3.04 Data

3.04.01 Types of Data Used in a GIS

Although the two terms, data and information, are often used interchangeably, they mean two different things. Data can be described as different observations which are collected and stored. Information is processed data which is useful in answering queries or solving a problem.

"Analogue data," "paper version" or "hard copy" are terms often used to denote any document or dataset produced on paper while "digital data" or "soft copy" refer to files processed by GIS software in the computer. The result of the computer manipulated data can be transformed into a paper format such as the printout of a map.

Geographic data are inherently a form of spatial data organized in a geographic database. This database can be considered as a collection of spatially referenced data that acts as a model of reality. There are two important components of this geographic database: its geographic position and its attributes or properties. In other words, spatial data (where is it?) and attribute data (what is it?)

Spatial Data

Spatial data pertains to the location and spatial dimensions of geographical entities, and data that can be linked to locations in geographic space, usually via features on a map.

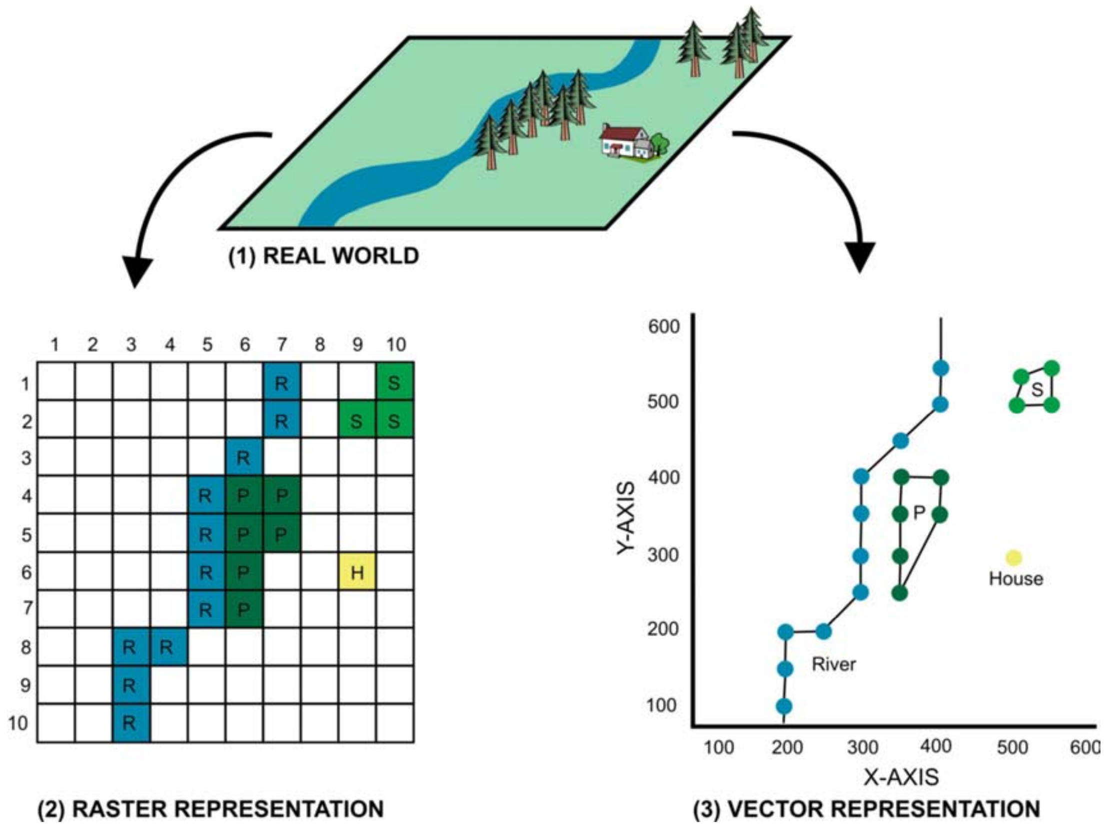
Attribute Data

Attribute data refer to the properties of a specific, precisely defined location. The data are often statistical but may be in text, images or multi-media. These are linked in the GIS to spatial data that define the location. They are often referred to as non-spatial data since they do not in themselves represent location information.

Spatial data can be represented into two fundamental approaches:

1. **Vector data** wherein objects or conditions in the real world are represented by points and lines and polygons that define their boundaries, much as if they were being drawn on a map. The position of each object is defined by its placement in a map space that is organized by a coordinate reference system, as shown below.

2. **Raster data** wherein the space is regularly subdivided into cells (usually square in shape), as shown in the figure below. The location of geographic objects or conditions is defined by the row and column positions of the cells they occupy. The area that each cell represents defines the spatial resolution available. The value stored for each cell indicates the type of object or condition that is found at that location in the raster model, and the homogeneous units are the cells.



Comparison of the Raster and Vector Models. The landscape in 1 is shown in a raster representation (2) and in a vector representation (3). The pine forest stand (P) and spruce forest stand (S) are features. The river is a line feature, and the house (H) is a point feature.

Some basic properties of raster and vector data are as follows:

- Each entity in a vector file appears as an individual data object. It is easy to record information about an object or to compute characteristics such as its exact length or surface area. It is difficult to derive this kind of information from a raster file because raster files contain little (and sometimes no) geometric information.
- Some applications can be handled much more easily with raster techniques than with vector techniques. Raster works best for applications where individual features are not important.

Comparison of Raster and Vector Data

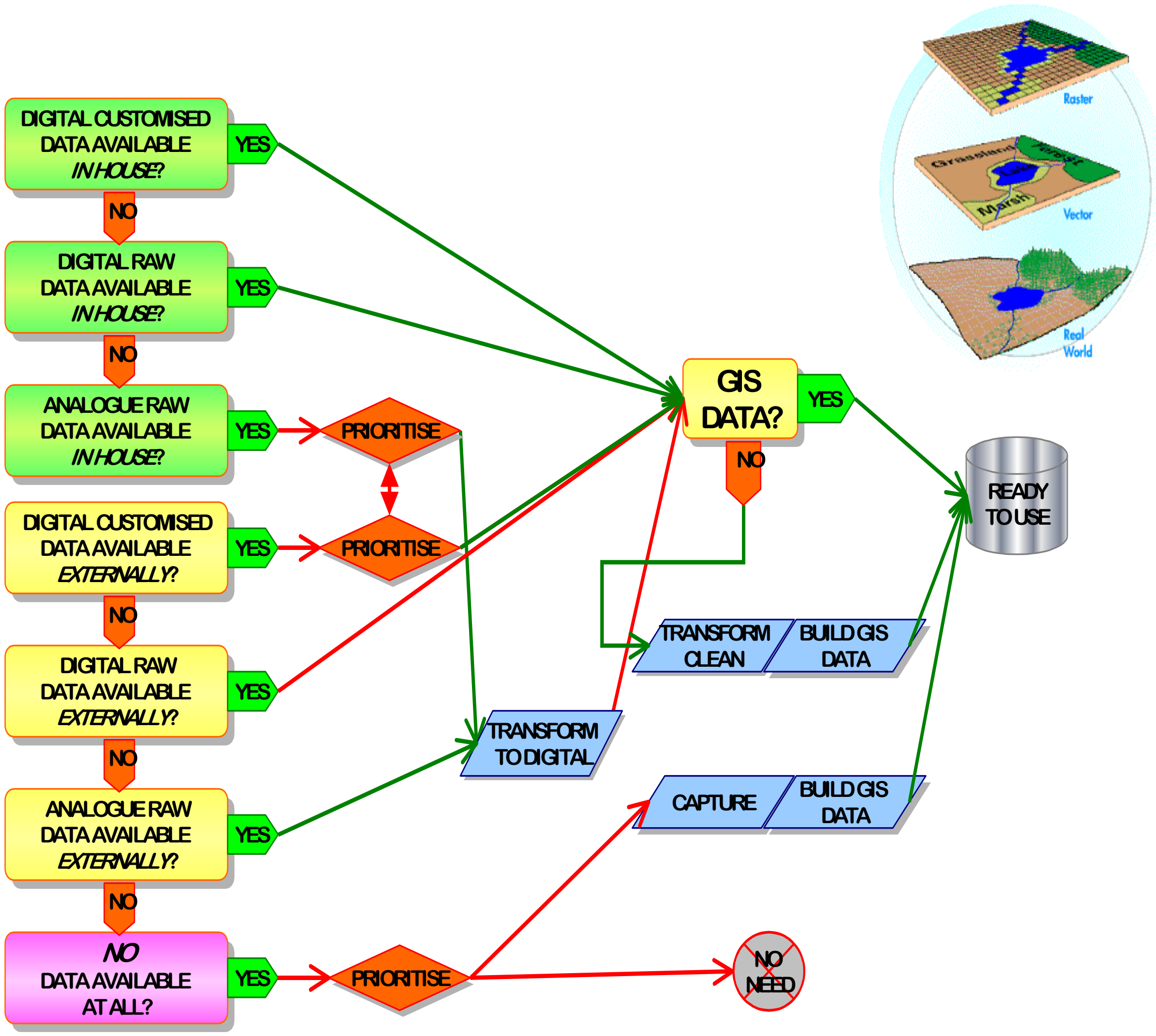
	Raster	Vector
Advantages	Good for complex analysis Efficient for overlays Data structure common for imagery	Compact data structure Efficient for encoding topology True representation of shape
Disadvantages	Large datasets Topology hard to represent Maps less "realistic"	Complex structure Overlay operations difficult Might imply false sense of accuracy

3.04.02 Data Preparation

Search for Data

Possibly the most important component of a GIS is the data. Geographic data and related attribute data can be collected in-house or acquired from a public agency or a commercial data provider. For the database building, standards for data acquisition and data entry, data maintenance and storage, data analysis and processing, data display and reporting have to be defined. By formulating and agreeing on a metadata base, specifications can be developed that facilitate the system integration.

The process of putting data into a GIS takes time. The process can be slow and laborious; and time equals money. Every year someone promises that next year there is going to be a faster, more intelligent scanning system that is going to get data into the system much easier. Things are indeed getting better and more and more data is becoming available in digital form, but the process of building a database still typically represents 80% of the first five-year costs of establishing a GIS. This is real money expenditure and that is where much of GIS time is going to be spent.



> *In this context one has to remember that the LGU is primarily an institution for data users, not data producers. Consequently, if customized GIS data is available on the market, it is better to purchase the data, instead of starting an in-house 'production line' to transfer data to a GIS format. The GIS Cookbook presents a collection of CLUP Data custodians to facilitate the data searching by LGUs in their CLUP preparation. There is an inventory of available or accessible attribute and spatial data that are needed in preparing the CLUP. (Chapter 4.17.01 in the Toolbox.)*

Data Capture

In the data capturing process the data are taken from the real-world (**primary source**), or from a **secondary source** such as a paper map, and entered into GIS software.

The Preparation of Primary (Attribute) CLUP Data

When the ‘Search-for-Data’ process starts, in some cases attribute data will not be stored in a digital format. The ArcView and most other GIS software have a tool to manipulate attribute data. However, for beginners in GIS it is recommended to use MS Excel for two reasons:

1. The custodian of the attribute data will (hopefully) be a representative from the specific sector (education, engineering, etc.) and the staff will most probably be familiar with the Microsoft Software package, which includes Excel.

2. The custodian of the spatial data will be the MPDO, and since the software is an expensive part of the GIS start package, it should be the unit that holds the GIS software license. As a consequence, it will be the MPDO who will assist the attribute data custodians in including the attribute datasets into the GIS. Furthermore, the MPDO will have to extend services by providing GIS browsers and producing print outs for the other GIS stakeholders so that they will be able to use the information in their tasks.

Eventually, the stakeholders will have the confidence to work with the attribute data in the GIS software and the methodology recommended during the ‘introductory phase’ mentioned above will cease to be a problem.

Aside from using Excel, it is also recommended that files to be used in the GIS should be stored in dBase file formats DBF4 (dBase IV). This is because in many instances, the dBase format can be used in the older version of the GIS software, for example ArcView 3.x. However, since the dBase IV format saves only the text and values as they are displayed in cells of the active worksheet, special attention is needed, as described in Chapter 4.20.01. How this is practiced is shown in Chapter 7.03.03.

The Preparation of Spatial CLUP Data

In the ‘Search-for-Data’ process, there will be instances wherein **primary data gathering** of spatial features will be done. It is recommended that a GPS be used in this activity. Chapter 4.19 in the Toolbox will show what to do in this case. There will also be analogue spatial data (paper drawn maps) that must be transformed into digital format. The process of capturing, processing and converting analogue spatial data into digital format is the same whether it is for basemap purposes or other maps for the CLUP. Chapter 4.21 will discuss these matters.

Map Accuracy and Level of Acceptance

> *GIS technology has broadened our view of a map. Instead of a static entity, a map is now a dynamic presentation of geographic data. The advantages are outstanding but there are also risks involved. In this case study, the importance of observing positional accuracy between the input data and the end product in the form of a CLUP map is shown.*

Six accuracy issues can be identified in a GIS:

1. Positional accuracy by which the location has been determined;
2. Attribute accuracy for the information describing a geometric element;
3. Logical consistency which means that lines are connected, polygons closed, etc.;
4. Completeness, which describes if the data is valid for the whole area or for parts of it;
5. Currentness that describes the time for data collection;
6. Lineage that describes all operations and manipulations that were used to produce the data (air photo i*nterpretation, digitizing, etc.)*

In the preparation of the CLUPs using GI Technology, **secondary source data** will be used. The LGU planner must rely on data captured by a national agency (e.g. geologic map, soil map, erosion map, flooding map, etc.). The source data will most likely be in a paper format, the data has been produced using manual methods, scales may vary, and little is known about the accuracy (few metadata is attached).

Chapter 4 in the Toolbox provides some metadata specifications for some of the data, but a lot more needs to be done to assist the planner. The source maps, in order to be useful in a CLUP GIS database, must be transformed into a digital layer. However, data from paper format will only be converted into digital format. Scanning and georeferencing are discussed in Chapter 4.21 where acceptance and accuracy should be observed in these processes. It is likely that errors inherent in the paper source will be also be transferred to the digital form including any errors that might have been incurred during scanning and georeferencing processes. The accuracy of the digital data will depend on the accuracy of the secondary source, and comparison would only be between secondary data sources. The way how to treat errors between primary and secondary sources will be discussed in the Toolbox.

How much error (errors from source and from scanning and georeferencing) is acceptable? The answer depends on how much accuracy the secondary source can provide. If the accuracy of a secondary source is not known, the data could be compared with other secondary sources which have similar features that are comparable.

However, one must be cautious in comparing data. Most secondary source data done manually would contain a lot of errors. It is also possible that there are secondary sources which were produced digitally like orthophotos and GPS surveys. These sources would have greater accuracy than all other secondary sources, and these secondary source data will have to be evaluated differently.

Lessons Learned

The spatial data, especially the data for the Base Map:

- must be captured with agreed and acceptable (positional) accuracy;
- must be properly georeferenced;
- must be defined in the right projection;

- must have enough information about how it was prepared (metadata)

It should not be expected of a planner to be able to assess whether ‘technicalities regarding the cartography’ are properly set from the beginning. There should be enough guaranties for the planner that the data has a workable standard so he can focus on his professional task, which is the actual planning and the preparation of the CLUP.

Metadata

> *Metadata is the term used to describe the summary information or characteristics of a set of data or "**data about data**".*

WITHOUT METADATA

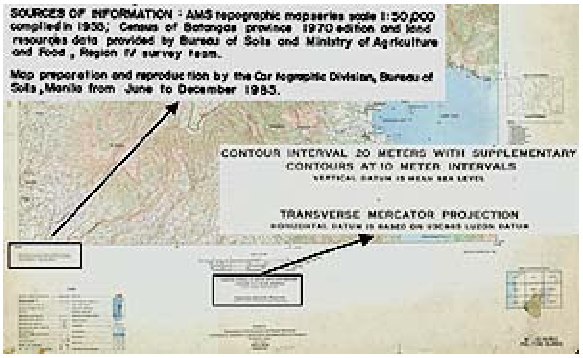
WITH METADATA

Metadata can be defined as geospatial data describing its characteristics in terms of content, quality, processing history, format etc, into a common set of terms and definitions. In simple words, metadata is “data about data”. A map legend on a paper map is a type of metadata that describes the different map elements, publishing date, projection and coordinate system, etc.

A common perception of GIS data is that it consists of two parts: spatial data (coordinates and topology), and attribute data (descriptive information). However, without proper documentation, GIS data will remain incomplete. It is thus equally important that GIS data also includes a metadata component. Metadata creation is typically considered to be an obligation of the data producer. The data user needs metadata to determine whether or not a particular data set exists, and to decide whether or not the data is appropriate for use. Proper metadata should describe the *who, what, when, where, why and how* regarding all aspects of a GIS data set.

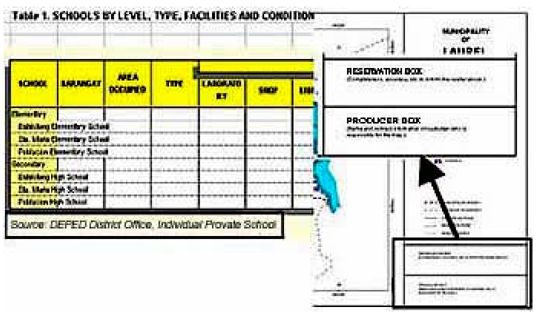
The use or creation of Metadata is often ignored or avoided. However, with the rise in use of digital data, the advantage of including metadata for datasets is increasingly recognized. Whereas cartographers rigidly provided metadata within a paper map’s legend, the evolution of computers and GIS has seen a decline in this practice. As organizations start to recognize the value of this ancillary information, they often begin to look at incorporating metadata collection within the data management process.

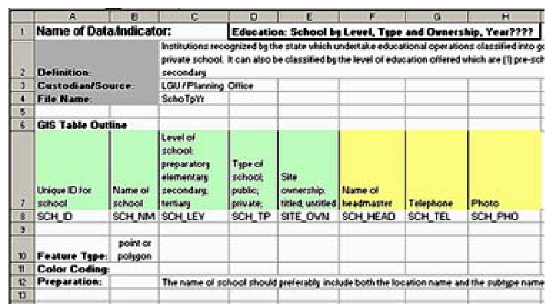
Metadata helps people who use geo spatial data find the data they need and determine the best way to use it. Metadata benefits the data-producing organization as well. As personnel change in an organization, undocumented data may lose their value. Incoming and newcomer staff may have little understanding of the contents and uses for a digital database and may find they can't trust results generated from these data. Lack of knowledge about other organizations' data can lead to duplication of effort. It may seem burdensome to add the cost of generating metadata to the cost of data collection, but in the long run the value of the data depends on its documentation.



Some metadata found on Source Maps

Sample Digital Metadata





Sample Metadata found here in the GIS Cookbook

In the GIS Cookbook there are Metadata Specifications and Standards for the attributes as well as the spatial datasets.

What Are Standards and Why Use Them?

> *The benefits of using GIS will be truly achieved once data is shared and exchanged between and among producers and users of geographic data. A prerequisite for such cooperation should be the capability of reading and interpreting the data among the exchanging entities. One basic condition is to* ***standardize*** *data, technically and conceptually*

Why do we need standards?

- A new project is formed - team building is required
- Improves options for coordination;
- Facilitates network operations and transfer of data;
- Enables cross sector working methodology;
- Increases transparency in the organisation;
- Increases opportunities to use staff in a 'slimming' organisation;
- Facilitates database operations;
- Improves the 'institutional memory' of the organisation;
- Facilitates searching for information

Paper Maps Means Conceptual Standards As Well The printed map, in itself, represents a standardized way of describing geographic information. With our knowledge, experience and intuition we understand a meaning, an image and properties of that road which is described with a certain symbol. It works pretty well as long as we deal with a certain map category. The problem is that the important aspects can easily draw in all information on the maps when performing analysis procedures by using a number of different thematic maps.

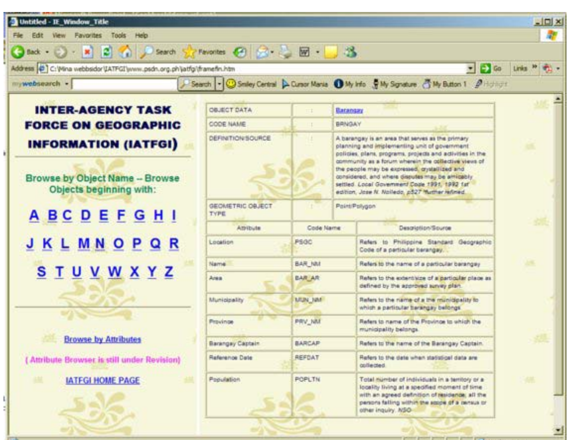
Computer Assistance Will Increase the Demands for Systematic Management of Data When changing to the digital world, there is a need to describe the tasks in a logical manner to get the computer to do what we want.

A Corporate Language GIS, as well as our own language, is created to transfer and disseminate information. A corporate language consists of a corporate vocabulary and a corporate grammar. In the computer world we talk about corporate feature names, feature definitions, attribute lists and uniformly defined data format and data base design. This is standardization

Use of Geographic Data Many organizations use many types of geographic data from numerous data vendors or producers. These data should be used together. Standardization concerning geographic data such as using the same projection is an absolute prerequisite.

As we use many data types from different producers it is also necessary with information about who is producing what, about data quality, about data capture methods etc. This is metadata. A uniform metadata structure also requires standardization, in order to easily understand the meaning of metadata.

Use of Geographic Data A standard is agreed upon by a group of users who have cooperated in order to standardize a certain thing. The work is approved by the standardization organization and appointed official standard. In addition to the official standards for geographic data, a certain group can decide to apply a standardized data description for a certain purpose. In this case the result will occur as a de facto-standard. This needs no approval by a standardization organization since it is just for the use of the internal organization that agreed on this standard. Today there are a number of official standards concerning geographic data. Those are developed within the International Standard Organization (ISO), for example ISO TC 211 (Global level).



There are also a lot of other unofficial standards. One example is the product de facto- standard established by

Microsoft as this company is dominating the software market for computers. Another strong player is Environmental Systems Research Institute (ESRI), the world leading vendor of GIS software.

In the Philippines, the Inter Agency Task Force on Geographic Information (IATFGI) has made serious effort to come up with technical standards for geodata.

The preparation of the GIS Cookbook has been coordinated with their recommendation and applicable metadata specifications have been adopted. However, the metadata specifications have been improved focusing not only on national government institutions but the local government data environment as well.

3.04.03 Data Management

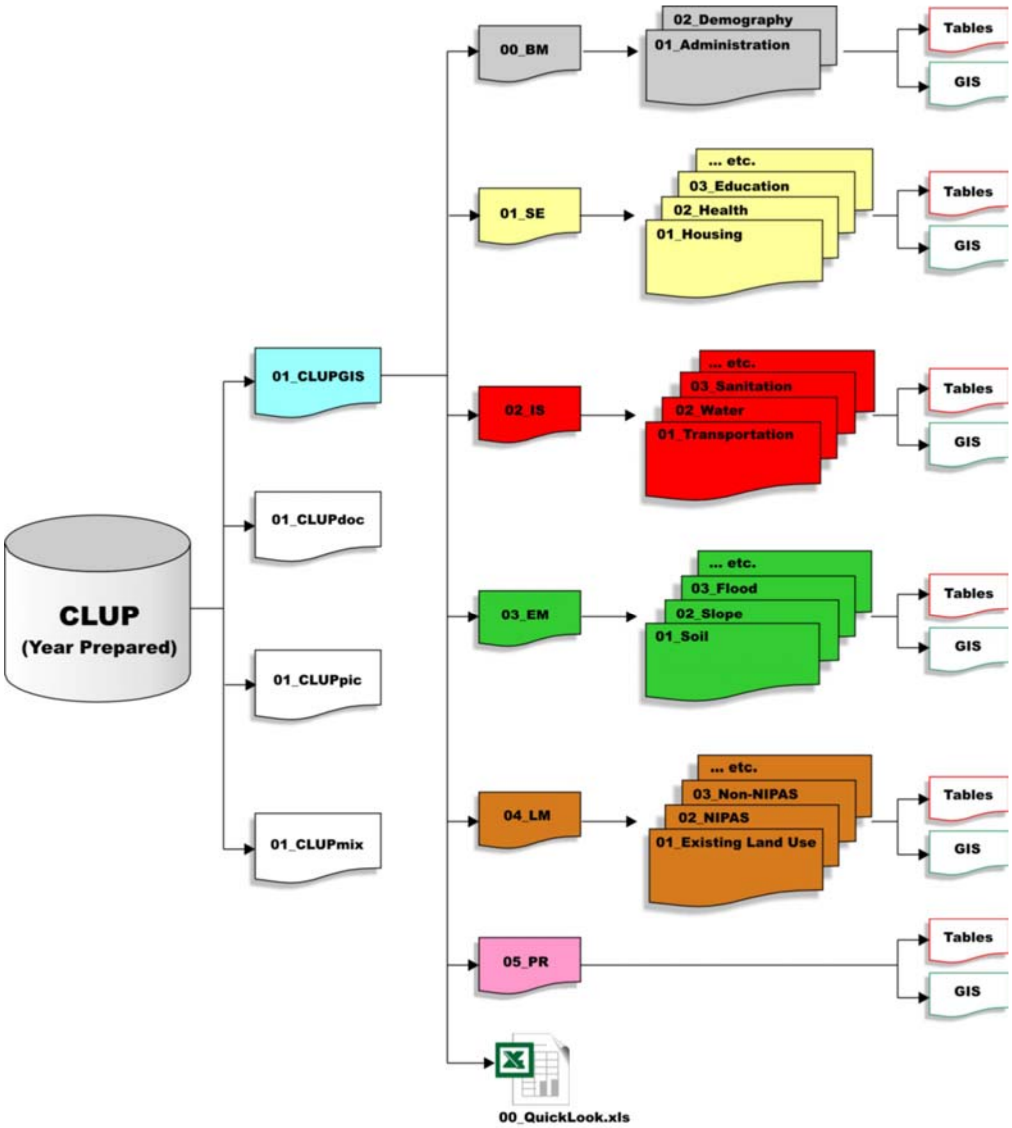
> *In a GIS it is very important that data is named and stored in a logical way otherwise it will be difficult to find, which version that should be used, and to maintain the information property.*
>
> *If there are no previous file and folder management standards in the municipality, the following guidelines should be followed. In case there is a recognized file and folder system in the municipality then that system's standards should be used. The guidelines are meant for a stand alone computer environment, with the files stored locally in one computer. In the case of networks, standards for data sharing should be applied.*

Guidelines for File and Folder Management

The goal is to minimize duplication of datasets and to have the data well organized and easily accessible. This will help avoid confusion during the CLUP preparation, and in future revisions of the CLUP.

To facilitate an overview of the folders, the subfolders should be organized in a specific order. They are automatically placed first in numerical and then alphabetical order. If you start with digits you can decide the appropriate order. It might not be necessary to use figures for all folders, but this is preferable for the most used or important folders. It is important to name the folders and files in a coherent way, so that will be easier to view the content of the drive. Using meaningful names and abbreviations can help see at a glance what each dataset is.

The folder structure described below is a proposed setup that can be used in the preparation of the CLUP. It is recommended for better organization and management of files in case no previous standard has been used by the municipality.



CLUP FOLDER STRUCTURE

All the files such as written reports and other documents, graphs and photos used in the narrative part of the CLUP and the geodata needed to build up the CLUP GIS, are organized into 4 folders, which then are divided into subfolders and sub-sub folders accordingly: **01_CLUPGIS**; **02_CLUPdoc**; **03_CLUPpic**; **04_CLUPmix**.

01_CLUPGIS – contains the data, mostly tables/spreadsheets that is needed for the GIS. The building stones of the GIS consist basically of spatial data (which configures the features on the map), and attribute data (which describes the specific map feature). For example, a school is represented as a point on the map (spatial data) and when you click on it one will find information on how many teachers, classrooms, etc. (attribute data) the school holds.

Aside from the geodata there are also (Excel) table data that have no GIS representation, and can be used in the narrative part of the CLUP report as tables or graphs originating from the spreadsheets.

The components of the CLUP GIS data are divided into sector folders which follow IATFGI recommendations on metadata as shown below.

Code	Name of folder	Content of sector folder
BM	Basic Information	Fundamental data sets that can be used to make the sector data described above more meaningful. Displaying or analyzing the base data with the sector data assists the user in making more effective and well-informed decisions.
SE	Socio-economic	The **'software issues'** which in a GIS context are combined due to international GIS standards and technical rationales. Data related to public services and economic development.
EN	Environment	The **'valid to'** tag, which identifies data that can be used to protect and develop environmental sustainability, conditions in the municipality.
IS	Infrastructure	Infrastructure is the **'man made features'** 'with layers, which depict the location, distribution, volume, standards and type of infrastructure utilities within the municipality.
LM	Land (use) Management	Land-use Management' symbolizes the **'price tag**' with layers which provide basis for zoning, land ownership, taxation and assessment of land values, which can be inputted to fiscal resources of the municipality.
PR	Project Management	Monitoring development activities in projects that have been initiated by the CLUP or have impact on the land use.

Each of the sector folders is divided by planning component subjects (Housing, Education, etc.) in order to differentiate between table files being used for preparatory activities (both for the GIS and to be inserted in the CLUP narrative text), and files that are being used in the GIS. Each planning component subject folder is further subdivided into two subfolders, namely '**Tables**' and '**GIS**.'

A 'Quick-look' file placed together with the sector subfolders in the CLUPGIS folder describes important information about the data, which could be of good use and facilitate understanding by a new user/custodian. Refer to Chapter 5.01.01 for more information about the 'Quick-look' file.

The GIS Cookbook does not give any recommendation how the data used in the CLUP Report should be organized. However, below are some general suggestions:

02_CLUPdoc – contains drafts of the CLUP document eventually divided into subfolders for drafts and final version. Each subfolder is recommended to have numbered subfolders corresponding to the division of chapter in the document, such as, 01_Introduction; 02_Baseline Studies; etc. .

03_CLUPpic – contains all types of imagery, such as photos, satellite imagery, aerial photos, graphic illustrations, etc. For easy reference it is recommended that all imagery used in the final version should be placed in a separate subfolder and if there are several images, these may be subdivided into chapters such as 02_CLUPdoc.

04_CLUPmix – contains miscellaneous files, preferably organized into subfolders according to the steps in Volume 1 prepared, such as minutes from meetings and consultations; correspondence, etc.

> *A preset directory that can be copied and inserted in the computer is also found in the Toolbox.*

Guidelines for Naming of Files

It is important to name the folders and files in a coherent way, so that it will be easier to view the content of the drive. Using meaningful names and abbreviations can help see at a glance what each dataset is.

The following guidelines are recommended, where the name of the folder or file should be:

- Clear and comprehensive;
- Not too long, not more than 40 characters (including space between words);
- Written following the sentence rule (start with capital letter);
- Acronyms with capital letters;
- No dots, slashes and backslashes. Only underscore can be used.

The following table sets out the characters that may NOT be used in file or folder names, as they are generally reserved by the operating system and will cause file retrieval problems if used:

Character	Description
/ Or \	Slashes ("/" or "\") – these are used by the operating systems to denote directories.
: Or;	Colons (":") or semi-colons (";")
*	Asterisks – used in search criteria as wildcards
%	Percent symbol
() or [] or{}	Brackets
.	Period – used to denote the file extension
?	Questions marks – bad form
=	Equals sign
" or '	Quotation marks
< or >	Greater than or less than signs
$	Dollar sign – this has a special usage for security permissions.
~	Tilde – used by the operating system to truncate files names that are too long.
!	Exclamation marks – bad form.

It is recommended that the geodata files be named as follows:

- Product/ feature name + year (2 digits) + eventual more detailed description about the feature + property, version or other property information. file extension
- For example: Admin96b_pline.shp. where
    1. 'Admin' is the code for an administrative feature;
    2. '96' indicates the year the data was captured/revised (for example when the CLUP was prepared);
    3. 'b' defines the type of administrative feature, namely a barangay (b is the coding for a barangay);
    4. '_pline' is the polyline version (as there is also a polygon version of the same feature needed for the base map)

> *The shape file format defines the geometry and attributes of geographically referenced features in as many as five files with specific file extensions that should be stored in the same project workspace. They are:*

> ***. shp*** *- the file that stores the feature geometry. Geographic features in a shapefile can be represented by points, (poly) lines, or polygons (areas).*
>
> ***.shx*** *- the file that stores the index of the feature geometry.*
>
> ***.dbf*** *- the dBASE file that stores the attribute information of features. When a shapefile is added as a theme to a view, this file is displayed as a feature table.*
>
> ***.sbn*** *and* ***.sbx*** *- the files that store the spatial index of the features. These two files may not exist until you perform theme on theme selection, spatial join, or create an index on a theme's Shape field. If you have write access to the source data directory, the index files will be persistent and remain after your Arc View session is complete. If you do not have write access to the source data directory, they will be removed when you close the project or exit Arc View.*
>
> ***.ain*** *and* ***.aih*** *- the files that store the attribute index of the active fields in a table or a theme's attribute table. These two files may not exist until you perform Link on the tables. If you have write access to the source data directory, the index files will be persistent and remain after your Arc View session is complete. If you do not have write access to the source data directory, they will be removed when you close the project or exit ArcView*
>
> ***.apr*** *is a project file in ArcView3*
>
> ***.mxd*** *is a map document in ArcGIS.*

Data Sharing

GIS and supporting technologies will lead to the development of decision support systems that facilitate the municipal planning process. By using indicators and alternative development scenarios it is possible to measure the performance of the LGU and future land-use.

Planning support systems like the CLUP GIS can measure and compare performances of different planning scenarios according to planner- or citizen-defined indicators for land use, transportation, education, natural resources, and employment, to name a few.

However, the ultimate goal is to bring together all potential players to work collaboratively on a common vision for their community. GIS-based planning support systems allow planners to quickly and efficiently create and test alternative development scenarios and determine their likely impacts on future land use patterns and associated population and employment trends, thus allowing public officials to make informed planning decisions. With a basic understanding and implementation of data sharing one can provide more information to local residents and the municipality without increasing capital or personnel costs. Employing these techniques will actually reduce the amount of time spent updating municipal management and planning data and increase accuracy and timeliness.

The idea that is advocated for in the GIS Cookbook is that much of the data presented in the CLUP tables (see Chapter 5 in the Toolbox) can be designed/formatted so they can be used both in the CLUP preparation and in the day-to-day work of the respective sector office (health, education, social welfare, building and business permits, etc.) that is responsible for providing the specific municipal service.

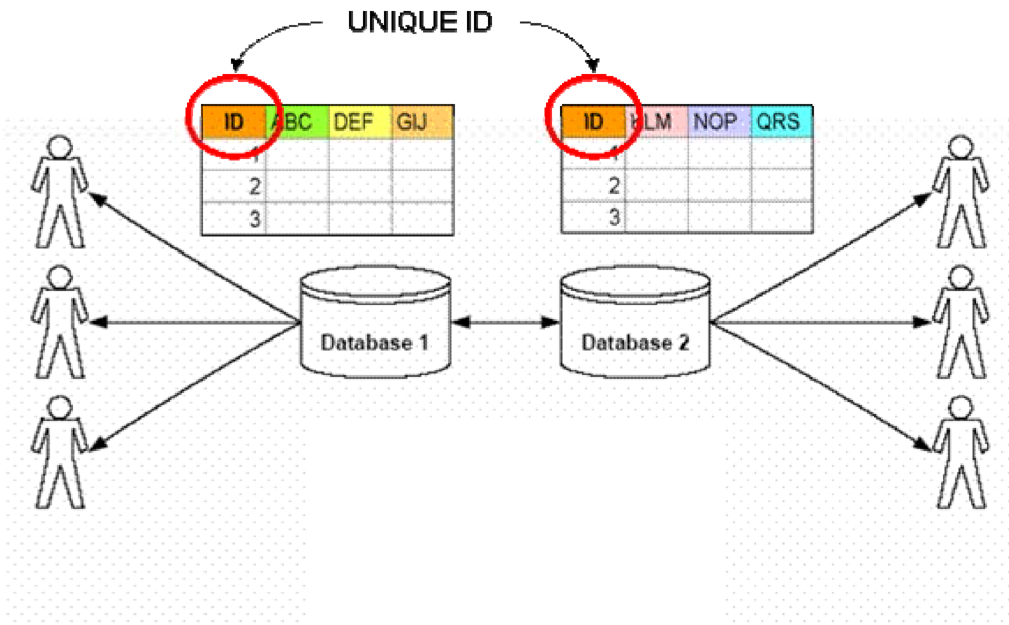
Once municipal offices (and other government agencies interacting with the LGUs) agree to share or replicate the data, they face the challenge of maintaining up-to-date datasets. Both attribute and spatial data are changing continuously as new social services, infrastructure, etc. are provided, or more accurate data is collected. To maintain up-to-date databases the various data “owners” (custodians) must exchange their most current datasets with those they share their data with.

This can be done in two ways:

1. Complete data load. This is the most straightforward approach. The current dataset is removed and completely replaced with the new dataset. However, this approach is sometimes impractical due to the volume of data, which may be difficult to distribute and take a prohibitively long time to reload, resulting in the database being inaccessible to the users for extended periods of time;

2. Change only updates. This approach requires smaller data volumes to be distributed as only the records that have changed (modifiications, deletions and additions) are exchanged. Change only updates also reduce the time for the data load because of the smaller data volume. The update process is more complex than the complete data load approach.

Corporate datasets and working databases may also have different data models (or schemas). Posting scripts are used to control the transfer of the data between the different databases, and these scripts must be capable of handling these different configuration issues and formats, as shown in the figure below.



Unique Feature Identifiers: To simplify the update process, unique ID's are used to keep track of joining tables, which features have changed, etc. Consequently all CLUP GIS tables, (see Chapter 5) have been given a field for a unique ID. For example, a school unit will always be identified with a unique alphanumeric ID which is referred to by all users and used when joining tables in a GIS. A good example on unique ID is to start from the coding of municipalities (and barangays) that is used by NSO (see Chapter 5.09.01 for more detailed information).

Data Ownership: It is important to clarify data ownership to eliminate potential conflicts.

For example, who 'owns' the table data for education? Which department is responsible for maintaining the school unit locations and attribute data about enrolment? Data ownership may also have to be shared. For example in a low-income municipality it might be the best solution that the planning unit takes responsibility for the data management of the spatial data, and see to it that the locations of schools are properly identified, while the school unit keeps records on the attributes such as number of classrooms and teachers, etc.

However, aside from agreeing about unique IDs and Data Custodianship, for municipal offices that share data with external users (those outside their administrative sphere of influence), "change only updates" result in a number of potential challenges that may include versioning, data transactions, data validation, coordinate systems and accuracy. Sometimes the CLUP/corporate datasets (shape files, Excel) are a different format to the external databases (ESRI Geodatabase, Oracle Spatial, MapInfo TAB, GeoMedia, AutoCAD, etc.). To cope with these issues there is a need for special GIS and IT knowledge.

In the Toolbox (Chapter 4.18), some examples illustrate the benefit of data sharing.

Data Security

Whatever the current value of the database, if it is properly maintained, this will increase considerably over the years. A successful GIS will be an integral part of daily operations. Over time, the value of information derived from the GIS database grows beyond a monetary cost to one measured by the functionality it provides to the work. Consequently, considerations for the protection of the GIS from damage will be necessary at some stage.

The possibility of the system and data being destroyed or severely damaged is real and deserves attention. The system is vulnerable to both deliberate and accidental damages. A disgruntled employee might purposely corrupt data, hackers may steal information, or a computer virus could find its way into the server. Natural disasters also pose a threat. Earthquakes, floods, fires, hurricanes, tornadoes, and lightning are all examples of natural hazards that could disrupt a GIS.

While deviant behavior and natural disasters are intriguing subjects, threats more common are found in day-to-day operations. Examples include coffee being spilt in the wrong place, a well-intentioned employee who accidentally deletes or corrupts a database, or a power disruption with no automatic battery backup.

When conducting a security review, the physical, logical, and archival security of the databases are examined.

Physical security measures protect and control access to the computer equipment containing the databases. Protection of database storage includes guarding against human intrusions (such as unauthorized personnel) and environmental factors (such as fire, flood, or earthquake).

Logical security measures protect and control access to the data itself. For example, users may be restricted to certain types of terminals, particular datasets, and particular functions. One common security measure is to ensure that only database management staff have editing and update rights to particular datasets.

Archival security is essential for many applications. Metadata, information about past coding and updating practices, the location of data, and the type of media on which data is stored, must be kept track of to allow for data recovery.

The table below illustrates the sections and subsections that might be included in a document that describes the security recommendations of systems and databases for a municipality. Recommendations are made that affect the current and future operations. This document will also help set priorities for actions and costs involved. Further, the security recommendations should be approved and a budget allocated to put the measures into effect.

Physical Security	Logical Security	Archival Security
Prevent access to main data storage from unauthorized entrances.	Develop a policy for terminal access	Establish an audit trail for copies of data
Review the construction plans for the office buildings to ensure appropriate errand climate control	Create an access matrix by document types	Establish an offsite backup facility
Upgrade fire protection	Review protection of storage media	Create and organize metadata
Initiate document sign out and follow up procedures	Implement virus protection standards	Purchase storage media

Backup Basics

There are many ways one can unintentionally lose information on a computer; a power surge, lightning, floods, for instance. Sometimes the equipment just fails. Backup copies of files kept in a separate place is a good practice to ensure that the information is still there when something happens to the original files in the computer.

Before making backup copies, a checklist of files for backup should be made. This will help determine what files to back up, and also provide a reference list which will be useful in retrieving backed-up files.

Backup copies should be stored in external storage media, such as an external hard disk drive or flash drive, CDs, DVDs, or some other storage formats

The size of the files needed for the CLUP database will be relatively modest providing not so much raster data is included. Consequently, the recommendation is that the CLUP folder should be written to a DVD/CD on a regular interval (like once a month) and the backup be kept in a safe environment outside the office.

3.04.04 Legal Implications on Data Capturing and Storing

Intellectual Property Rights (IPR)

IPR is currently governed by Republic Act No. 8293, known as the **Intellectual Property Code of the Philippines (IPC)**, which was enacted and signed into law in 1997, and took effect on January 1, 1998.

It consists of Copyright and Related Rights, Trademarks and Service Marks, Geographical Indications, Industrial Designs, Patents, Lay-out Designs (Topographies) integrated circuits and Protection of undisclosed information.

Copyright and Related Rights

Copyright – is the protection extended to expressions and not to ideas, procedures, and methods of operation or mathematical concepts as such. These expressions may be in the forms of literary, scholarly, scientific and artistic works.

Related Rights – is the protection extended to derivative works, to include among others, dramatizations, translations, adaptations, abridgements, arrangements, and other alterations of literary or artistic works.

Programs / Software

Computer programs are protected by the IPC. The Code expressly protects computer programs as literary works. It also protects copyright in the manuals and packaging, which accompany the software.

Some notable points about the IPC law are:

- It protects databases and tables;
- It grants an exclusive rental right to the copyright owner;
- It recognizes and expands the protection of an author's moral rights, i.e., the right of an author to preserve the integrity of his work and name;
- It penalizes the possession of infringing software for the purpose of sale.

IPC allows reproduction of backup copies or adaptation of a computer programs without authorization of the author / copyright owner provided that the copy is necessary for:

a. the use of the computer program in conjunction with a computer for the purpose, and to the extent, for which the computer program has been obtained;

b. archival purposes, and, for the replacement of the lawfully owned copy of the program in the event that the lawfully obtained copy of the computer program is lost, destroyed or rendered unusable.

Such copy must be destroyed in the event that continued possession of the copy of the computer program ceases to be lawful.

Enforcement

The Intellectual Property Rights Code protects the owner's copyright, giving him the exclusive right to do certain things with the work, which in this case consists of the computer program, the manuals, and the packaging. These "exclusive rights" include:

- The right to copy the whole program or a substantial part of it.
- The right to adapt or translate the program.
- The right to rent the program to another person.

This means that one may only copy, adapt or rent a computer program if the copyright owner gives the permission to do this. This permission is given in the form of license. Every purchase of a legitimate copy of a computer program entitles one to receive a license agreement.