatial technologies

The CGR is not meant to replace any other types of geospatial technologies but to complement them.

As such, there is a need for the deployment of a CGR to be accompanied by the choice and use of the appropriate geospatial technologies that ensure the quality of the geospatial data stored in the CGR.

The geospatial technology in question refers to:

- GIS software to prepare the geospatial data as a CGR is not meant to provide all the editing functionalities that GIS software can provide (topological check, snapping…).
- Global Navigation Satellite System (GNSS)-enabled devices to collect geographic coordinates in the field as a CGR is not meant to directly support field data collection.

When it comes to remote sensing (RS), even if satellite images can be used as basemaps in the CGR, a CGR is not meant to treat raw images which can be done in GIS software.

It is therefore important to ensure that the organizations having the curation mandate over content stored in the CGR have not only access to both the necessary geospatial technology and the related Standard Operating Procedures (SOPs) to ensure their proper use in the CGR context but also the necessary skilled staff to operate them.

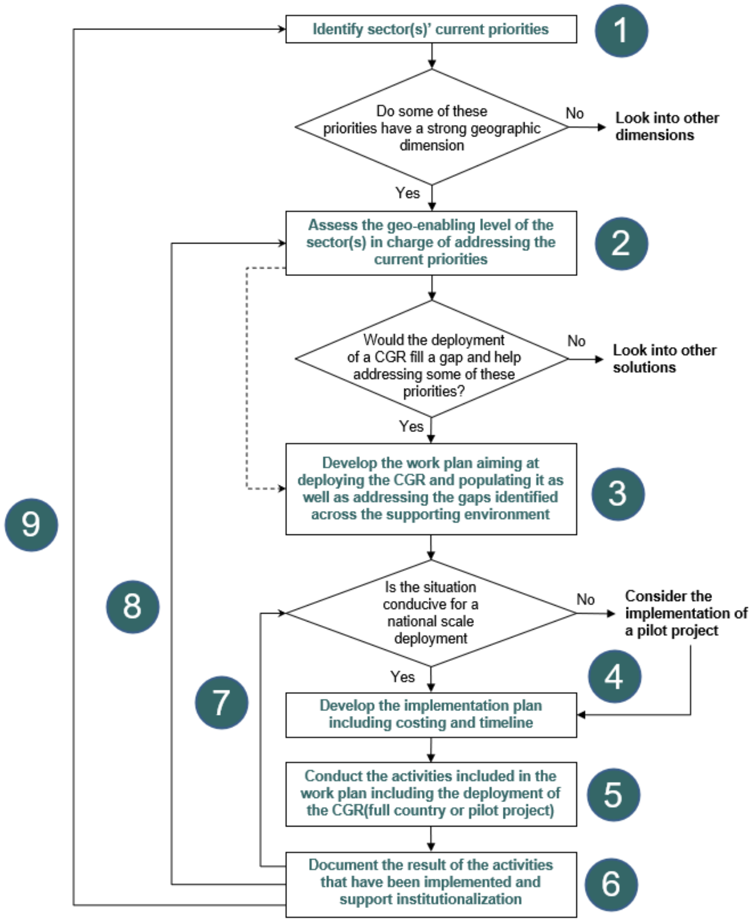
4.8 Considerations for in-country deployment

The long-term sustainability of a CGR solution depends greatly on the level of maturity of the supporting environment illustrated in Figure 1 and described in Section 4, as well as the priorities it addresses both cross-sectorally and within a sector. Identifying the potential gaps in such environment and aiming at addressing them in parallel to the deployment of a CGR will not only contribute to its long-term sustainability but also to the geo-enablement of the sector(s) involved in its deployment and/or use.

Additionally, the long-term use of the CGR content will only be possible if it is of the highest quality possible not only at the time of uploading it but also during the entire course of its existence.

In view of the above, it is important for the deployment of a CGR to take place as part of a larger process aiming at addressing all the elements that will ensure cost-effectiveness and long-term sustainability. The process in question could look like the one reported in Figure 15 – a process which is adapted from the HIS geo-enabling framework implementation process [2, 3, 1] and expanded to be applicable to other sectors as well as across sectors.

Figure 15 – CGR in-country deployment process



This process begins by identifying the current priorities of the sector(s) involved in the process (Step 1) and to determine if there is a strong geographic dimension to any of them.

If this is the case, Step 2 consists of assessing the level of geo-enablement of the sector(s) in charge of addressing the current priorities having a strong geographic dimension across all the dimensions reported in Figure 1. The quick assessment tool developed to assess the level of geo-enablement of the HIS [1] and in routine immunization programs [2] can be adapted for use as a starting point.

If the deployment of a CGR would indeed fill a gap and help address some of the priorities identified during Step 1, then the next step will consist in developing the work plan aiming at deploying the CGR and populating it with content as well as addressing the gaps identified across the supporting environment (Step 3). While the HIS geo-enabling toolkit [1], the guidance on the use of geospatial data and technologies in immunization programs [2], and the implementation support guide for the development of a national georeferenced community health worker master list hosted in a registry [8] will help in defining the list of activities to be anticipated across the supporting environment as well as the development of master lists, the following activities are also to be considered when it comes to the deployment of a CGR:

- Setup (server, software)
- Configuration (organization, users, localization)
- Population (upload of existing content)
- Testing and, if needed, enhancement
- Training (central and subnational level)
- Institutionalization (updating mechanism, technical capacity, ongoing maintenance, resources)

Once the activities are defined, it is important to determine if the result of the assessment indicates a situation that is conducive to a national scale deployment or if a pilot project covering only part of the country would be the most appropriate approach at this stage. When making this decision, it is important to consider that the implementation of the work plan also represents a perfect opportunity to strengthen the technical capacity of the sectors involved in the process. As such, implementing a pilot project at first might provide a proof of concept to demonstrate the value of geo-enabling the sector(s) through a CGR in addition to contributing to capacity strengthening. An example of the content and result of a pilot project implementation for the health sector in Myanmar is described in the following story map: https://arcg.is/1KL18X0.

The implementation plan, including costing and timeline, can then be developed (Step 4) and the activities included in the work plan conducted and monitored (Step 5).

Once completed, the result of the activities should be documented and used to support institutionalization by demonstrating the value of the process that has been implemented as well as establishing updating mechanisms for the CGR content.

From there:

- The situation could be re-evaluated to see if it is now conducive to a national scale deployment in the case of a pilot project implementation (Step 7)
- The geo-enabling level of the involved sector(s) could be re-evaluated to determine which part of the initial gap has been filled and re-adjust the original work plan (Step 8). In this case, there would be no need to re-evaluate the relevance of the CGR deployment (dashed arrow between Steps 2 and 3)

- The priorities of the involved sector(s) should be re-evaluated on a regular basis to see if they have changed and, as such, would require an expansion of the CGR content to include additional geographic object types and/or hierarchies to address them (Step 9). In this case as well, there would be no need to re-evaluate the relevance of the CGR deployment.

5. Conclusion

This guidance builds on the first version released in August 2017 to provide more information regarding the content, functional requirements, and supporting environment associated with the concept of Common Geo-Registry (CGR).

While the CGR concept has not changed since the first version of the guidance, the lessons learned during the development and deployment of the first IT solution aiming at operationalizing this concept confirmed the necessity to expand it beyond the health sector and demonstrated that the technology was allowing to expand the list of functionalities to be included in a CGR for the benefit of a larger group of users.

By capturing these lessons learned, this new version and associated assessment matrix provide a more comprehensive guide for countries and development partners to assess if the IT solution they are considering using, or developing, to serve as CGR is appropriate and, if not, identify gaps to be addressed.

At the same time, this guidance aims at promoting the proper integration of the geographic and time dimensions in information systems across sectors and within the health sector in countries. In the long term, such an integration supported by a CGR should result in all the sectors better benefiting from what geography, geospatial data, and geospatial technologies have to offer in support of geographically based planning and decision making and therefore a more systemic approach to solving developmental problems.

The content of this guidance is meant to continue expanding as the CGR concept is implemented in new countries especially in cross-sectoral settings as this is where the full potential of a CGR is revealed.

References

[1] Health GeoLab Collaborative (2018): HIS Geo-Enabling Toolkit: https://www.healthgeolab.net/DOCUMENTS/HIS_geo-enabling_toolkit.pdf

[2] UNICEF, GAVI, HGLC. 2020. Guidance on the use of geospatial data and technologies in immunization programs, overview and managerial consideration for in-country strengthening: https://www.unicef.org/media/58181/file

[3] Gavi, UNICEF & HealthEnabled. 2021. Leveraging geospatial technologies and data to strengthen immunisation programmes: rapid guidance for investment planning. April 2021: https://www.gavi.org/news/document-library/leveraging-geospatial-technologies-and-data-strengthen-immunisation

[4] DAMA UK Working group on data quality dimensions (2013): The six primary dimensions for data quality assessment: Defining data quality dimensions

[5] Ebener S. (2016): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.1 Documenting the process and defining the data needs. Health GeoLab Collaborative document: http://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_1.pdf

[6] Ebener S. (2016): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.2 Defining the vocabulary, the data set specifications and the ground reference. Health GeoLab Collaborative document: https://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part2_2.pdf

[7] USAID, WHO (2018): Master facility list resource package: guidance for countries wanting to strengthen their MFL: https://apps.who.int/iris/rest/bitstreams/1244370/retrieve

[8] CHAI, CHIC, HGLC, LivingGoods, The Global Fund, UNICEF (2021): Implementation support guide: Development of a national georeferenced community health worker master list hosted in a registry: https://www.unicef.org/media/113081/file/National-Georeferenced-Community-Health-Worker-Master-List-Hosted-in-a-Registry-2021.pdf

[9] Ebener S. (2016): Guidance for the management and use of geospatial data and technologies in health. Part 1 - Introduction to the data-information-knowledge-decision continuum and the geospatial data management chain. Health GeoLab Collaborative document: http://www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part1.pdf

[10] Ebener S., Pantanilla I., Mercado C.E., Maude R. (2018): Guidance for the management and use of geospatial data and technologies in health. Part 2 - Implementing the geospatial data management cycle: 2.5 Cleaning, validating, and documenting the data - 2.5.1 Documenting the data using a metadata profile. Health GeoLab Collaborative document: www.healthgeolab.net/DOCUMENTS/Guide_HGLC_Part_2_5_1.pdf

Annex 1 – Glossary of terms

Administrative unit	Delineated geographical areas within a particular sovereign state or territory created for the purpose of administration
Authoritative	Having, marked by, or proceeding from authority
Basemap	Collection of GIS data and/or orthorectified imagery that form the background setting for a map. The function of the basemap is to provide background detail necessary to orient the location of the map
Catchment area	A geographical area delineated around an institution or business, such as a health facility, from where the population utilizes its services
Classification table	Table organizing and categorizing data elements according to pre-defined criteria
Common Geo-Registry	IT solution that allows the simultaneous hosting, management, regular update, and sharing of the lists as well as associated hierarchies and geospatial data for the geographic objects core to development in general and public health in particular.
Community	A particular area or place considered together with its inhabitants.
Conceptual data model	A data model that represents an abstract view of the real world. A conceptual model represents the human understanding of a system
Data dictionary	A collection of names, definitions, and attributes about data elements that are being used or captured in a database or information system
Data element	Fundamental data structure in a data processing system. Any unit of data defined for processing is a data element. For example: Full name, Type, Address, etc. are each a separate data element. A data element is defined by its size (in characters) and type (alphanumeric, numeric only, true/false, date, etc.)
Data flow	A path for data to move from one part of the information system to another
Data flow diagram	A diagram used to graphically represent the flow of data in a business information system
Data management	All the disciplines related to managing data as a valuable resource. This covers, but is not limited to: data collection, cleaning, validation, documentation as well as the generation of data products (graphs, tables, and maps)
Functional requirement	Any requirement which specifies what the system should do
Functionality	The range of operations that can be run on a computer or other electronic system.
Geo-enable	To apply geospatial capabilities to a business process in order to establish the authoritative spatial location of any piece of information and

	data, and enable contextual spatial analysis
Geographic coordinates	A measurement of a location on the earth's surface expressed in degrees of latitude and longitude.
Geographic data	Information describing the location and attributes of things, including their shapes and representation. Geographic data is the composite of spatial data and attribute data (Examples: Number of doctors attached to the location of the health facility in which they are working, Number of Malaria cases attached to the administrative district in which they have been observed)
Geographic feature	Naturally and artificially created features on the earth. Natural geographic features consist of landforms and ecosystems. For example, terrain types and physical factors of the environment are natural geographic features. Conversely, human settlements or other engineered forms are considered types of artificial geographic features
Geographic information	Spatial and/or geographic data organized and presented to create some value and to answer questions
Geographic Information System (GIS)	An integrated collection of computer software and data used to view and manage information about geographic places, analyze spatial relationships, and model spatial processes.
Geographic object	A computer representation of a geographic feature (e.g., point, line, polygon)
Geographic object type	A programming domain abstraction for a common geo-registry interface that defines a type of geographic object (e.g., health facility, village, …) including its attributes and geometry type
Geography	The field of science devoted to the study of the lands, the features, the inhabitants, and the phenomena of Earth
Geospatial data	Also referred to as spatial data, information about the locations and shapes of geographic features and the relationships between them, usually stored as coordinates and topology (e.g., geographic location of health facilities, boundaries of administrative units, …)
Geospatial technologies	Refers to equipment used in visualization, measurement, and analysis of Earth's features, typically involving such systems as Global Navigation Satellite System (GNSS), Geographical Information Systems (GIS), and remote sensing (RS)
Global Navigation Satellite System (GNSS)	A satellite navigation system with global coverage
Governance	The processes, structures, and organizational traditions that determine how power is exercised, how stakeholders have their say, how decisions are taken, and how decision-makers are held to account

Health Facility	Fixed physical structure where healthcare is provided
Health Information System (HIS)	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
Hierarchy	An arrangement or classification of things according to relative importance or inclusiveness
Information	Data processed, organized, structured or presented in a given context so as to make it useful
Information Technology (IT) solution	A set of related software programs and/or services that are sold as a single package
Master list	Unique, authoritative, officially curated by the mandated agency, complete, up-to-date, and uniquely coded list of all the active (and past active) records for a given type of geographic feature (e.g., health facilities, administrative units, villages)
Metadata	Information that describes the content, quality, condition, origin, and other characteristics of data or other pieces of information.
National Spatial Data Infrastructure (NSDI)	Data infrastructure implementing a framework of geographic data, metadata, users, and tools that are interactively connected in order to use spatial data in an efficient and flexible way
Non-master list	List of active (and past active) records for a type of geographic object that is being used in a production environment for data collection and decision support by software applications without being curated by the mandated governmental entity.
Primary key	Data element (or combination of data elements) designated to uniquely identify each record in a table (i.e., the data elements containing the unique identifier)
Raster	A spatial data model that defines space as an array of equally sized cells arranged in rows and columns, and composed of single or multiple bands
Registry	An IT solution that allows storing, managing, validating, updating, and sharing the master list for a specific type of geographic object. It is the “container” for the master list.
Remote Sensing (RS)	Collecting and interpreting information about the environment and the surface of the earth from a distance, primarily by sensing radiation that is naturally emitted or reflected by the earth's surface or from the atmosphere, or by sensing signals transmitted from a device and reflected back to it. Examples of remote-sensing methods include aerial photography, radar, and satellite imaging. Screen reader support enabled.

Reporting divisions	Subnational divisions that are different from administrative units and that are used for decision making in the health sector and/or the reporting of health-related statistics
Requirement	That which is required; a thing demanded or obligatory:
Rule	One of a set of explicit or understood regulations or principles governing conduct within a particular activity or sphere
Settlement	A place, typically one that has hitherto been uninhabited, where people establish a community.
Specification	An act of describing or identifying something precisely or of stating a precise requirement.
Standard	A required or agreed level of quality or attainment
Standard Operating Procedure (SOP)	A set of step-by-step instructions compiled by an organization to help workers carry out routine operations.
Unique identifier (UID)	Data element in a relational database that is unique for each record
User story	Tool used in Agile software development to capture a description of a software feature from an end-user perspective
Vector	A coordinate-based data model that represents geographic features as points, lines, and polygons
Village	A clustered human settlement or community, larger than a hamlet but smaller than a town
Workflow	The sequence of industrial, administrative, or other processes through which a piece of work passes from initiation to completion.

Annex 2 – Example of data dictionary for a health facility master list

Data element group	Data element code	Data element description	Data element type	Data element size	Mandatory for adding a new record
Uniquely identify	HF_ID	Official unique identifier of the health facility	Varchar	8	X
	HF_N_EN	Official name of the health facility (English)	Varchar	150	X
	HF_N_MM	Official name of the health facility (local language)	Varchar	150	
	OP_DATE	Date at which the health facility has been officially opened	Date	10	X
Classify	HF_T_EN	Type of the health facility spelt in full (English)	Varchar	25	X
	HF_OWN	Ownership type (public, private) with indication of the entity owning the health facility	Varchar	50	X
	STATUS	Status of the health facility in the master list	Char	1	X
Locate	HF_ADD	Street number and name	Varchar	100	
	SR_C	Official unique identifier of the province in which the facility is located	Varchar	6	X
	SR_N_EN	Official name of the province in which the health facility is located (English)	Varchar	100	
	SR_N_MM	Official name of the province in which the health facility is located (local language)	Varchar	100	X
	DI_C	Official unique identifier of the district in which the health facility is located	Varchar	10	X
	DI_N_EN	Official name of the district in which the health facility is located (English)	Varchar	100	

	DI_N_MM	Official name of the district in which the health facility is located (local language)	Varchar	100	X
	VI_C	Official unique identifier of the Village in which the health facility is located	Varchar	6	
	VI_N_EN	Official name of the Village in which the health facility is located (English)	Varchar	100	
	VI_N_MM	Official name of the Village in which the health facility is located (local language)	Varchar	100	
	LAT	Latitude of the health facility	Num Floating	8	
	LONG	Longitude of the health facility	Num Floating	9	
	S_COOR	Source and method used to obtain the geographic coordinates of the health facility	Varchar	50	
	AC_COOR	Qualitative measure of the accuracy level for the geographic coordinates	Varchar	6	
Contact	HD_NA_EN	Full name of health facility head (English)	Varchar	50	
	HD_NA_MM	Full name of health facility head (local language)	Varchar	50	
	HD_PS_EN	Position of the Head of Facility (English)	Varchar	50	
	HD_POS_MM	Position of the Head of Facility (local language)	Varchar	50	
	LD_NBR_H	Landline telephone number of the health facility head	Varchar	50	
	LD_NBR_G	General landline telephone number at which the health facility can be reached	Varchar	50	
	MB_NBR_H	Cell phone number of the health of the health facility	Varchar	15	

	MB_NBR_G	General cell phone number at which the health facility can be contacted	Varchar	15	
	FAX_NBR	Official fax number at which the health facility can be contacted.	Varchar	30	
	EMAIL_H	Email address of the health facility head	Varchar	50	
	EMAIL_G	General email address of the health facility	Varchar	50	
	WEB	URL of the official website of the health facility	Varchar	200	

Annex 3 – Rights by CGR user roles

Legend for numbers associated to each role in the table:

1. For all the organizations
2. For his organization (any geographic object type)
3. For his organization (only for the geographic object types over which the user has the curation mandate)
4. For other organizations having given access to the restricted content

Few examples on how to read the information for each role:

- Cell highlighted in grey: The SA role can view the information of an organization for all the organizations.
- Cell highligthed in blue: The RA can manage data elements (including classification tables) for the restricted geographic object types and/or hierarchies for his organization (any geographic object type)

Step in the CGR data/information flow	SA	RA	RM	RC	Action (can)	Applicability (visibility)	CGR component
Step 1 (setup the CGR)	1	1			View the information of an ...	N/A	Organization
	1				Create an ...	N/A	
	1				Edit the information of an	N/A	
	1				Delete an ...	N/A	
	1	2*	2	2	View the account information of the RA, RM and RC ...	N/A	User
	1	2			Create a ... account	N/A	
	1	2			Edit the information of a .... Account	N/A	
	1	2			Delete a ... account	N/A	
	1				View the available ...	N/A	Localization
	1				Import a	N/A	
	1				Edit a ...	N/A	
	1				Export a ...	N/A	
	1				Delete a ...	N/A	
Step 2 (define geographic object types, data elements, classification tables and hierarchies)	1	1	1	1	View (including metadata) the ...	Public	Geographic objects types and/or hierarchies
	1	2, 4	2, 4	2, 4		Restricted	
	1	2	2	2		Private	
	1	2			Create (including metadata) ...	Public	
	1	2				Restricted	
	1	2				Private	
	1	2			Manage data elements (including classification tables) for the ...	Public	
	1	2				Restricted	
	1	2				Private	
	1	2			Edit the metadata of the ...	Public	
	1	2				Restricted	
	1	2				Private	
	1	1	1	1	Export the metadata (including classification tables) for the ...	Public	
	1	2, 4	2, 4	2, 4		Restricted	
	1	2	2	2		Private	
	1	2			Delete the ...	Public	
	1	2				Restricted	
	1	2				Private	

* The RA of each organization can also see the contact information of the SA

Step in the CGR data/information flow	SA	RA	RM	RC	Action (can)	Applicability (visibility)	CGR component
Step 3 (upload lists and geospatial data), Step 4 (Curate lists and geospatial data), Step 5 (Manage lists and spatial data in the working environment), Step 6 (generate lists and geospatial data versions), Step 7 (Share and access content); Step 8 (submit, validate and implement change requests)	1	2	3		Upload list and geospatial data (including metadata) in the ...	Public part	Working environment
	1	2	3			Restricted part	
	1	2	3			Private part	
	1	1	1	1	View the content (including metadata) of the ...	Public part	
	1	2, 4	3, 4	3, 4		Restricted part	
	1	2	3	3		Private part	
	1	2	3		Curate the content of the ...	Public part	
	1	2	3			Restricted part	
	1	2	3			Private part	
	1	2	3		Edit the content (including metadata) of the ...	Public part	
	1	2	3			Restricted part	
	1	2	3			Private part	
	1	2	3	3	Export the content (including metadata) of the ...	Public part	
	1	2	3	3		Restricted part	
	1	2	3	3		Private part	
	1	2	3	3	Submit change requests for the ...	Public part	
	1	2	3	3		Restricted part	
	1	2	3	3		Private part	
	1	2	3		Validate and implement change requests for the ...	Public part	
	1	2	3			Restricted part	
	1	2	3			Private part	
	1	2	3	3	Generate versions (including metadata) of ...	Public	Lists and geospatial data
	1	2	3	3		Restricted	
	1	2	3	3		Private	
	1	1	1	1	View the content of versions (including metadata) of ...	Public	
	1	2, 4	2, 4	2, 4		Restricted	
	1	2	2	2		Private	
	1	1	1	1	Export the content of versions (including metadata) of ...	Public	
	1	2, 4	2, 4	2, 4		Restricted	
	1	2	2	2		Private	
	1	2	3		Delete versions (including metadata) of ...	Public	
	1	2	3			Restricted	
	1	2	3			Private	
	1	2	3		Capture historical changes for the ...	Public	Geographic object types
	1	2	3			Restricted	
	1	2	3			Private	
	1	1	1	1	View historical changes for the ...	Public	
	1	2, 4	2, 4	2, 4		Restricted	
	1	2	2	2		Private	
	1	1	1	1	Export historical changes for the ...	Public	
	1	2, 4	2, 4	2, 4		Restricted	
	1	2	2	2		Private	

Annex 4 – CGR minimum required functionalities

1. CGR setup (11 required functionalities)

At the platform level:

1. Manage different organizations (create, edit, inactivate)
2. Support different data governance models (centralized, decentralized, federated)
3. Capture organization information
4. Provide organization information to users who need it
5. Manage localization for the user interface (select, import, edit, export)
6. Invite organizations to the CGR

At the organization level:

7. Define roles having specific rights down to the geographic object type level (e.g., administrator, maintainer, contributor)
8. Manage user accounts (create, edit, inactivate)
9. Assign a role to each user
10. Provide user information to users who need it
11. Invite users to the CGR

2. Define geographic object types, data elements, classification tables, and hierarchies (27 required functionalities)

Geographic object types:

1. Handle as many geographic object types as needed to cover the defined conceptual data model
2. Attach each geographic object type to a specific organization
3. Manage geographic object types (add, edit, delete)
4. Manage metadata for each geographic object type
5. Differentiate between master and non-master geographic object types
6. Specify access rights for each geographic object type (public, restricted, or private)
7. Define the geometry for each geographic object type (point, line, polygon)
8. Handle changes over time for each geographic object type down to a specific date (temporal validity of a given geographic object type)
9. Group geographic object types for the purpose of generating lists at the group level
10. Manage (add, edit, delete) core CGR data elements (uniquely identify, classify, locate, and contact) and store them in a data dictionary
11. Handle the format necessary to manage the core CGR data elements (text, date, Boolean, enumerated) as well as multiple languages
12. Manage metadata for each data element
13. Specify if the values for a particular data element are meant to be unique
14. Specify access rights for each data element (public, restricted, or private)
15. Specify if a particular data element is mandatory when adding new geographic objects
16. Create, edit, delete, and export the classification table associated with the enumerated data elements

Hierarchies:

17. Handle as many hierarchies as needed to cover the defined conceptual data model
18. Attach a hierarchy to a specific organization
19. Differentiate between master and non-master hierarchies
20. Specify access rights for each hierarchy (public, restricted, or private)
21. Manage hierarchies (add, edit, delete)
22. Graphically visualize the structure and content of a hierarchy
23. Handle changes over time down to a specific date (temporal validity of a hierarchy)
24. Manage metadata for each hierarchy
25. Use the same geographic object type in several hierarchies (no duplication of information)
26. Combine geographic object types from different organizations in a given hierarchy
27. Inherit the structure of other hierarchies having a geographic object type in common
28. Export a hierarchy in the form of a report

3. Import lists and geospatial data (6 required functionalities)

1. Import lists separately from geospatial data
2. Import the data either from the user's computer or from an external system
3. Specify the period over which the data being imported is valid (start/end date)
4. Specify the type of import (information for new records, update of existing information (overwrite), filling of value gaps)
5. Specify for which data elements the data is being imported, including the parents of the geographic object across the hierarchies in which it is included for the concerned organization
6. Match the data elements in the file being imported with those defined at the geographic object type level

4. Curate lists and geospatial data (9 required functionalities)

1. Identify and correct data element format validity issue (e.g., text imported against a date format data element)
2. Identify and correct hierarchy inconsistencies (correct parent attribution across the concerned hierarchies)
3. Identify and correct duplicates across geographic objects of the same type
4. Identify and correct temporal gaps for a given data element (missing values over a given period)
5. Identify and correct gaps for a given data element across all geographic objects of a same type (e.g., health facilities for which the type is missing)
6. Inform users about changes in the upper part of a hierarchy (e.g., creation of a new administrative unit requires the health facilities to be re-attributed according to the new structure)
7. Identify and correct orphans in a given hierarchy (e.g., new village not yet attached to a health facility as part of a catchment area)
8. Identify when the information for a given data element has been updated for the last time across all the geographic objects
9. Prevent from sharing the content from the working environment until it has been completely curated

5-7. Manage and share CGR content (29 required functionalities)

Lists:

1. Manage (create, edit, delete) as many lists as needed
2. Attach each list to a specific organization
3. Differentiate between master and non-master lists
4. Specify access rights (public, restricted, or private)
5. Make the distinction between the lists in the working environment and those that are being shared with users outside the organization
6. Use a unique set of geographic objects to generate any list (no duplication of geographic objects across the CGR)
7. Integrate multiple hierarchies in the same list or a different hierarchy in separate lists
8. Manage the metadata of each list
9. Access the data dictionary for the data elements included in the list
10. Manage lists at the geographic object level (add, edit, inactivate)
11. Handle data element changes over time down to a specific date
12. Create, visualize, and edit lists for any given temporal validity (date, period) without losing other data
13. Persist different versions of lists for each temporal validity to preserve what was referenced in the past (e.g. 01.01.2020-v1, 01.01.2020-v2, 01.01.2020-v3)
14. Explore the content of a list for a given temporal validity (date), including search and filter
15. Export lists for a given temporal validity (date, period) in a format readable by users together with metadata and data dictionary
16. Access the list and associated information, including metadata through a REST or streaming API

Geospatial data:

17. Manage (create, edit, delete) geospatial data associated with any list
18. Attach each geospatial data to a specific organization
19. Enable/disable geometry editing by geographic object type
20. Differentiate between master and non-master geospatial data
21. Specify access rights (public, restricted, or private)
22. Make the distinction between the geospatial data in the working environment and those that are being shared with users outside the organization
23. Manage the metadata of each geospatial data
24. Access the data dictionary for the data elements included in the attribute table of the geospatial data
25. Manage geospatial data at the geographic object level (add, edit, delete geometry)
26. Handle changes in geometry over time down to a specific date
27. Create, visualize, and edit geospatial data for any given temporal validity (date, period) without losing other data
28. Persist different versions of geospatial data for each temporal validity to preserve what was referenced in the past (e.g. 01.01.2020-v1, 01.01.2020-v2, 01.01.2020-v3)
29. Explore geospatial data on a map for a given temporal validity (date) including search and filter
30. Export geospatial data in a usable format for a given temporal validity (date or period) with metadata and data dictionary
31. Access the geospatial data with metadata through a REST or streaming API

Historical changes:

1. Capture and export changes over time for given geographic objects in the form of a report
2. Provide a REST API that supports queries on historical changes

8. Submit and treat change requests (6 required functionalities)

1. Submit change requests for any of the data elements associated with a given geographic object type (add, edit, inactivate) and a given temporal validity (date, period)
2. Review, accept, or reject the change requests that have been submitted down to the data element level
3. Provide the necessary information for the request to be reviewed (information before and after the change, temporal validity of the new information, reason for the request, etc.)
4. Provide the necessary information to follow up with the user who submitted the request if needed
5. Consult and monitor the change requests that have been submitted (including filtering and search)
6. Modify the change requests that have been submitted as long as they have not been reviewed