Prepared under the:

Local Government GIS Demonstration Grant

Supported by:

Local Government Records Management Improvement Fund
Local Government Records Services
State Archives and Records Administration

Project Team:

Erie County Water Authority
Mr. Paul Becker, Project Manager

National Center for Geographic Information and Analysis
State University of New York at Buffalo
Dr. Hugh Calkins, Project Director
Ms. Carmelle J. Côté
Ms. Christina Finneran

GIS Resource Group, Inc.
Mr. Graham Hayes, President
Mr. Thomas Murdoch, Vice-President

For More Information, Contact:

Local Government Technology Services
State Archives And Records Administration
9B38 Cultural Education Center
Albany, New York 12230
Phone: (518) 474-4372
Fax: (518) 473-4941

GIS DEVELOPMENT GUIDE

Volume 1

Table of Contents

MANAGER'S OVERVIEW

NEEDS ASSESSMENT

Needs Assessment cont'd

Figures

CONCEPTUAL DESIGN OF THE GIS

Part 1 Data Modeling

Part 2: Spatial Data Standards and Metadata Requirements

Figures

GIS DEVELOPMENT GUIDE: MANAGER'S OVERVIEW

1 INTRODUCTION

This guide is the first of a set of technical support documents to assist local governments indeveloping a GIS. The set of guides describes procedures and methods for planning the GIS,evaluating potential data sources, testing available hardware and software and planning for itsacquisition, building the GIS data base, developing GIS applications, and planning for the long termmaintenance of the GIS system and data base. These guides are intended to provide advice on howbest to accomplish the GIS development tasks for all levels of local government - from large,urbanized counties to small rural towns to special-purpose districts.

Realistically, large comprehensive GISs will be developed by the larger units of government (counties and cities) individually or, most likely as the leader in a cooperative multi-participant effort. These would involve the individual operating units within that government and/or the smaller units of local government within the common land area of the larger leading unit. Typically, we would expect to see county government taking the lead, but also covering the interest of all other governmental units within the county. Occasionally, there will be situations where smaller units of government (town, special purpose district, or limited purpose GIS application) may have to "go-it-alone" in developing the GIS. These guidelines have been written to mainly address the first case - a county leading a consortium or cooperative effort. Thus, we would expect the GIS development team of a county to be the primary user of these guidelines, in the sense of actually performing the tasks outlined in each document. ***However, this does not mean the other participants in a GIS should stop reading these guidelines at this point. It is critically important for all expected participants in a cooperative GIS venture to fully understand the development process. If a smaller unit of government is to reap the benefits of a county-level GIS, they must actively participate in the planning and development effort.***

The procedures are applicablefor use in first-time creation of a GIS, for restructuring an on-going GIS development project, andfor the review and further development of an existing GIS. The subject matter of the guidesidentifies the necessary tasks in a GIS development program, describes appropriate methods toaccomplish each task and, where applicable, provides examples and illustrations of documents orother products that result from each task.

The guidelines are designed for use by general-purpose local governments (city, county, town, or village), special purpose governments (utilities, school districts, etc.), and by those who provide assistanceto local governments (consultants, academic units, etc.). The guides address the technical stepsrequired to create a GIS, the management tasks required to ensure successfuldevelopment of the GIS, and the policy issues that should be considered for the effective use of theGIS.

The Role Of Management

Although GIS is often viewed as an arena for the technically sophisticated computerprofessional, the development of a successful government-based multi-participant GIS is very dependent on propermanagement participation and supervision. Normal, common-sense management practices are asnecessary in a GIS project as in any other major undertaking. In fact, our experience has shown thatthe recommended management actions may be the most critical aspect of the GIS developmentprocess. ***GIS development is a process of technological innovation and requires managementattention appropriate to this type of activity*** - active as opposed to passive management involvement in theproject. Historically, much of the disillusions and disappointment with GIS projects stems not froma failure of the technical components of the GIS but rather from a lack of understanding of theprocess of technology innovation and the lack of realistic expectations of all parties associated withthe project (GIS technicians, potential users, managers, and elected/appointed officials).

Applying The GIS Development Guides By Local Governments In New York State

The overall procedure contained in the GIS Development Guides is very comprehensive and canrequire considerable time, effort and dollars to complete. This raises the questions:

- Does all ofthis have to be done?
- What level of detail is appropriate?
- How can smaller governments,villages and towns, special purpose districts, or a single department in a larger jurisdiction, get through this process?

Does everything have to be done? . . level of detail?

Basically, yes. However, the steps in the GIS development process are frequently done in an iterative manner over an extended time period. Also, the steps are not completely independent of one another and so some back-and-forth does happen. It is often useful to make a "first-cut" run through the entire process, writing down what is already known and identifying the major questions that need to be answered. The person who will be managing the development process may be able to do this "first-cut" description in 1 to 2 days. This can be very helpful in getting a feel for the scope of the whole process and then can be used as a decision tool for continuing. The number of times the process is conducted, the amount of detail, and the resources needed to complete the study can be balanced in this way. If the intended implementation will be limited or small, the planning effort and documents can be sized accordingly. It is important, however, that each step be considered and completed at some level. The companion ***GIS Design software*** package that accompanies these guides provides a structure and makes it easy to record the information developed during the planning process - application descriptions, data model, data dictionary, metadata, logical database design, and record retention information.

How can smaller units of local government, such as villages and small towns complete a GIS Plan?

The best situation for a village, small town, or even a smaller, rural county is to be a partner with a larger unit of government, a county, regional agency or utility company that is conducting and/or leading a GIS planning exercise. Participating in a regional GIS cooperative, or joining an existing one, will provide access to GIS technical expertise and spatial data created by other agencies. Additionally, if one is a partner in a larger group, the activities directed toward the evaluation and selection of the GIS hardware and software may not need to be completed. One would simply use the same GIS system in use by the larger agency or group. Only the activities aimed at defining applications (uses) and identifying the needed data would need to be done by the smaller unit of government. In such a situation, the larger unit of government assumes the leadership role for the area-wide GIS and should have the technical expertise to assist the smaller unit. In situations where a larger effort does not exist, a village or town government may want to look at a GIS installation in a similar village or town elsewhere in the state. Given the similarities in local governments within the state, the adoption of the GIS plan of another unit is not unreasonable. That plan should be carefully reviewed by the intended participants in the GIS to ensure applicability. After modifying and validating the plan, a schedule for GIS hardware, software and data acquisition can be prepared consistent with available resources. If a good plan is prepared, there is no reason data acquisition (the most expensive part of a GIS) cannot be stretched over a long time period. Significant data already is available from state and federal agencies at reasonable costs. These data can form the initial GIS database, with locally generated data added later. A list of state and federal data sources is contained in the ***Survey of Available Data Guide***.

Content Of This Guide

This guide presents an overview of the GIS development process. This process is presented as asequence of steps conducted in a specific order. Each step is important in itself, but moreimportantly, information needed to complete subsequent steps is assembled and organized in eachprevious step. ***The underlying philosophy of the entire series of documents is to concentrate on the GIS data.*** As well as being the most expensive part of any GIS, the data must be collected, stored, maintained, and archived under an integrated set of activities in order to ensure continued availability and utility to the initial users as well as future users, including the general public. Defining and documenting data elements from their initial definition in the needs assessment through to proper archiving of the GIS database according to state requirements is the constant theme of these guidelines.

2 GEOGRAPHIC INFORMATION SYSTEMS: DEFINITIONS AND FEATURES

Basic Definition Of A Geographic Information System (GIS)

A geographic information system (GIS) may be defined as "...a computer-based informationsystem which attempts to capture, store, manipulate, analyze and display spatially referenced andassociated tabular attribute data, for solving complex research, planning and management problems"(Fischer and Nijkamp, 1992). GISs have taken advantage of rapid developments inmicroprocessor technology over the past several decades to address the special challenges of storingand analyzing spatial data. Geographers have referred to GISs as simultaneously providing "...thetelescope, the microscope, the computer and the Xerox machine" for geographic and regionalanalysis (Abler, 1987).

Unique Features Of A GIS - Why Planning Process Is Needed

GIS belongs to the class of computer systems that require the building of large databasesbefore they become useful. Unlike many micro-computer applications where a user can begin useafter the purchase of the hardware and software, ***the use of a GIS requires that large spatial databasesbe created, appropriate hardware and software be purchased, applications be developed, and allcomponents be installed, integrated and tested before users can begin to use the GIS.*** These tasks are large and complex, so large in fact,as to require substantial planning before any data, hardware or software is acquired. The focus ofthe GIS Development Guides is to describe the GIS planning process and to provide examples ofhow to accomplish the recommended planning tasks.

History Of Technology Innovations - GIS Is A Technology Innovation

It is useful to note that GIS is, at present, a technological innovation. The adoption of technologicalinnovations (i.e., the development of a GIS for a local government) is not always a straightforwardprocess, such as one might expect with the installation of something that is not new. Severalproblems are likely to occur such as:

- Staff not fully understanding the technology prior toextensive training
- Development time estimates differing from actual task times
- Greateruncertainty about costs
- A greater likelihood that programmatic changes will be needed duringthe development phases, etc.

The significant management point here is that these are normalconditions in the adoption of a new technology. Management needs to anticipate that such eventswill happen, and when they do, take appropriate management actions.

The adoption of computer technology by an organization either GIS or other applications, introduces fundamental change into the organization in its thinking about data. Prior informationtechnology allowed data to be collected and related to activities and projects individually. Organized stores of data were the exception rather than common practice. This led to duplicate datacollection and storage (as in different departments) and to the possibility of erroneous data existing in one ormore locations. One of the goals of computer systems and database development is to eliminateredundant data collection and storage. The principle is that data should be collected only once andthen accessed by all who need it. This not only reduces redundancy; it also allows for more accuratedata and a greater understanding of how the same data is used by multiple departments. The necessary condition for successful computer system and database development is fordifferent departments and agencies to cooperate in the development of the system. A databasebecomes an organization-wide resource and is created and managed according to a set of databaseprinciples.

3 ENTERPRISE-WIDE GIS: THE CORPORATE DATABASE

The role of a GIS in a local government setting is more than simply automating a few obvious tasksfor the sake of efficiency. A local government (or several cooperating governments) should ***viewthe GIS project as an opportunity to introduce fundamental change into the way its business isconducted.*** As with the adoption of management and executive information systems in the businessworld, the adoption of GIS effectively reorganizes the data and information the government collects,maintains and uses to conduct it affairs. This can, and arguably should, lead to major changes inthe institution, to improve both effectiveness and efficiency of operations.

A key factor in the success of computer system adoption in the business world is the concept of the"enterprise" or "corporate" database. As implied by the name, ***the corporate database is a single,organization-wide data resource.*** The advantages of the corporate database are first, that all usershave immediate and easy access to up-to-date information and, secondly that the construction of thedatabase is done in the most efficient manner possible. Typically, the corporate database eliminatesredundant collection and storage of information and the keeping of extra copies of data and extrareference lists by individual users. Here, we are recommending the use of ***corporate database concept*** to integrate GIS data for all units of local government participating in a cooperative GIS program.

An effective corporate database does require cooperation on the part of all users, both for thecollection and entry of data in the database and in developing applications in a shared data context. This may result in some individual applications or uses being less efficient, however the overallbenefits to the organization can easily outweigh these inefficiencies. Greater emphasis must,however, be placed on maintaining a high quality of data and services to users, mainly to offset theperceived loss of control that accompanies sharing an individual's data to another part of theorganization.

The corporate database concept can be used in the governmental situation, for either single units ofgovernment or between several governmental entities in the same region. The benefits associatedwith the corporate database can be achieved if governmental units are willing to cooperate andshare a multi-purpose regional GIS database. Such an arrangement has some technicalrequirements; however, ***establishing the corporate database is much more a question of policy , management cooperation and coordination.***

4 POLICY ISSUES IN GIS DEVELOPMENT

There are several policy issues that need to be addressed early in the GIS planning process:

GIS Project Management

Adequate management attention has already been mentioned in this document. As GIS is still anevolving new technology, the individuals involved (management, users, GIS staff) may have verydifferent expectations for the project, some based on general perceptions of computing, which mayor may not be correct. This, along with the long time period for developing the GIS, makes it veryimportant for substantial involvement of management in the project. Several factors associated withsuccessful GIS projects are:

- Emphasize advantages of GIS to individual users and entire organization
- Require high level of competency by all participants
- Ensure high level of management commitment from all management levels in the organization
- Require participation in team building and team participation within & between departments
- Ensure minimum data quality and access for all users
- Require development team to set realistic expectations
- Minimize time between user needs assessment and availability of useful products.
- Develop positive attitude toward change within organization
- Ensure level of technology is appropriate for intended uses
- Highly visible Pilot Project that is successful

Data Sharing

The ***sharing of data among government agencies is a virtual necessity*** for a successful, long-term GIS. Not even the most affluent jurisdictionswill be able to justify "going-their-own-way" and not taking advantage of what data areavailable from other sources or not sharing their database with other governmental units. This, then, raises several questions that must beconsidered during the planning of the GIS:

- What will be the source for each data item?
- How will sharing be arranged? . . purchase? . . license? . . other agreement?
- Who will own the data?
- How will new GIS data be integrated with existing data files (legacy systems)?
- Who will be responsible for updates to the data?
- How will the cost of the data (creation and maintenance) be allocated?
- Who will provide public access to the data?
- Who will be responsible for data archiving and retention? . . of the original? ..of copies?

These questions do not, at this time, have good answers. Currently, the Freedom ofInformation regulations require that all government data be made available to thepublic at minimal cost (cost of making a copy of the data). No distinction is made onthe basis of the format of the data (eye-readable or digital), the amount of data, or theintended use of data. Thus, the question of sharing the cost of a GIS database cannotbe addressed in general. If data can be obtained free from another agency, why enter into anagreement to pay for it? The answer is, of course, that the creating agency will not beable to sustain the GIS database under these circumstances. However, at this time , ***the set of state laws and regulations applicable to GIS data are not adequate to resolvecost issues and to facilitate regional data sharing cooperatives.*** New legislation willbe required. The New York State Temporary GIS Council did submitrecommendations on these issues to the Legislature in March 1996. Additionally, theNew York State Archives and Records Administration is currently in the process ofpreparing record management and retention schedules suitable for GIS data, both inindividual agencies and for shared databases. The New York State Office of Real Property Services has been designated as the GIS representative on the Governor's Task Force for Information Resource Management. One of the charges that has been given to the Task Force is to design a cohesive policy for the coordination of geographic information systems within New York building on the work of the Temporary GIS Council. Further information should be availablein late-1996 that should clarify the issues associated with arranging for data sharingamong governments.

5 MANAGEMENT ISSUES IN GIS DEVELOPMENT

Expected Benefits From The GIS

Local government need for, and use of, a GIS falls into several categories: ***maintaining publicrecords, responding to public inquiries for information, conducting studies and makingrecommendations to elected officials (decision-makers), and managing public facilities and services(utilities, garbage removal, transportation, etc.)***. The GIS tasks that meet these uses are:

- Providing regular maps
- Conducting spatial queries and displaying the results
- Conductingcomplex spatial analyses

Many of these tasks are already done by local government, althoughby manual means. The GIS is able to perform these tasks much more efficiently. Some of theanalytical tasks cannot be performed without a computer due to their size and complexity. In thesecases, the GIS improves local government effectiveness by providing better information to plannersand decision-makers.

Benefits from using a GIS fall into the two categories of: ***efficiency and effectiveness***. Existingmanual tasks done more efficiently by the GIS result in a substantial savings of staff time. In thelocal government context, the largest savings come from answering citizen inquiries of many types. Depending on the size of the government, savings using the ***query function*** of a GIS can rangefrom 2 person-years for a smaller town, to 5-8 person years for a large town, to 10 or more person-years for a large county. Estimates of potential time savings can be derived by measuring the timeto respond to a query manually and by GIS and multiplying the difference by the number ofexpected queries. This information is usually gathered during the Needs Assessment. Effectivenessbenefits are more difficult to estimate. The GIS may be used to accomplish several tasks that werenot previously done due to their size and complexity (e.g., flow analysis in water and sewer systems,traffic analysis, etc.). As these are essentially new tasks, a comparison between manual and GISmethods is not possible. While not measurable, the benefits from these applications can besubstantial. Generally categorized as ***better planning, better or more effective decision-making***,these applications support more effective investment of government resources in physicalinfrastructure where relatively small performance improvements can translate into large dollarsavings. GIS also provides an effective way to communicate the problem and solution to the general public and other interested parties

Resources Required To Develop A GIS

Developing a GIS involves investment in five areas: computer hardware, computer software,geographic data, procedures and trained staff. The acquisition of the computer hardware andsoftware are often incorrectly viewed as the most expensive activity in a GIS program. Research, someconducted at the National Center for Geographic Information and Analysis at SUNY-Buffalo, hasdemonstrated that ***developing the geographic database (which includes some of the procedure andstaff costs) can account for 60% to 80% of the GIS development costs.*** Continuing costs foroperation and maintenance are also dominated by the data costs. Coordination of GIS programs,particularly among several local government agencies, can minimize the cost of databaseconstruction and maintenance, and can provide for the greatest use of the database, which givesmaximum benefits from the investment.

Staffing Requirements For A GIS

Staffing for a GIS is a critical issue. In general, it is not easily feasible to directly expand the localgovernment staff positions to fill the GIS need. There are three areas where expertise is needed:

- Management of the GIS project ***(GIS project manager)***
- GIS database skills (usually called a ***database administrator***)
- Application development for database and users (***a GIS software analyst)***

Initial creation of the GIS database (digitizing) will require an appropriatelysized clerical staff, dependent on the amount of data to be converted. ***Alternatives to staff expansionare consultants and data conversion firms.*** GIS database conversion is a front-end staff need thatcan easily be contracted-out (good quality specifications need to be written for this task). If at allpossible, the three functions of ***GIS manager, GIS software analyst and GIS databaseadministrator*** should be fulfilled by staff personnel, either by hiring or by retraining existingprofessionals. When necessary, during the start-up phases of GIS development, the GIS analyst anddatabase administrator functions can be done under consultancy arrangements, ***PROVIDED THATA FULL-TIME GIS MANAGER IS AVAILABLE ON STAFF***.

The second need is for training of users in general computing, database principles, and GIS use. These topics are covered in training courses offered by most GIS vendors, and after the GISsoftware has been selected, they are the best source for user training.

Management Decision Points in the GIS Development Program

The "decision" to develop a GIS is made incrementally. The information needed to determine thefeasibility and desirability of developing a GIS is not available until several of the planning stepshave been completed. The key decision points are:

- Decision to investigate GIS for the organization - the initial decision to begin the process. This is an initial feasibility decision and is based on the likelihood that a GIS will be useful and effective. It is fairly important to identify the major participants at this point - both departments within agencies and the group of agencies, particularly key agencies, the agencies who represent a majority of the uses and who will contribute most of the data.

- Decision to proceed with detailed planning and design of the database - at this time, theapplications, data required, and sources of the data have been identified. Applications can be prioritized and scheduled and the benefits stream determined. Also, applications to be tested during the pilot study and the specific questions to be answered by the pilot study will have been determined. A preliminary decision will need to be made as to which GIS software will be used to conduct the pilot study.

- Decision to acquire the GIS hardware and software - this decision follows the preparation of the detailed database plan, the pilot study and, if conducted, the benchmark test. This is the first point in the development process where the costs of the GIS can reasonably be estimated, the schedule for data conversion developed, and targets for users to begin use determined.

6 GEOGRAPHIC INFORMATION SYSTEMS: THE DEVELOPMENT CYCLE

Developing a GIS is more than simply buying the appropriate GIS hardware and software. ***Thesingle most demanding part of the GIS development process is building the database.*** This tasktakes the longest time, costs the most money, and requires the most effort in terms of planning andmanagement. Therefore the GIS development cycle presented here emphasizes database planning. Most local governments will acquire the GIS hardware and software from a GIS vendor. Choosingthe right GIS for a particular local government involves matching the GIS needs to thefunctionality of the commercial GIS. For many agencies, especially smaller local governments, choosing a GIS will require help from larger, more experienced agencies, knowledgeable university persons and from qualified consultants. By completing selected tasks outlined in these guidelines local governments can prepare themselves to effectively interact and use expertise from these other groups.

The GIS development cycle ***starts with the needs assessment where the GIS functions andthe geographic data needed are identified.*** This information is obtained through interviewingpotential GIS users. Subsequently, surveys of available hardware, software and data are conductedand, based in the information obtained, detailed GIS development plans are formulated.

It is important to involve potential users in all stages of GIS development. They benefit from thisinvolvement in several ways:

- Describing their needs to the GIS analysts
- Learning what the GISwill be capable of accomplishing for them
- Understanding the nature of the GIS developmentcycle - the time involved and the costs.

Potential users need to understand that there may besignificant time lags between the first steps of Needs Assessment and the time when the GIS canactually be used. Mostly, this is due to the size of the database building task, which can take up toseveral years in a large jurisdiction.

In addition to understanding that database development takes substantial time, users and managersneed to appreciate that GIS is a new technology and its adoption often involves some uncertaintythat can cause time delays, on-going restructuring the development program, and the need to resolveunforeseen problems. This set of guideline documents describes the GIS development process in away that will minimize problems, time delays, cost overruns, etc.; however, the occurrence of thesesituations cannot be completely avoided. ***The GIS project team and management simply have to beaware that some unforeseen events will happen.*** GIS development must be viewed as a process rather than a distinct project.

Estimating and planning for the cost of the GIS is a somewhat difficult task. First, it is necessaryto recognize that the GIS database will likely be the single most costly item - if a local governmentdevelops all of the data itself from maps, etc., this cost can be as much as 70 - 80 % of the totalsystem cost. Thus, ***acquiring digital data from other GIS systems, government sources or the privatesector can be very cost effective.*** Participating in, or organizing a regional data sharing cooperativeor district, can also lead to reduced data costs. When planning for the GIS database, ***long term datamaintenance and retention costs must be estimated as well as the initial start-up costs.*** Cooperation betweenagencies with similar data needs may provide the most effective way to achieve long-term datamaintenance, retention, and archiving.

7 TASKS FOR GIS DEVELOPMENT AND USE

The GIS development cycle is a set of eleven steps starting with the needs assessment and endingwith on-going use and maintenance of the GIS system. These steps are presented here as a logicalprogression with each step being completed prior to the initiation of the next step. While this viewis logical, it is not the way the world always works. Some of the activities in the process mayhappen concurrently, may be approached in a iterative manner, or may need to be restructureddepending on the size and character of the local government conducting the study and the resourcesavailable to plan for the GIS. ***The GIS development cycle is based on the philosophy that one firstdecides what the GIS should do and then as a second activity decides on how the GIS will accomplisheach task.*** Under this philosophy, the needs are described first, available resources are inventoriedsecond (data, hardware, software, staff, financial resources, etc.), preliminary designs are createdand tested as a third major set of activities, and lastly the GIS hardware and software are acquiredand the database is built.

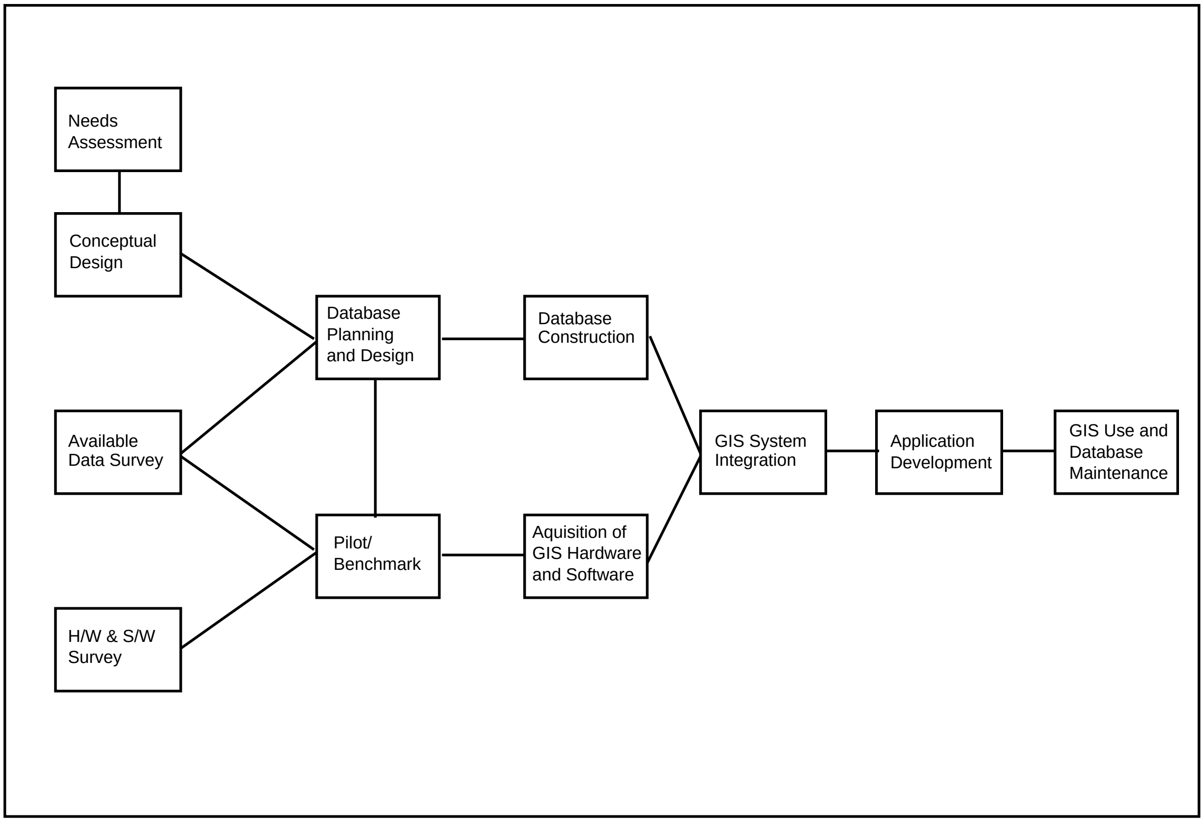


Figure 1 - GIS Development Process

Figure 1 shows the GIS development cycle, which is described in terms of 11 major activities. Prior to initiating these studies, the responsible staff in local governments should attend introductory GIS seminars and workshops, GIS conferences, and meetings of specific GIS users' groups, to obtain a broad overview of what GIS is and how others are using these systems.

The 11 steps of the GIS development cycle are:

1. Needs Assessment
2. Conceptual Design of the GIS
3. Survey of Available Data
4. Survey of GIS Hardware and Software
5. Detailed Database Planning and Design
6. Database Construction
7. Pilot Study/Benchmark Test
8. Acquisition of GIS Hardware and Software
9. GIS System Integration
10. GIS Application Development
11. GIS Use and Maintenance

These tasks are one way of dividing up the entire set of activities that must be accomplished to builda successful GIS. While there are other ways of expressing and organizing these activities, thisparticular structure has been chosen because it emphasizes data development - data definition, datamodeling, data documentation, data capture and storage, and data maintenance and retention.

Theimportant point to be made here is not the order or structure of the tasks, but rather that, one wayor another, ***all of these tasks must be completed to have a successful GIS***.

In some situations, different methods may be more appropriate than those presented in these guides,or a different level of detail may fit the particular situation of a unit of local government. Nomatter how simple or complex a given GIS environment is, all of the above tasks should becompleted ***at an appropriate level of detail***. In the specific guides of this set, examples of differentlevels of detail will be provided.

The starting point is the needs assessment. It is assumed that the local government has decided thata GIS may be justified and it is reasonable to expend the resources to further study the problem. Afinal assessment of the costs and benefits will not be made until several tasks have been completedand the nature and size of the resulting GIS can be estimated. In the process presented here, thisfinal feasibility assessment is made as part of the detailed database planning and design activity.

Each of the major portions of the development cycle identified and briefly described below is furtherdescribed in a subsequent guideline document.

Needs Assessment

The GIS needs assessment is designed to produce two critical pieces of information:

- Thelist of GIS functions that will be needed
- A master list of geographic data.

These twoinformation sets are extracted from a set of GIS application descriptions, a list of important data,and a description of management processes. Standard forms are used to document the results ofuser interviews. The information gained in the needs assessment activity goes directly intothe Conceptual GIS Design activity.

Conceptual Design of the GIS System

The conceptual design of the GIS system is primarily an exercise in database design. Itincludes formal modeling (preparation of a data model) of the intended GIS database and the initialstages of the database planning activity. ***Database planning is the single most important activity inGIS development.*** It begins with the identification of the needed data and goes on to cover severalother activities collectively termed the ***data life cycle*** - identification of data in the needs assessment,inclusion of the data in the data model, creation of the metadata, collection and entry of the data into thedatabase, updating and maintenance, and, finally, retention according to the appropriate recordretention schedule (Figure 2). A complete data plan facilitates all phases of data collection,maintenance and retention and as everything is considered in advance, data issues do not becomemajor problems that must be addressed after the fact with considerable difficulty and aggravation. Theproduct of the conceptual design activity is a ***data model*** which rigorously defines the GIS databaseand supports the detailed database planning activity.

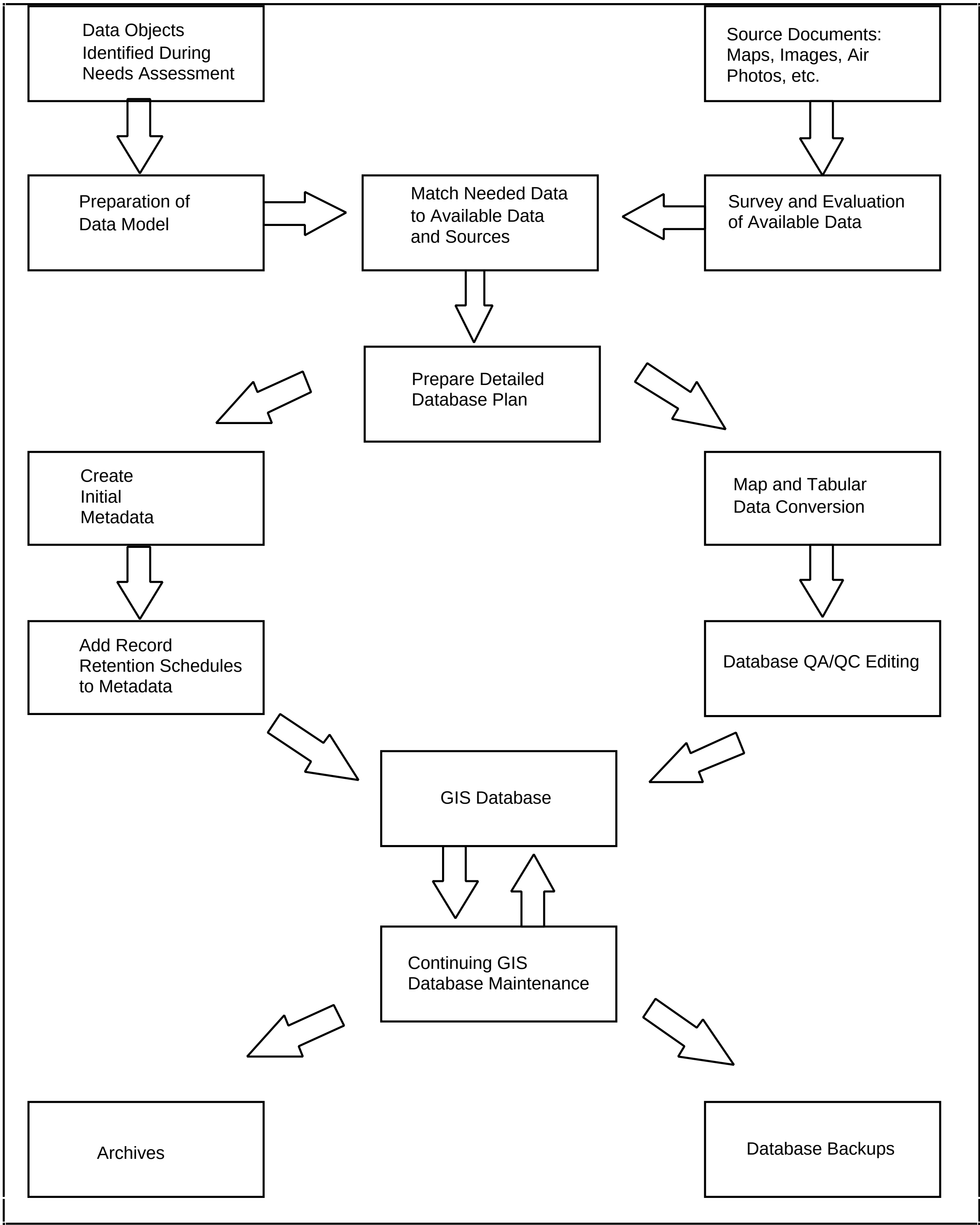


Figure 2 - Life Cycle of a GIS Database

The conceptual design of the GIS also includes identification of the basic GIS architecture (type ofhardware and GIS software), estimates of usage (derived from the Needs Assessment), and scopingthe size of the GIS system. All of this is done with reference to the existing data processingenvironments (legacy systems) that must interface with the GIS. This guideline also includes a section on metadata and data standards.

Survey Of Available Data

A survey of available data can commence once needed data have been identified in the NeedsAssessment. This task will ***inventory and document*** mapped, tabular and digital data within thelocal government as well as data available from other sources, such as federal, state, or other localgovernments and private sector organizations. The entries in this inventory may include other GISsystems within the local area from which some of the needed data may be obtained. If there existsan organized data sharing cooperative or other mechanism for government data sharing, it shouldbe investigated at this time. There also exists the possibility that one or more of the commercial GISdatabase developers may be able to supply some of the needed data and should therefore beinvestigated. The documentation prepared at this point will be sufficient to evaluate each potentialdata source for use in the GIS. Information collected at this point will also form part of themetadata for the resulting GIS database.

Survey Of Available GIS Hardware And Software

Almost all local government GIS programs will rely on commercially available GIS software. Asa result, a survey of the available GIS systems needs to be conducted. During this activity, the GISfunctionality of each commercial GIS system can be documented for later evaluation.

Detailed Database Design And Planning

The detailed database planning and design task includes the following activities: developing alogical or physical database design based on the data model prepared earlier, evaluating the potentialdata sources, estimating the quantities of geographic data, estimating the cost of building the GISdatabase and preparing the data conversion plan. Concurrent with the detailed planning for thedatabase, pilot studies and/or benchmark testing that are desired can be executed. Informationgained from these studies and tests will be needed to estimate the size of the equipment (disk space,main memory etc.) and to determine how much application development will be necessary. Subsequently, plans for staffing, staff training, equipment acquisition and installation, and usertraining must be completed. After the preparation of all these plans, the entire cost of the GIS willbe known and the final feasibility assessment can be made.

Pilot Study And Benchmark Tests

Pilot studies and benchmark tests are intended to demonstrate the functionality of the GIS software -simply put, what the commercial GIS from the vendor can do. These tests are useful to demonstrate to potential users andmanagement what the GIS will do for them. Also, performance data of the GIS system can bedetermined.

GIS Database Construction

Database construction (sometimes referred to as "database conversion") ***is the process of buildingthe digital database from the source data - maps and tabular files.*** This process would have beenplanned during the previous activity and the main emphasis here is management of the activity andquality assurance/quality control of the converted data. The conversion process is often "contracted-out" and involves large quantities of source maps and documents. Close and effective managementis the critical factor in successful data conversion.

GIS System Integration

Unlike many other computer applications, a GIS is not a "plug and play" type system. The severalcomponents of a GIS must be acquired according to well documented specifications. The databasemust be created in a ***careful and organized manner***. Once all the individual components have beenacquired, they must be integrated and tested. Users must be introduced to the system, trained asnecessary, and provided with adequate assistance to begin use of the GIS. Parts of the GIS which may appear to work fine individually may not work properly when puttogether. The GIS system staff must resolve all the problems before users can access the GIS.

GIS Application Development

"Application" is a general term covering all things that "go on" in a GIS. First, there are "databaseapplications." These are all the functions needed to create, edit, build, and maintain the database,and are usually carried out by the GIS systems staff. Some users may have responsibility forupdating selected parts of the GIS database, however the entire database should be under the controlof a ***"database administrator."*** Other applications are termed "user applications." ContemporaryGISs provide many simple applications as part of the initial software package (e.g., map display,query, etc.). ***More complex applications,*** or ones unique to a particular user, ***must be developedusing a macro-programming language***. Most GISs have a macro-programming language for thispurpose (e.g., Arc Macro Language (AML) in ARC/INFO™. and Avenue in ArcView™). Theapplications needing development by the GIS systems staff will have been described during theNeeds Assessment on the GIS Application forms.

GIS System Use And Maintenance

After having described the rather large task of creating a GIS, we can now say that use andmaintenance of the GIS and its database will likely require as much attention as was needed to initiallybuild it. Most GIS databases are very dynamic, changing almost daily, and users will immediatelythink of additional applications that they would like to have developed. Formal procedures forall the maintenance and updating activities need to be created and followed by the GIS system staffand by all users to ensure continued successful operation of the GIS.

8 SUMMARY

This document has presented an overview of the GIS development process, with anemphasis on data and database issues. All of the tasks and issues identified in thisdocument will be described in detail in the remaining eleven guidelines of this series. Theprocedures are presented as "guides," and not as a "cookbook recipe" which must berigorously followed. Each of the major tasks in the GIS development process and theinformation generated within the task ***should be addressed*** in any specific project. Themethods and forms used in this series can be used, or alternatives can be developed,as appropriate to the situation. The one matter to always keep in mind is that ***the GISplan is a document to communicate user needs to a GIS analyst.*** The componentsof the plan must contain:

- Descriptions of applications that are understandable to theuser
- A logical translation of user requirements to system specifications
- Detailed specification suitable for system development

Following the recommendations in these guidelines cannot, unfortunately, guarantee success. Many of the factors, outside the control of the GIS development team, will affect the ultimate success of the GIS - success being defined as use of the GIS by satisfied users. However, the authors of these guidelines believe that attempting to develop a GIS without following these, or similar procedures, substantially raises the probability of an unsuccessful GIS project - either one that is not useful or one that substantially exceeds both cost and development time estimates.

Finally, although presented here as an independent activity, GIS development must recognize and interface with other computer systems in local government, such as E911, police and fire dispatch, facilities management systems, etc. . The GIS must not be viewed as independent of the other systems, but integrated with them, no matter how difficult, to form a true ***corporate database for local government.***

REFERENCES

1) Fischer, Manfred M. and Nijkamp, Peter, "Geographic Information System, Spatial Modeling, and Policy Evaluation," Berlin & New York: Springer-Verlag, 1993, pg 42.

2) Abler, R.F., 1987, "The National Science Foundation National Center for Geographic Information and Analysis" International Journal of Geographical Information Systems, 1, no. 4, 303-326.

SUGGESTED READINGS

1. Antenucci, John C., et.al., ***Geographic Information Systems: A Guide to the Technology,*** New York: Van Nostrand Reinhold, 1991 (ISBN 0-442-00756-6)

2. Aronoff, Stan, ***Geographic Information Systems: A Management Perspective,*** Ottawa: WDL Publications, 1989 (ISBN 0-921804-00-8)

3. Burrough, P.A., ***Principles of Geographical Information Systems for Land Resources Assessment,*** Oxford: Oxford University Press, 19865. (ISBN 0-19-854563-0); ISBN 0-19-854592-4 paperback).

4. Huxhold, William E., ***An Introduction to Urban Geographic Information Systems***, Oxford: Oxford University Press, 1991 (ISBN 0-19-506534-4)

5. Korte, George B., ***A Practioner's Guide: The GIS Book***, Sante Fe: OnWord Press, 1992 (ISBM 0-934605-73-4)

6. Laurini, Robert and Derek Thompson, ***Fundamentals of Spatial Information Systems***, London: Academic Press Limited (ISBN: 0-12-438380-7)

7. Montgomery, Glenn E., and Harold C. Schuck, ***GIS Data Conversion Handbook***, Fort Collins: GIS World, Inc. (ISBN 0-9625063-4-6)

GIS INFORMATION SOURCES

Scholarly journals

There are a number of scholarly journals that deal with GIS. These are published on an on-going basis.

Cartographica - Contact: Canadian Cartographic Association
Cartography and Geographic Information Systems - Contact: American Cartographic Association
International Journal of Geographical Information Systems - Contact: Keith Clark at CUNY Hunter College, New York City
URISA Journal - Contact: Urban and Regional Information Systems Association

Trade magazines
There are a number of trade magazines that are focused on GIS. They are:

GIS World
GIS World Inc.
155 E. Boardwalk Drive
Suite 250, Fort Collins, CO 80525.
Phone: 303-223-4848
Fax: 303-223-5700
Internet: info@gisworld.com

Business Geographics
GIS World, Inc.
155 E. Boardwalk Drive, Suite 250
Fort Collins, CO 80525.
Phone: 303-223-4848
Fax: 303-223-5700.
Internet: info@gisworld.com.

Geo Info Systems
Advanstar Communications
859 Williamette St.
Eugene, OR., 97401-6806
Phone: 541-343-1200
Fax: 541-344-3514
Internet:geoinfomag@aol.com
WWW site:http://www.advanstar.com/geo/gis

GPS World
Advanstar Communications
859 Williamette St.
Eugene, OR., 97401-6806
Phone: 541-343-1200
Fax: 541-344-3514
Internet:geoinfomag@aol.com
WWWsite:http://www.advanstar.com/geo/gis

Conference Proceedings

American Congress on Surveying and Mapping (ACSM)
5410 Grosvenor Lane
Bethesda, MD, 20814
Phone: 301-493-0200
Fax: 301-493-8245

American Society for Photogrammetry and Remote Sensing (ASPRS) & (GIS/LIS)
5410 Grosvenor Lane
Bethesda, MD, 20814
Phone: 301-493-0290
Fax: 301-493-0208

Association of American Geographers (AAG)
1710 Sixteenth St. N.W.
Washington D.C., 20009-3198
Phone: 202-234-1450
Fax: 202-234-2744

Automated Mapping/Facility Management International (AM/FM International)
14456 East Evans Ave.
Aurora, CO, 80014
Phone: 303-337-0513
Fax: 303-337-1001

Canadian Association of Geographers (CAG)
Burnside Hall, McGill University
Rue Sherbrooke St. W
Montreal, Quebec H3A 2K6
Phone: 514-398-4946
Fax: 514-398-7437

Canadian Institute of Geomatics (CIG)
206-1750 rue Courtwood Crescent
Ottawa, Ontario, K2C 2B5
Phone: 613-224-9851
Fax: 613-224-9577

Urban And Regional Information Systems Association (URISA)
900 Second St. N.E., Suite 304
Washington, D.C. 20002
Phone: 202-289-1685
Fax: 202-842-1850

Glossary

Accuracy - Degree of conformity with a standard, or the degree of correctness attained in a measurement. Accuracy relates to the quality of a result. If accuracy is relative, the position of a point is defined in relation to another point. It is less expensive to build a GIS in the context of relative accuracy. If accuracy is absolute, the position of a point is defined by a coordinate system. Building a GIS in the context of absolute accuracy requires use of the global positioning system.

Accuracy Requirement - statement of how precise the desired results must be to support a particular application.

Adjoining Sheets - Maps that are adjacent to one another at the corners and on one or more sides.

Aerial - Relating to the air atmosphere, being applicable in a descriptive sense to anything in space above the ground and within the atmosphere.

Aerial Photography - The method of taking photographs from an aerial platform (aircraft). (1.) Vertical photography, some times called orthophotography (see entry) is used for photogrammetric mapping and requires a high degree of accuracy. (2.) Oblique photography is used for general information, sometimes to verify certain attributes, but does not provide accurate measurements for photogrammetric mapping.

Aerial Survey - A survey utilizing aerial photography or from remote sensing technology using other bands of the electromagnetic spectrum such as infrared, gamma or ultraviolet.

Algorithm - A set of instructions; ordered mathematical steps for solving a problem like the instructions in a computer program.

Alignment - Relates to survey data transposed to maps. The correct position of a line or feature in relation to other lines or features. Also the correct placement of points along a straight line.

Alphanumeric - A combination of alphabetic letters, numbers and or special characters. A mailing address is an alphanumeric listing.

Analog Data - Data represented in a continuous form, not readable by a computer.

Area - level of spatial measurement referring to a two-dimensional defined space; for example, a polygon on the earth as projected onto a horizontal plane.

Attribute - 1. A numeric, text, or image data field in a relational data base table that describes a spatial feature such as a point, line, node, area or cell. 2. A characteristic of a geographic feature described by numbers or characters, typically stored in tabular format, and linked to the feature by an identifier. For example, attributes of a well (represented by a point) might include depth, pump type, location, and gallons per minute.
AM/FM - Automated mapping/facilities management. A GIS designed primarily for engineering and utility purposes, AM/FM is a system that manages databases related to spatially distributed facilities.

Base Data - set of information that provides a baseline orientation for another layer of primary focus, e.g., roads, streams, and other data typically found on USGS topographic and/or planimetric maps.

Base Line - A surveyed line established with more than usual care upon which surveys are based.

Base Map - A map showing planimetric, topographic, geological, political, and/or cadastral information that may appear in many different types of maps. The base map information is drawn with other types of changing thematic information. Base map information may be as simple as major political boundaries, major hydrographic data, or major roads. The changing thematic information may be bus routes, population distribution, or caribou migration routes.

Base Station - a GPS receiver on a known location that may broadcast and/or collect correction information for GPS receivers on unknown locations.

Bench Mark - A relatively permanent point whose elevation above or below an adopted datum is known.

Beta Test - Hardware or software testing performed by users in a normal operating environment; follows alpha testing, which is generally done in the developer's facility.

Bezier - (computer graphics) A curve generated by a mathematical formula in CAD (see entry) programs that maintains continuity with other Bezier curves.

Binary - The fundamental principal behind digital computers. Binary means two, computer input is converted into binary numbers made up of O and 1 (see bit).

BIT: (computers) a binary digit with a value of either 1 or 0.

Block (Tax) - A group of municipal tax lots that can be isolated from other parcels by a boundary, usually a roadway, waterway or properly labeled lot line.

Boundary Line - A line along which two areas meet. In specific cases, the word "boundary" is sometimes omitted, as in "state line", sometimes the word "line" is omitted, as in "international boundary", "county boundary", etc. The term "boundary line" is usually applied to boundaries between political territories, as "state boundary line", between two states. A boundary line between privately owned parcels of land is termed a property line by preference, or if a line of the United States public land surveys, is given the particular designation of that survey system, as section line, township line, etc.

BPS - Bits per second, the speed of data transfer.

Buffer A zone of a given distance around a physical entity such as a point, line, or polygon.

CAD/CADD - (Computers) Computer-Aided Design/ Computer-Aided Design and Drafting. Any system for Computer-Aided rather than manual drafting and design. Displays data spatially. on a predefined coordinate grid system, allowing data from different sources to be connected and referenced by location. Speeds conventional map development process by 1. permitting replication of shapes, floor plans, etc. from an electric library rather than requiring every component to be drawn from scratch. 2. Plotters and terminal screens are faster and more accurate than manual drafting. 3. Portions of drawings can be edited, enlarged, etc. quickly. 4. Related information can be stored in files and added to drawings in layers.

CAD - (Communication) Computer-Aided Dispatching. Used with emergency vehicles, CAD can be very sophisticated. Online maps of a city can display emergency vehicles as moving dots on the map, their status (enroute to an emergency, awaiting a call, call completed, returning to base, etc.) indicated by different colors. (The acronym for computer-aided dispatch is sometimes confused with computer-aided design.)

Cadastre - a record of interests in land, encompassing both the nature and extent of interests. Generally, this means maps and other descriptions of land parcels as well as the identification of who owns certain legal rights to the land (such as ownership, liens, easements, mortgages, and other legal interests). Cadastral information often includes other descriptive information about land parcels.

Cadastral - Relating to the value, extent and ownership of land for tax purposes. Cadastral maps describe and record ownership. Also called property map.

Cadastral Survey - A survey relating to land boundaries and subdivisions, made to create units suitable for transfer or to define the limitations to title. Derived from "cadastre", and meaning register of the real property of a political subdivision with details of area, ownership, and value. The term cadastral survey is now used to designate the surveys for the identification and resurveys for the restoration of property lines; the term can also be applied properly to corresponding surveys outside the public lands, although such surveys are usually termed land surveys through preference. See also boundary, survey.

Cartographic (Planimetric) Features - Objects like trees or buildings shown on a map or chart.

Cartography - The technology of mapping or charting features of Earth's topography.

Centroid - The "center of gravity" or mathematically exact center of an irregular shaped polygon; often given as an x, y coordinate of a parcel of land.

Clearinghouse - a physical repository structure used to accumulate and disseminate digital data and information concerning that data. In the GIS context a clearinghouse can contain all or a portion of spatial, metadata and informational data.

Client - A software application that works on your behalf to extract some service from a server somewhere on the network. Basic idea, think of your telephone as a client and the telephone company as a server.

COGO - Acronym for Coordinate Geometry achieved via a computer program.

Computer-aided Design or Drafting (CAD) - A group of computer software packages for creating graphic documents.

Control Point - A point in a network, identifiable in data or a photograph, with a given horizontal position and a known surface elevation. It is correlated with data in data set or photograph.

Contour - An imaginary outline of points on the ground which are at the same altitude relative to mean sea level.

Contour Line - A line on a map or chart that connects to points which are at the same elevation.

Contour Map - A map that defines topography (hypsography) by interpreting contour lines as relief.

Control - Also called ground control. A system of survey marks or objects called control points that have established positions and/or elevations verified by ground survey. The marks, or control points, serve as a reference correlating other data such as contour lines (see entry) determined from aerial surveys.

Conversion - 1. The translation of data from one format to another (e.g., TIGER to DXF; a map to digital files).S 2. Data conversion when transferring data from one system to another (E.g., SUN to IBM).s

Coordinate - The position of point is space in respect to a Cartesian coordinate system (x, y and/or z values). In GIS, a coordinate often represents locations on the earth's surface relative to other locations.

Coordinate System - The system used to measure horizontal and vertical distances on a planimetric map. In a GIS, it is the system whose units and characteristics are defined by a map projection. A common coordinate system is used to spatially register geographic data for the same area. See map projection

CRT - Cathode Ray Tube. A computer screen or monitor.

CTG - Center for Technology in Government

Data Capture - series of operations required to encode data in a computer-readable digital form (digitizing, scanning, etc.)

Data Dictionary - description of the information contained in a data base, e.g., format, definition, structure, and usage. It typically describes and defines the data elements of the data base and their interrelationships within the larger context of the data base.

Data Element - specific item of information appearing in a set of data, e.g. well site locations.

Data Model 1. A generalized, user-defined view of the data related to applications. 2. A formal method for arranging data to mimic the behavior of the real world entities they represent. Fully developed data models describe data types, integrity rules for the data types, and operations on the data types. Some data models are triangulated irregular networks, images, and georelational or relational models for tabular data.

Data Quality - refers to the degree of excellence exhibited by the data in relation to the portrayal of the actual phenomena

Data Sets - a collection of values that all pertain to a single subject.

Data Standardization - the process of achieving agreement on data definitions, representation, and structures to which all data layers and elements in an organization must conform.

Data Structure - organization of data, particularly the reference linkages among data elements.

Database -usually a computerized file or series of files of information, maps, diagrams, listings, location records, abstracts, or references on a particular subject or subjects organized by data sets and governed by a scheme of organization. "Hierarchical" and relational" define two popular structural schemes in use in a GIS. For example, a GIS database includes data about the spatial location and shape of geographic entities as well as their attributes.

Database Management System (DBMS) - 1. The software for managing and manipulating the whole GIS including the graphic and tabular data. 2. Often used to describe the software for managing (e.g., input, verify, store, retrieve, query, and manipulate) the tabular information. Many GlSs use a DBMS made by another software vendor, and the GIS interfaces with that software.

Datum - a mathematical reference framework for geodetic coordinates defined by the latitude and longitude of an initial point, the azimuth of a line from this point, and the parameters of the
ellipsoid upon which the initial point is located.

DEC - Department of Environmental Conservation

Differential Correction - the method (usually done through post processing) of using two GPS receivers, one on a known location and one on an unknown location, using information from the one on the known location to correct the position of the unknown location.

Digital Accuracy - refers to the accuracy of digital spatial data capture.

Digital Elevation Model (DEM) - a file with terrain elevations recorded at the intersections of a fine grid and organized by quadrangle to be the digital equivalent of the elevation data on a topographic base map.

Digital Data - a form of representation in which distinct objects, or digits, are used to stand for something in the real world--temperature or time--so that counting and other operations can be performed precisely. Data represented digitally can be manipulated to produce a calculation, a sort, or some other computation. In digital electronic computers, two electrical states correspond to the Is and Os of binary numbers, which are manipulated by computer programs.

Digital Exchange Format (DXF) 1. ASCII text files defined by Autodesk, Inc. (Sausalito, CA) at first for CAD, now showing up in third-party GIS software . 5 2. An intermediate file format for exchanging data from one software package to another, neither of which has a direct translation for the other but where both can read and convert DXF data files into their format. This often saves time and preserves accuracy of the data by not reautomating the original.

Digital Line Graph (DLG) 1. In reference to data, the geographic and tabular data files obtained from the USGS for exchange of cartographic and associated tabular data files. Many non-DLG data may be formatted in DLG format. 2. In reference to data, the formal standards developed and published by the USGS for exchange of cartographic and associated tabular data files. Many non-DLG data may be formatted in DLG format.

Digital Map - A machine-readable representation of a geographic phenomenon stored for display or analysis by a digital computer; contrast with analog map.

Digital Orthophoto - A geographically correct digital image with the same accuracy as a vector digital map, but preserving the information content of the original photography.

Digital Orthophoto Quarter-Quad (DOQ) - a 3.75 minute square distortion free image of the surface of the earth. The imagery has been geographically and photographically rectified to remove all distortion, and meet requirements of the USGS.

Digital Terrain Model (DTM) - A computer graphics software technique for converting point elevation data into a terrain model displaced as a contour map, sometimes as a three-dimensional "hill and valley" grid view of the ground surface.

Digitize - A means of converting or encoding map data that are represented in analog form into digital information of x and y coordinates.

Digitized Terrain Data - Transposed elevation information from maps or photographs to X-Y-Z digital coordinates for storage on magnetic media.

Digitizer - A device used to capture planar coordinate data, usually as x and y coordinates, from existing analogmaps for digital use within a computerized program such as a GIS; Also called a digitizing table.

Digitizing - refers to the process of manually converting an analog image or map or other graphic overlay into numerical format for use by a computer with the use of a digitizing table or tablet and tracing the input data with a cursor (see also scanning).

DIME - Dual Independent Map Encoding Provides vector data such as streets to census data addresses. Superseded by Topologically Integrated Geographic Encoding and Referencing (see TIGER).

DIME File - A geographic base file produced by the U.S. Census Bureau with Dual Independent Map Encoding. Now being superseded by TIGER files (see below).

DLG - See Digital Line Graph

DOB - Division of the Budget

DOQ - See Digital Orthophoto Quarter-quad

DOT - Department of Transportation

DTF - Department of Taxation and Finance

Edge Match - An editing procedure to ensure that all features crossing adjacent map sheets have the same edge locations, attribute descriptions, and feature classes.

Federal Information Processing Standards (FIPS) - official source within the federal government for information processing standards. They were developed by the Institute for Computer Sciences and Technology, at the National Institute of Standards and Technology (NIST), formerly the National Bureau of Standards.

Federal Geographic Data Committee (FGDC) - established by the Federal Office of Management and Budget, is responsible for the coordination of development, use, sharing, and dissemination of surveying, mapping, and related spatial data.

Fifth Generation Computer - A computer designed for applications of artificial intelligence (Al). Some elements of spatial data management, especially the CADD output side, are beginning to integrate Al computing.

FOIL - Freedom of Information Law

Format - 1. The pattern in which data are systematically arranged for use on a computer. 2. A file format is the specific design of how information is organized in the file. For example, DLG, DEM, and TIGER are geographic data sets in particular formats that are available for many parts of the United States 6

File Transfer Protocol (FTP) - a standard protocol that defines how to transfer files from one computer to another.

Fortran - A high-level programming language and compiler originally designed to express math formulas. Developed in 1954 by IBM it is still the most widely used language for scientific and engineering programming.

GBF/DIME - See Geographic base file/dual independent map encoding

Geocode - The process of identifying a location as one or more x, y coordinates from another location description such as an address. For example, an address for a student can be matched against a TIGER street network to locate the student's home.

Geodetic Monumentation - a permanent structure that marks the location of a point taking into account the earth's curvature.

Geographic - Pertains to the study of the Earth and the locations of living things, humans and their effects.

Geographic Base File/dual Independent Map Encoding (GBF/DIME) - A data exchange format developed by the US Census Bureau to convey information about block-face/street address ranges related to 1980 census tracts. These files provide a schematic map of a city's streets, address ranges, and geostatistical codes relating to the Census Bureau's tabular statistical data. See also TIGER, created for the 1990 census.

Geographic Database - Efficiently stored and organized spatial data and possibly related descriptive data.

Geographic Information Retrieval and Analysis (GIRAS) - Data files from the US Geological survey. GIRAS files contain information for areas in the continental United States, including attributes for land use, land cover, political units, hydrologic units, census and county subdivisions, federal land ownership, and state land ownership. These data sets are available to the public in both analog and digital form.

Geographic Information System (GIS) - An organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information. Certain complex spatial operations are possible with a GIS that would be very difficult, time-consuming, or impractical otherwise.

Geographic Object - A user-defined geographic phenomenon that can be modeled or represented using geographic data sets. Examples include streets, sewer lines, manhole covers, accidents, lot lines, and parcels.

Geographical Resource Analysis Support System (GRASS) - 1. A public-domain raster GIS modeling product of the US Army Corps of Engineers Construction Engineering Research Laboratory. 2. A raster data format that can be used as an exchange format between two GlSs.

Georectify - the process of referencing points on an image to the real world coordinates.

Georeference - To establish the relationship between page coordinates on a paper map or manuscript and known real-world coordinates

Geospatial - a term used to describe a class of data that has a geographic or spatial nature.

Geostationary Satellite: An earth satellite that remains in fixed position in sync with the earth's rotation.

GIS - Geographic information system. A computer system of hardware and software that integrates graphics with databases and allows for display, analysis, and modeling.

Grid-Cell Data - Grid-cell data entry places a uniform grid over a map area, and the area within the cell is labeled with one attribute or characteristic, such as elevation averaged over all points. Grid cells can be layered with differing types of information.

Global Positioning System (GPS) - a system developed by the U.S. Department of Defense based on 24 satellites orbiting the Earth. Inexpensive GPS receivers can accurately determine ones position on the Earth's surface.

Ground Truth - Information collected from a survey area as remote sensing data is being collected from the same area (see control).

Hierarchical - A way of classifying data, starting with the general and going to specific labels.

Hydrography - Topography pertaining to water and drainage feature.

Hypsography - 1) The science or art of describing elevations of land surfaces with reference to a datum, usually sea level. 2) That part of topography dealing with relief or elevation of terrain.

Image - A graphic representation or description of an object that is typically produced by an optical or electronic device. Common examples include remotely sensed data such as satellite data, scanned data, and photographs. An image is stored as a raster data set of binary or integer values representing the intensity of reflected light, heat, or another range of values on the electromagnetic spectrum. Remotely sensed images are digital representations of the earth.

Imagery - a two dimensional digital representation of the earth's surface. Examples are a digital aerial photograph, a satellite scene, or an airborne radar scan.

Index - A specialized lookup table or structure within a database and used by an RDBMS or GIS to speed searches for tabular or geographic data.

Infrastructure - The fabric of human improvements to natural settings that permits a community, neighborhood, town, city metropolis, region, state, etc., to function.

Initial Graphics Exchange Specification (IGES) An interim standard format for exchanging graphics Polygon data among computer systems.'

Internet - a system of linked computer networks, worldwide in scope, that facilitates data communication services such as remote login, file transfer, electronic mail, and newsgroups. The Internet is a way of connecting existing computer networks that greatly extends the reach of each participating system.

Internet Protocol (IP) - the most important of the protocols on which the Internet is based. It allows a packet to traverse multiple networks on the way to its final destination.

Interpolate - Applied to logical contouring by determining vertical distances between given spot elevations.

IT - Information Technology

Land Information System (LIS) - the sum of all the elements that systematically make information about land available to users including: the data, products, services, the operating procedures, equipment, software, and people.

Land Information System (LIS) - NJ State 45:8-28(e) - Any computer coded spatial database designed for multi-purpose public use developed from or based on property boundaries.

Latitude - The north-south measurement parallel to the equator.

Layer- A logical set of thematic data, usually organized by subject matter.

Layers - refers to the various "overlays" of data each of which normally deals with one thematic topic. These overlays are registered to each other by the common coordinate system of the database.

Longitude - The angular distance, measured in degrees, cast or west from the Greenwich meridian, or by the difference in time between two reference meridians on a globe or sphere.

Lot Number - A numerical parcel designation, that when combined with a block number is unique to a single parcel of land within a given municipality.

Manual Digitizing - Conversion of an analog measurement into a digital form by using a manual device such as a calculator.

Map - A representation of a portion of the earth, usually drawn on a flat surface. (From Latin mappa, a napkin, sheet or cloth upon which maps were drawn.)

Map Projection - A mathematical model for converting locations on the earth's surface from spherical to planar coordinates, allowing flat maps to depict three dimensional features. Some map projections preserve the integrity of shape; others preserve accuracy of area, distance, or direction.

Map Units - The coordinate units in which the geographic data are stored, such as inches, feet, or meters or degrees, minutes and seconds.

Metadata - data describing a GIS database or data set including, but not limited to, a description of a data transfer mediums, format, and contents, source lineage data, and any other applicable data processing algorithms or procedures .

NCGIA - National Center for Geographic Information Analysis

Network Analysis - Addresses relationships between locations on a network. Used to calculate optimal routes, and optimal locations for facilities.

NSGIC - National States Geographic Information Council

NSDI - National Spatial Data Infrastructure

OPRHP - Office of Parks, Recreation and Historical Preservation

ORPS - Office of Real Property Services

Orthophoto - A photograph of the earth's surface in which geographic distortion has been removed.

Overlay - A layer of data representing one aspect of related information.

Parcel - Generally refers to a piece of land that can be designated by number.

Photogrammetry - The system of gathering information about physical objects through aerial photography and satellite imagery.

Plane-Coordinate System A system for determining location in which two groups of straight lines intersect at right angles and have as a point of origin a selected perpendicular intersection.

Planimetric Map - A map which presents the horizontal positions only for the features represented; distinguished from a topographic map by the omission of relief in measurable form. The natural features usually shown on a planimetric map include rivers, lakes and seas; mountains, valleys and plains; and forests, prairies, marshes and deserts. The culture features include cities, farms, transportation routes and public-utility facilities; and political and private boundary lines. A planimetric map intended for special use may present only those features which are essential to the purpose to be served.
Plat: A scale diagram void of cultural, drainage and relief features, showing only land boundaries and subdivisions together with data essential to its legal description.

Plotter - Equipment that can plot a graphic file using multiple line weights and colors. Types available today are: pen, laser, and electrostatic plotters.

Point Data - level of spatial definition referring to an object that has no dimension, e.g., well or weather station.

Points - Items such as oil wells, utility poles, etc. Specific objects with exact location noted.

Polygon - A vector representation of an enclosed region, described by a sequential list of vertices or mathematical functions.

Positional Accuracy - term used in evaluating the overall reliability of the positions of cartographic features relative to their true position.

Precision - refers to the quality of the operation by which the result is obtained, as distinguished from accuracy.

Protocol - a definition for how computers will perform when talking to each other. Protocol definitions range from how bits are placed on a wire to the format of an electronic mail message. Standard protocols allow computers from different manufacturers to communicate; the computers can use completely different software, providing that the programs running on both ends agree on what the data means.

Quadrangle - A four-sided region, usually bounded by a pair of meridians and a pair of parallels.

Quality Control - process of taking steps to ensure the quality of data or operations is in keeping with standards set for the system.

Raster - A grid-type data format used to interpret gray-scale photographs and satellite imagery. Imagery is stored as dots or pixels, each with a different shade or density.

Raster Data - Machine-readable data that represent values usually stored for maps or images and organized sequentially by rows and columns. Each "cell" must be rectangular but not necessarily square, as with grid data.

RDBMS - See relational database management systems.

Rectified - referencing points, lines, and/or features of two dimensional images to real world geographic coordinates, to correct distortion in the image.

Rectify - The process by which an image or grid is converted from image coordinates to real-world coordinates. Rectification typically involves rotation and scaling of grid cells, and thus requires resampling of values.

Registration - the procedure used to bring two maps or data layers into concurrence via known ground location control points or the procedure of bringing a map or data layers into concurrence with the earth's surface.

Relational Database Management System (RDBMS) - A database management system with the ability to access data organized in tabular files that may be related together by common field (item). An RDBMS has the capability to recombine the data items from different files, thus providing powerful tools for data usage.

Remote Sensing - Recording imagery or data and information from a distance. Photography is a form of remote sensing. Satellites provide a remote sensing platform for developing geology and soils analysis with sensors sensitive to various bands of the electromagnetic spectrum.

Resolution - 1. The accuracy at which the location and shape of map features can be depicted for a given map scale. For example, at a map scale of 1:63,360 (1 inch=1 mile), it is difficult to represent areas smaller than 1/10 of a mile wide or 1/10 of a mile in length because they are only 1/1 0-inch wide or long on the map. In a larger scale map, there is less reduction, so feature resolution more closely matches real world features. As map scale decreases, resolution also diminishes because feature boundaries must be smoothed, simplified, or not shown at all. 2. The size of the smallest feature that can be represented in a surface. 3. The number of points in x and y in a grid (e.g., the resolution of a USGS one-degree DEM is 1.201 x 1.201 mesh points).2

Rubber-sheet - A procedure to adjust the entities of a geographic data set in a non-uniform manner. From- and to-coordinates are used to define the adjustment.

SARA - State Archives and Records Administration

Scale - the relationship between a distance on a map and the corresponding distance on the earth. Often used in the form I :24,000, which means that one unit of measurement on the map equals 24,000 of the same units on the earth's surface.

Scanner - A scanner is an optical device that recognizes dark and light dots on a surface and converts this recognition into a digital file. However, scanners generally do not create a map database in a logically correct format, so additional computer-aided manipulation and often manual editing are used to add intelligence required by a specific GIS platform.

Scanning - Also referred to as automated digitizing or scan digitizing. A process by which information originally in hard copy format (paper print, mylar transparencies, microfilm aperture cards) can be rapidly converted to digital raster form (pixels) using optical readers.

Schematic Map - A map prepared by electronically scanning or digitizing in which the lines are not dimensionally or positionally accurate.

SDTS - Spatial Data Transfer Standard

SED - State Education Department

SEMO - State Emergency Management Office

Server - software that allows a computer to offer a service to another computer. Other computers contact the server program by means of matching client software. Also a computer using server software.

Source Material - data of any type required for the production of mapping, charting, and geodesy products including, but not limited to, ground-control aerial and terrestrial photographs, sketches, maps, and charts; topographic, hydrographic, hypsographic, magnetic, geodetic, oceanographic, and meteorological information; intelligence documents; and written reports pertaining to natural and human-made features.

Spatial Data - data pertaining to the location of geographical entities together with their spatial dimensions. Spatial data are classified as point, line, area, or surface.

Spatial Index - A means of accelerating the drawing, spatial selection, and entity identification by generating geographic-based indexes. Usually based on an internal sequential numbering system

Spatial Model - Analytical procedures applied with a GIS. There are three categories of spatial modeling functions that can be applied to geographic data objects within a GIS: (1) geometric models (such as calculation of Euclidian distance between objects, buffer generation area, and perimeter calculation); (2) coincidence models (such as a polygon overlay); and (3) adjacency models (pathfinding, redistricting, and allocation). All three model categories support operations on geographic data objects such as points, lines, polygons, TlNs, and grids. Functions are organized in a sequence of steps to derive the desired information for analysis.

Stakeholders - Any constituency in the environment that is affected by an organization's decisions and policies.
Standards - In computing, a set of rules or specifications which, taken together, define the architecture of a hardware device, program, or operating system.

State Plane Coordinate System - The plane-rectangular coordinate systems established by the United States Coast and Geodetic Survey (now known as National Ocean Survey), one for each state in the United States, for use in defining positions of geodetic stations in terms of plane-rectangular (X and Y) coordinates. Each state is covered by one or more zones, over each of which is placed a grid imposed upon a conformal map projection. The relationship between the grid and the map projection is established by mathematical analysis. Zones of limited east-west dimension and indefinite north south extent have the transverse Mercator map projection as the base for the state coordinate system, whereas zones for which the above order of magnitude is reversed have the Lambert conformal conic map projection with two standard parallels. For a zone having a width of 158 statute miles, the greatest departure from exact scale (scale error) is 1 part in 10,000. Only adjusted positions on the North American datum of 1927 and NAD 1983 may be properly transformed into plane coordinates on a state system. All such geodetic positions which are determined by the National Ocean Survey are transformed into state plane-rectangular coordinates on the proper grid, and are distributed by that bureau with the geodetic positions. State plane coordinates are extensively used in recording land surveys. and in many states such use has received approval by legislative enactment.

SUNY - State University of New York

System - A group of related or interdependent elements that function as a unit.

Tax Map - An accurate map of a municipal territory prepared for the purpose of taxation. Showing among other things, the location and width of streets, roads, avenues and each individual lot of land within the municipality.

Text Data - Information in a GIS system such as property owners' names and lot dimensions.

Thematic Layer - mapping categories, consisting of a single type of data such as population, water quality, or timber stands, intended to be used with base data.

Thematic Map A map that illustrates one subject or topic either quantitatively or qualitatively.

Theme - A collection of logically organized geographic objects defined by the user. Examples include streets, wells, soils, and streams.

TIGER - supersedes DIME (see entry) files.

TIGER - See Topologically Integrated Geographic Encoding and Referencing

Topographic Map - A map of land-source features including drainage lines, roads, landmarks, and usually relief, or elevation.

Topologically Integrated Geographic Encoding and Referencing data (TIGER) - A format used by the US Census Bureau to support census programs and surveys. It is being used for the 1990 census. TIGER files contain street address ranges along lines and census tract/block boundaries. These descriptive data can be used to associate address information and census/demographic data to coverage features.

Topology - The spatial relationships between connecting or adjacent coverage features (e.g., arcs, nodes, polygons, and points). For example, the topology of an arc includes its from- and to- nodes and its left and right polygons. Topological relationships are built from simple elements into complex elements: points (simplest elements), arcs (sets of connected points), areas (sets of connected arcs), and routes (sets of sections) that are arcs or portions of arcs). Redundant data (coordinates) are eliminated because an arc may represent a linear feature, part of the boundary of an area feature, or both. Topology is useful in GIS because many spatial modeling operations don't require coordinates, only topological information. For example, to find an optimal path between two points requires a list of which arcs connect to each other and the cost of traversing along each arc in each direction. Coordinates are only necessary to draw the path after it is calculated .

Transformation - The process of converting data from one coordinate system to another through translation, rotation, and scaling .

Transmission Control Protocol (TCP) - One of the protocols on which the Internet is based.

Vectors - Lines defined by "x", "y" and "z" coordinate endpoints. Roads, rivers, contour lines, etc. presented as vector lines.

Vector Data - A coordinate-based data structure commonly used to represent map features. Each linear feature is represented as a list of ordered x, y coordinates. Attributes are associated with the feature (as opposed to a raster data structure, which associates attributes with a grid cell). Traditional vector data structures include double-digitized polygons and arc-node models.

Vector Display: A vector display on a computer screen is produced by drawing vectors on the screen. A raster display, in contrast, is produced on a screen as rows of dots of "on" or "off' which produce the picture.

Wide Area Network (WAN) - a network that uses high-speed, long distance communications networks or satellites to connect computers over distances greater than those traversed by local area networks (LANs)--about 2 miles.

Workstations and Terminals A workstation is a device or a combination of devices integrated to provide the user with graphic data entry, display, and manipulation. These devices are used for map digitizing and map-related applications, geographic analysis and ad hoc query. Most systems still use some type of inexpensive edit-query workstations or terminals to provide low-cost access to both maps and related data.

GIS DEVELOPMENT GUIDE: NEEDS ASSESSMENT

1 INTRODUCTION

A needs assessment is the first step in implementing a successful GIS within any localgovernment. A needs assessment is a systematic look at how departments function and the spatial data needed to do their work. In addition to the final needs assessment report that is generated, intangible benefits are realized by an organization. Conducting a GIS needs assessment fosters cooperation and enhanced communication among departments by working together on a common technology and new set of tools. Finally, the needs assessment activity itself serves as a learning tool where potential users in each participating department learns about GIS and how it can serve the department.

A needs assessment is required if the local government will be adopting a GIS throughout the organization. Without a complete needs assessment each department might proceed to adopt their own system and database which may or may not be compatible with those of another department. The largest benefit for a local government adopting a GIS is to realize efficiencies from common "base data" and the sharing of data among departments.

At the conclusion of a needs assessment, an organization will have all of the information needed to plan the development of a GIS system. This information can be grouped into the following categories:

- ***Applications to be developed***. - In evaluating the responsibilities and work flow within a department, certain tasks are identified that can be done more efficiently or effectively in a GIS. These tasks will form the basis of GIS applications. Application descriptions prepared as part of the needs assessment will describe these tasks.

- ***GIS Functions required***. - For each application identified, certain GIS functions will be required. These will include standard operations such as query and display, spatial analysis functions such as routing, overlay analysis, buffering, and possibly advanced analysis requiring special programming.

- ***Data needed in the GIS database***. - Most departments in local government use data that has a spatial component. Much of this data are hardcopy maps or tabular data sets that have a spatial identifier such as addresses and zip codes or X-Y values (latitude-longitude, state plane coordinates, or other coordinate system). A needs assessment will identify how this information will be used by GIS applications.

- ***Data maintenance procedures***. - By looking at the work flow and processes within and between departments, responsibility for data creation, updates and maintenance will become apparent.

Note: The needs assessment procedure refers to a *local government* and its *departments* as the organizational units. In a multi-agency GIS cooperative, the same activities described would be carried out by all participants, at the appropriate level of detail as determined by the role each participant would play in the resulting GIS cooperative.

Once all of this information is collected and analyzed for each department and published in a report, it can be used as a blueprint for implementing the GIS. The GIS coordinating group within the organization will use it to:

- Design the GIS database
- Identify GIS software that will meet the government's needs
- Prepare an implementation plan
- Start estimating the benefits and costs of a GIS

A common mistake in performing a needs assessment is to simply take an inventory of the maps and spatial data currently used in each department. There are two major problems with this approach. First, this does not allow the GIS coordinating group to evaluate how a GIS could be used to enhance the work of each department and the agency as a whole . By looking at the department functions and what the department does or produces, the GIS coordinating group and potential users develop an understanding of the role GIS can play in the organization. The existing data and maps do need to be inventoried and may well be used in building the GIS, however such an inventory should be separate from the needs assessment.

The second major problem with the "data inventory" approach is that it tends to focus only on data internal to the organization. Local governments rely heavily on data from outside sources - federal agencies, state agencies, business, etc. The need for these data is better determined by looking at the potential GIS applications and how data will be used by each application. It can then be determined what data should be acquired from other sources.

2 CONDUCTING A NEEDS ASSESSMENT

The most significant aspect of a needs assessment is to document the findings in a standard and structured manner. It is very important to adopt (or develop) a standard method to be used for the description of all the GIS tasks, processes and data that will be included in the needs assessment. These forms will be used in needs assessment to identify the three kinds of GIS requirements:

- ***GIS applications*** - these will be tasks that can be performed by the GIS when a user requires them, such as preparing a map, processing a query, or conducting some particular GIS analysis. GIS applications can be described using the five page GIS Applications forms included with this guide as Appendix A.

- ***GIS activity*** - these are situations where information needs to be kept on some activity or process important to the user, such as issuing building permits, conducting public health inspections, etc. A GIS activity can be described using pages 1 and 4 of the GIS Applications forms - the main application form and the data flow diagraming (DFD) form.

- ***GIS data*** - there will be certain categories of spatial data that are important to keep, but which will not appear in any GIS application or activity identified in any application description. A separate method must be developed to systematically record the need for such data. Other GIS data needed but not included in either of the above categories, can be entered directly into the master data list.

The main method used to collect the information to enter onto the forms is individual interviews. Potential users of the GIS can be identified by management and by examination of the organization chart. A series of one-on-one interviews is the best way to identify the users needs. During the interview, the user can usually identify documents that can provide additional information to the GIS analyst.

- The needs assessment activity is composed of two main parts
- Interviewing and documenting the needs of potential GIS users
- Compiling the results of the needs assessment into the ***master data list*** and the ***list of GIS functions***. These two lists respectively are used to prepare the ***GIS data model*** and the ***GIS specifications*** (activities described under Conceptual Design).

The interview process should identify and describe all anticipated uses of the GIS. The next section briefly describes the major categories of GIS use, followed by a detailed description on how to complete the needs assessment forms.

3 LOCAL GOVERNMENT USES OF GIS

The use of geographic information systems by local government falls into five major categories:

- Browse
- Simple display (automated mapping);
- Query and display;
- Map analysis; and
- Spatial modeling.

Browse

This function is equivalent to the human act of reading a map to find particular features or patterns. Browsing usually leads to identification of items of interest and subsequent retrieval and manipulation by manual means. For single maps, or relatively small areas, the human brain is very efficient at browsing. However, as data volumes increase, automated methods are required to effectively extract and use information from the map.

Simple Display

This GIS function is the generation of a map or diagram by computer. Such maps and diagrams are often simple reproduction of the same maps used in a previous manual orientedGIS environment. Examples of this type of use are preparation of a 1:1000-scale town map, a sketch of an approved site plan, maps of census data, etc.

Query And Display

This function supports the posing of specific questions to a geographic database, with the selection criteria usually being geographic in nature. A typical simple query would be: "draw a map of the location of all new residential units built during 1989" A more complex query might be: "draw a map of all areas within the town where actual new residential units built in 1989 exceeds growth predictions." Such a query could be part of a growth management activity within the town. Queries may be in the form of regular, often asked questions or may be ad hoc, specific purpose questions. The ability to respond to a variety of questions is one of the most useful features of a GIS in its early stages of operation. In the long run, other more sophisticated applications of the GIS may have a higher value or benefit, but to achieve these types of benefits, users must be familiar with the GIS and its capabilities. Such familiarization is achieved through the use of a GIS for the simpler tasks of query and display.

Map Analysis (Map Overlay)

This involves using the analytical capabilities of GIS to define relationships between layers of spatial data. Map analysis is the super-imposition of one map upon another to determine the characteristics of a particular site (e.g., combining a land use map with a map of flood prone areas to show potential residential areas at risk for flooding). Map analysis (often termed overlay or topological overlay) was one of the first real uses of GIS. Many government organizations, particularly those managing natural resources, have a need to combine data from different maps (vegetation, land use, soils, geology, ground water, etc.). The overlay function was developed to accomplish the super-imposition of maps in a computer. The data are represented as polygons, or areas, in the GIS data base, with each type of data recorded on a separate "layer." The combination of layers is done by calculating the logical intersection of polygons on two or more map layers. In addition to combining multiple "layers" of polygon-type data, the map overlay function also permits the combination of point data with area data (point-in-polygon). This capability would be very useful in a town for combining street addresses (from the Assessor's files) with other data such as parcel outlines, census tract, environmental areas, etc. Many facility siting problems, location decisions, and land evaluation studies have successfully used this procedure in the past.

Spatial Modeling

This application is the use of spatial models or other numerical analysis methods to calculate a value of interest. The calculation of flow in a sewer system is an example of spatial modeling. Spatial modeling is the most demanding use of a GIS and provides the greatest benefit. Most spatial modeling tasks are very difficult to perform by hand and are not usually done unless a computerized system, such as a GIS, is available. These models allow engineers and planners to evaluate alternate solutions to problems by asking "what if" type questions. A spatial model can predict the result expected from a decision or set of decisions. The quality of the result is only as good as the model, but the ability to test solutions before decisions have to be made usually provides very useful information to decision makers. Once again, this type of use of a GIS will evolve over time, as the GIS is implemented and used.

A closely related computer capability is a CAD system (computer aided design). CAD systems are used to prepare detailed drawings and plans for engineering and planning applications. While CAD systems functions are different from GIS functions, many commercial CAD products have some of the functionality normally found in a GIS. There are, however, significant differences between a CAD system and a GIS, mainly in the structure of the data base. There may be some need for CAD-type capabilities in a particular local government, so this forms another category of use.

In general, geographic information in local government is used to:

- Respond to public inquiries,
- Perform routine operations such as application reviews and permit approvals, and
- Provide information on the larger policy issues requiring action by the town board.

These are typical local government activities which benefit from a geographic information system. The development of GIS will facilitate the present geographic information handling tasks and should lead to the development of additional applications of benefit to the local government.

There are also other computer systems in local governments that perform GIS-like functions, such as Emergency 911, underground utility locator systems, school bus routing systems, etc. The variety and diversity of GIS applications are what make the definition of a GIS very difficult. Basically, any computer system where the data have one or more spatial identifiers or that perform spatial operations can be classed as a GIS. For example, a system containing street addresses and census tract codes and that has the ability to place a given street address in the proper census tract is a GIS whether or not map boundaries are part of the system. There are two important points here:

- A large proportion of local government data does have one of more spatial identifiers, and therefore has the potential of being part of a GIS.
- Other, existing systems with GIS data or performing GIS-like functions must be integrated into the overal system design. GIS should not be developed as a separate system.

Whether a local government unit is considering or planning a "full, multi-purpose GIS" or is only interested in a limited or single function system, the database planning and design considerations are the same. Only the magnitude of the analysis and design activities differ. Some GIS users believe that smaller and simpler applications, such as a school bus routing system, do not require a formal planning activity. There are, however, several reasons to conduct such a planning activity for the smaller applications:

- To ensure that the user requirements will be fully met

- To develop documentation, especially data documentation (metadata), needed to use and maintain the GIS

- To be in a position to participate in data sharing programs with other agencies as additional applications are developed

- To create a permanent record of the data and its use to document agency plans and decisions, and to meet data retention and archiving requirements.

- To use as a base for building a larger, multi-function at some later date.

The level of effort needed to complete a GIS plan can be kept commensurate with the scope and size of the intended GIS. Further, the GIS planning software tool that accompanies these guidelines provides an easy and convenient way to create the recommended documentation.

4 DATA USED BY LOCAL GOVERNMENT

There are many kinds of data used by local government that can be included in a GIS. Data in a GIS can be one of two types: spatial data and non-spatial data. Spatial data is that data which is taken from maps, aerial photographs, satellite imagery, etc. It is composed of spatial entities, relationships between these entities, and attributes describing these entities. Non-spatial data is usually tabular data taken from tables, lists, etc. Most of the time, the non-spatial data will be linked to one or more spatial entities by keys (unique identifiers associated with the spatial data and non-spatial data). For example, the tax map would represent the spatial data while the real property inventory is non-spatial data, which is linked to the entities(parcels) on the tax map.

Spatial data is commonly represented by geometric objects (points, line, and polygons). Non-spatial data containing a spatial reference is also considered spatial data. One of the most common forms of this type of data in local government are records and files referenced by street address.

Examples of local government data that have been used with GIS include:

- Tax parcels
- Real property inventories
- Infrastructure data
  - Water system
  - Sewer system
  - Electric
- Census data
- Land use maps
- Zoning maps
- Planimetrics
  - Right-of-way
  - Waterways (streams)
  - Building Outlines
- Permit records

The operations required in a GIS must meet the data handling requirements of the spatial data as well as those of the non-spatial data. The most common use of a GIS in local government is the query based on attribute keys and then displayed in map form.

5 DOCUMENTING GIS NEEDS

The GIS needs are documented using the following forms (full-page sized copies of all forms are included in Appendix A):

The GIS Application Description (5 pages) used to:
- Describe products (mostly map displays) produced by the GIS
- Describe activities supported by the GIS

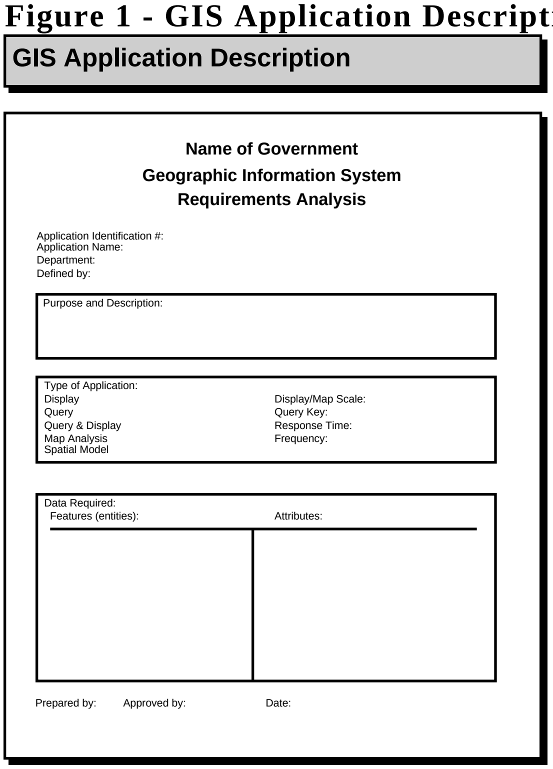
The Master Data List

Most GIS applications can be described using the GIS Application Description. In cases where these forms are not appropriate, any other systematic description of the need can be used. If more appropriate, different forms can be developed as long as the same information can be systematically recorded: the data required and the GIS functions need to develop the GIS product.

GIS Application Description

The set of forms used to document a GIS contains five pages:

Figure 1 - GIS Application Descriptions


GIS Application Description

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Purpose and Description:

Type of Application:
Display — Display/Map Scale:
Query — Query Key:
Query & Display — Response Time:
Map Analysis — Frequency:
Spatial Model

Data Required:
Features (entities): — Attributes:

Prepared by: Approved by: Date:

GIS Application Description (Page A- 1)

Use to enter:
- Application identification
- Description of purpose
- Type of application, map scale, query key, frequency, and required response time
- Data needed by the application
  - Entities (features)
  - Attributes of entites

Map Display

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Graphical Output Sample: *Screen:* *Hard Copy:*

Symbols/Legend

Prepared by: Approved by: Date:

Map Display (Page A- 2)

Used to draw a sample of any maps to be produced by the application (including the legend showing symbols for each feature). This can be a hand sketch, although it should be drawn to the scale of the output desired.

Table Display

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Report Layout/Format: *Screen:* *Hard Copy:*

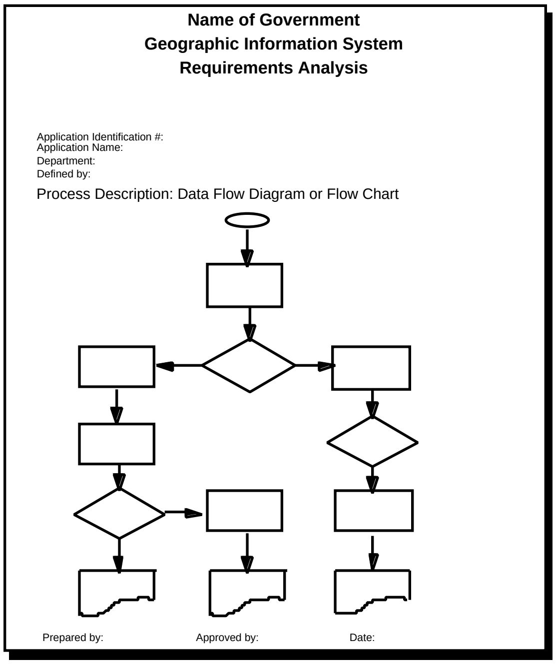
HEADINGS				
SUB-HEADINGS				
SUB-TOTALS/TOTALS:				

Prepared by: Approved by: Date:

Table Display (Page A- 3)

Used to show samples of any tables to be produced by the application (used only if tables are needed in the application). If any entries in the table involve complex calculations, these should be described using either a Data Flow Diagram (page 4) or other separate pages.

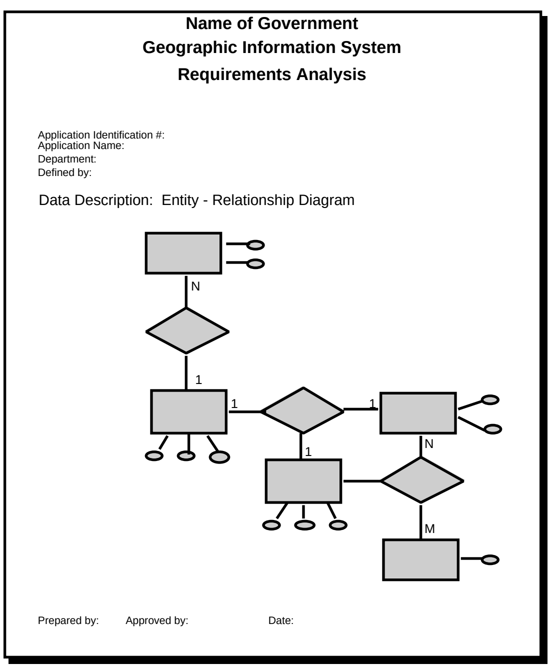
Data Flow Diagram



Data Flow Diagram (Page A-4)

Used to draw a data flow diagram or flow chart when an application is complex. This chart is usually drawn by the GIS analyst or someone else familiar with the diagraming techniques, and is used to document complex calculations or descriptions of activities that will need GIS support.

Entity-Relationship Diagram



Entity-Relationship Diagram (Page A- 5)

Used to draw an entity-relationship (E-R) diagram of the data used in the application. This drawing is usually done by the GIS analyst or someone else familiar with the E-R technique, and is only done for more complex GIS applications.

6 DOCUMENTING AN ACTIVITY-TYPE USE OF THE GIS

Some GIS applications in local government do not involve the production of maps and tables. For example, a GIS may be used to record and store information about a building permit application, a subdivision plat, a site plan, etc. Many activities of local government are simply the processing of permits from individuals or firms. If any of these activities will also generate GIS data, they should be described for the needs assessment. Two techniques available for describing processes are flow charts and data flow diagrams.

A completed application description for a local government activity of this type can be entered on pages 1 and 4 of the GIS Application Description forms. Page 4 - Data Flow Diagram would appear as follows:

Data Flow Diagram Example

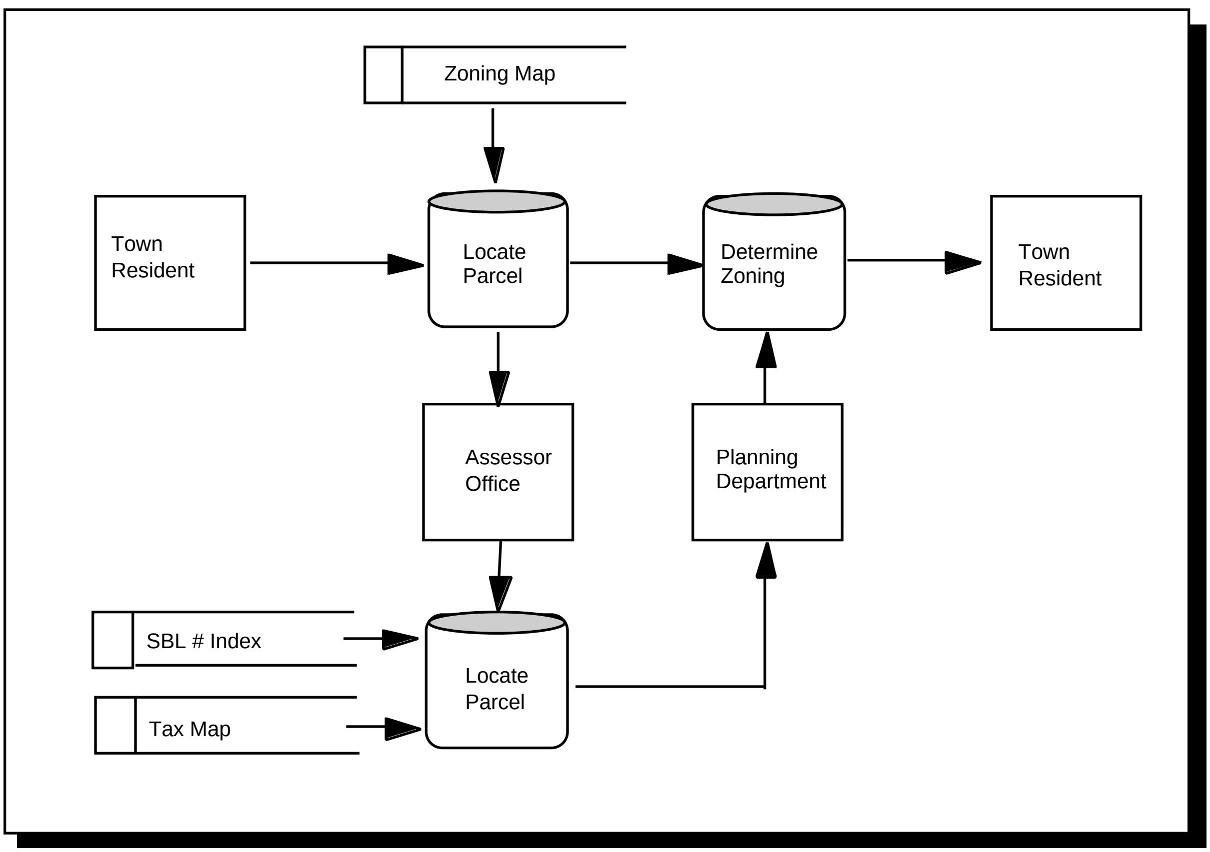


Figure 2 - Data Flow Diagram Example

This example shows a data flow diagram that has three participants (town resident, planning department, and assessor's department) that uses three parts of the database (zoning map, section-block-lot number index, and tax map), to answer a zoning inquiry. Appendix D contains a brief description of the data flow diagraming method.

7 THE MASTER DATA LIST

The master data list is a composite of all data entities (features) and their attributes that have been entered in the data section of the GIS Application Description (Page 1). Other data identified by users as "needed," but not included in any application description may be entered directly into the master data list.

Master Data List

Entity	**Attributes**	**Spatial Object**
Street_segment	name, address_range	Line
Street_intersection	street_names	Line
Parcel	section_block_lot#, owner_name, owner_address, site_address, area, depth, front_footage, assessed_value, last_sale_date, last_sale_price, size owner_name, owner_address, assessed_value (as of previous January 1st)	Polygon
Building	building_ID, date_built, building_material, building_assessed_value	Footprint
Occupancy	occupant_name, occupant_address, occupancy_type_code	None
Street_segment	name, type, width, length, pavement_type	Polygon
Street_intersection	length, width traffic_flow_conditions, intersecting_streets	Polygon,
Water_main	type, size, material, installation_date	Line
Valve	type, installation_date	Node
Hydrant	type, installation_date, pressure, last_pressure_test_date	Node
Service	name, address, type, invalid_indicator	None
Soil	soil_code, area	Polygon
Wetland	wetland_code, area	Polygon
Floodplain	flood_code, area	Polygon
Traffic_zone	zone_ID#, area	Polygon
Census_tract	tract#, population	Polygon
Water_District	name, ID_number	Polygon
Zoning	zoning_code, area	Polygon

Figure 3 - Master Data List

8 CONDUCTING INTERVIEWS

Individual interviews are the most effective way of finding out from users their potential GIS applications. Before starting interviews, a briefing session for all potential users should be held. During this meeting, the interviewers should describe the entire needs assessment procedure to all participants. The main activities will be:

- Conduct "start-up" seminar or workshop
- Interview each potential user
- Prepare documentation (forms) for each application, etc
- Review each application description with the user
- Obtain user approval of and sign-off for each application description

An introductory seminar or workshop with all potential users in attendance is useful to prepare the way for user interviews. At the beginning of a project, many users may not have much knowledge about GIS or how it might help them. Also, the interview team may be from outside the organization and may not be very familiar with the structure of the particular local government. The start-up seminar should address the following topics:

- Definitions:
  - What is a GIS?
  - How is a GIS used by local government? (Typical applications)

- Interview procedure to be followed:
  - What the interviewee will do?
  - What is expected from the interviewee?
  - Who approves the application descriptions?
  - How the information from the application descriptions will be used?

- Group discussion: It is often useful to have the group identify an initial set of GIS applications as candidates for further documentation. The discussion of possible applications between interviewers and users will start to reveal what is suitable for a GIS application. One or more applications can be described in the process by the group so everyone sees how the process will work.

It is preferable to interview users individually rather than in groups. This provides a better opportunity to explore the ideas of each person and also prevents other individuals from dominating any particular meeting. Group meetings easily lose focus on specific GIS applications and therefore do not provide the detailed information needed to adequately describe the GIS applications.

Conducting an interview is not an easy task. Some potential users may have a good grasp of GIS and how they might use one. However, often potential users do not have complete knowledge of the capabilities of a GIS and therefore may not be able to readily identify GIS applications. In these cases, the interviewer (GIS analyst) needs to help the user explore his/her job activities and responsibilities to identify GIS opportunities. The GIS analyst should usually begin an interview with a review of the procedure, then ask the user to identify and describe potential applications. When specific GIS applications cannot be easily identified, it is helpful if potential users describe, in general, his/her job functions and responsibilities and the role their department plays in the whole organization. From this discussion, the GIS analyst can usually identify potential GIS applications and then explore these for possible inclusion in the needs assessment.

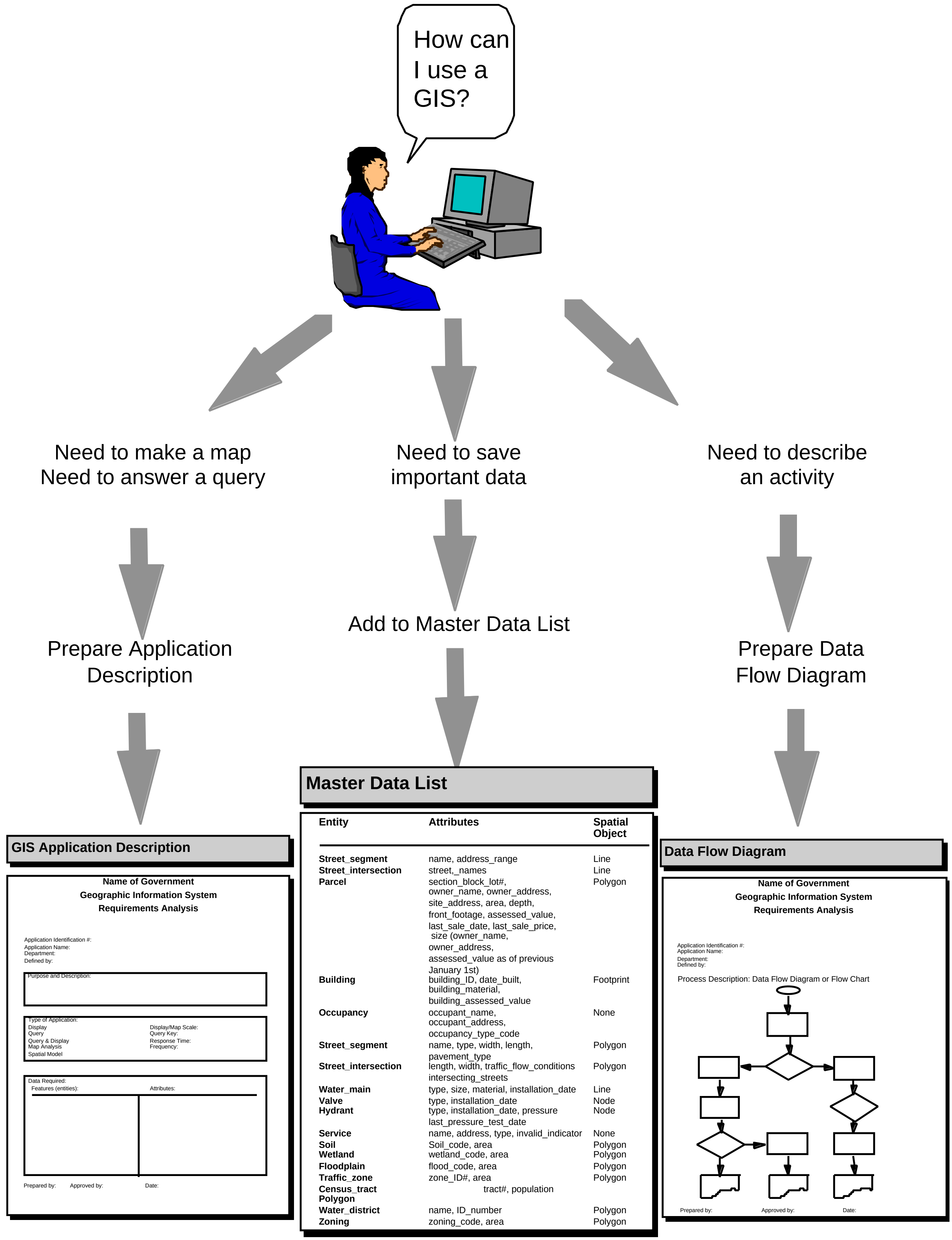

How can I use a GIS?
Need to make a map
Need to answer a query
Need to save important data
Need to describe an activity
Prepare Application Description
Add to Master Data List
Prepare Data Flow Diagram
GIS Application Description
Name of Government
Geographic Information System
Requirements Analysis
Application Identification #:
Application Name:
Department:
Defined by:
Purpose and Description:
Type of Application:
Display
Query
Query & Display
Map Analysis
Spatial Model
Display/Map Scale:
Query Key:
Response Time:
Frequency:
Data Required:
Features (entities):
Attributes:
Prepared by:
Approved by:
Date:
Master Data List
Entity | Attributes | Spatial Object
Street_segment | name, address_range | Line
Street_intersection | street,_names | Line
Parcel | section_block_lot#, owner_name, owner_address, site_address, area, depth, front_footage, assessed_value, last_sale_date, last_sale_price, size (owner_name, owner_address, assessed_value as of previous January 1st) | Polygon
Building | building_ID, date_built, building_material, building_assessed_value | Footprint
Occupancy | occupant_name, occupant_address, occupancy_type_code | None
Street_segment | name, type, width, length, pavement_type | Polygon
Street_intersection | length, width, traffic_flow_conditions intersecting_streets | Polygon
Water_main | type, size, material, installation_date | Line
Valve | type, installation_date | Node
Hydrant | type, installation_date, pressure last_pressure_test_date | Node
Service | name, address, type, invalid_indicator | None
Soil | Soil_code, area | Polygon
Wetland | wetland_code, area | Polygon
Floodplain | flood_code, area | Polygon
Traffic_zone | zone_ID#, area | Polygon
Census_tract Polygon | tract#, population
Water_district | name, ID_number | Polygon
Zoning | zoning_code, area | Polygon
Data Flow Diagram
Name of Government
Geographic Information System
Requirements Analysis
Application Identification #:
Application Name:
Department:
Defined by:
Process Description: Data Flow Diagram or Flow Chart
Prepared by:
Approved by:
Date:

Figure 4 - Interviewing and Documenting Needs of a Potential GIS User

9 PREPARING THE NEEDS ASSESSMENT REPORT

The needs assessment report consists of the application descriptions, the master data list, and several summary tables. A list of all applications summarizing the type and frequency of use is the first table.

App #	Application Name	Type	Frequency
1	Zoning Query	Query & Display	85 / day
2	Customer Phone Inquiry	Query & Display	100/day
3	Fire Dispatch Map	Query & Display	86/day
4	Fire Redistricting Map	Map Analysis	1/year
5	Crime Summary Map	Query & Display	12/month
6	Patrol Dispatch Map	Query & Display	133/day
7	Complaint Summary Map	Query & Display	624/year
8	Subdivision Development Map	Query & Display	No estimate
9	Counter Query Map	Query & Display	85/day
10	Land Use/Land Value	Map Display	1/year
11	Assessed Value Map	Query & Display	144/year
12	Grievance Map	Query & Display	2500/year
13	Comparable Value Map	Query & Display	No estimate
14	Built/Vacant Map	Display	1/year
15	Water and Sewer Line Map	Query & Display	30/month
16	Hydrologic Profile Map	Spatial Model	1440/year
17	Sewer System Flow Analysis	Spatial Model	12/year
18	Emergency Repair Map	Query & Display	110/year
19	Storm Drainage Map	Spatial Model	700/year
20	Fire Flow Test Map	Spatial Model	260/year
21	Easement Map	Query & Display	520/year
22	Zoning Map	Query & Display	50/day
23	Floodplain Map	Query & Display	50/day
24	Youth League Residency	Check Query & Display	3500/year
25	Mosquito Control Area Map	Query & Display	50/year
26	Site Plan Approval Process	Query & Display	200/year
27	Census Data Map	Display	48/year
28	Population Density Map	Map Analysis	50/year
29	Land Use Inventory	Display	24/year
30	Retail Space Projection	Spatial Model	24/year
31	Office Space Projection	Spatial Model	12/year
32	Traffic Volume Map	Query & Display	24/year

Figure 5 List of GIS Applications

This table contains selected GIS applications from the Town of Amherst, N.Y. Needs Assessment

GIS Application by Department by Type

Department	Display	Query & Display	Map Analysis	Spatial Model	Total
Fire Dispatch	0	3	1	1	5
Police	0	4	0	0	4
Assessor	2	5	0	0	7
Engineering	3	7	0	0	15
Building	2	2	0	0	4
Recreation	0	2	0	0	2
Highway	0	10	0	1	11
Planning	10	12	3	3	28
Total	17	45	4	10	76

GIS Application by Dept. by Frequency

Department	Display	Query & Display	Map Analysis	Spatial Model	Total
Fire Dispatch	0	94,170	1	100	94,271
Police	0	49,637	0	0	49,637
Assessor	2	23,894	0	0	23,896
Engineering	18	2,049	0	3,452	5,519
Building	250	25,000	0	0	25,250
Recreation	0	3,520	0	0	3,520
Highway	0	1,475	0	10	1,485
Planning	718	2,536	80	40	3,374
Total	988	202,281	81	3,602	206,952

Figure 6 - Table Summarizing Applications Example

The data from the first table can be used to prepare tables summarizing applications by department and the frequency of applications by department.

Numbers in these tables are from the Town of Amherst, N.Y. needs assessment and represent the estimates of GIS use per year. These numbers will be used during the database Planning and Design phase to estimate usage and benefits, of the GIS. In this example, for the Town of Amherst, it is estimated that 2.5 minutes of staff time will be saved for each query giving a total savings of 4.03 years staff time/year (202,281 times 2.5 minutes divided by 60 minutes/hour divided by 2088 hours per year).

The last table relates GIS applications to the data used by each application.

Application/Data Item Matrix:

	Roads	Buildings	Land Parcels	Water Mains	Fire Hydrants	Wetland Areas			
# 1 Leak Detection Map		x		x	x				
#2 Customer Service Report	x		x	x					
#3 Pressure Test Map		x		x	x				
#4 Hydraulic Model Analysis				x	x				
#5 Work Crew Schedule	x	x	x						

This matrix is very useful in planning and scheduling data conversion. If applications are prioritized, then data needed by high priority applications can be scheduled for conversion early in the conversion process. Also, if some data is not available for some reason, it is possible to determine the affected applications.

Figure 7 - GIS Applications/Data Matrix

The last step in compiling the needs assessment report is to extract the list of GIS functions needed from the application descriptions. This list will include the standard function types of display and query and display plus any other functions included in a data flow diagram or flow chart. Typical examples of such GIS functions are: calculate distance between objects, determine the shortest path through a network, etc. Figure 8 is an example of a GIS functions list.

GIS Functions/Procedures List

GIS Functions Needed:

GIS Functions (from Applic. Desc.)	Generic GIS Functions	Candidate GISs: ARC/INFO		INTERGRAPH		SYSTEM 9	
		Basic	Macro	Basic	Macro	Basic	Macro
DISPLAY	SCREEN DISPLAY	ARCPLOT		YES		YES	
	PLOTTER DISPLAY	ARCPLOT		YES		YES	
	GENERATE REPORT	INFO		YES		YES	
QUERY	ATTRIBUTE QUERY	INFO		YES		YES	
SPATIAL QUERY	SPATIAL SEARCH	IDENTIFY		YES		YES	
MAP ANALYSIS	OVERLAY	ARC		YES		YES	
	BUFFER	ARC		YES		YES	
	RECLASSIFY	ARC		YES		YES	
SHORTEST PATH	SHORTEST PATH	NETWORK		NO			YES
ROUTE	ROUTE	NETWORK		NO			YES
HYDRAULIC MODEL			AML		YES		YES

Figure 8 - GIS Function List

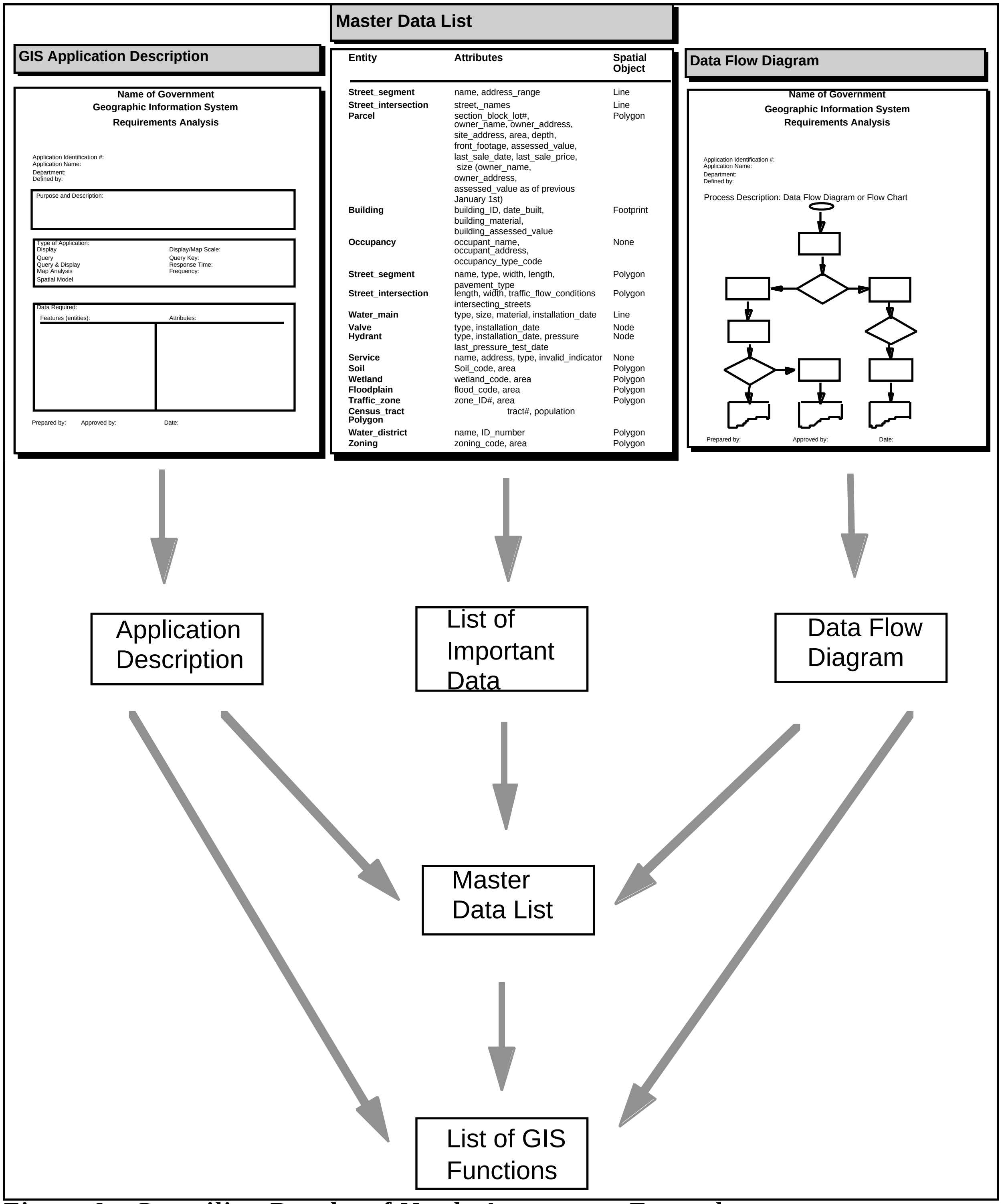


Figure 9 - Compiling Results of Needs Assessment Example

The list of GIS functions and the master data list will be used in subsequent tasks to design the database and prepare the GIS specifications.

10 SUMMARY

The procedure presented in the guideline for preparing a needs assessment is based on documenting GIS applications in a standard format. The components of this format are structured to facilitate communication between potential GIS users and the GIS analyst, and to provide specific and detailed information to the GIS analyst for designing the GIS. The first page of the application description is the most critical to the GIS analyst as it contains the list of data and an indication of the GIS functionality required by the application. If additional information on the GIS functionality is needed, than a flow chart or data flow diagram can be developed (page 4 of the application description). For the potential user, the map display and report format describe output he/she will receive. These pages should be sufficiently detailed for the user to approve or sign-off as to the correctness of the application description. It is, of course, very important that the entire GIS application description be internally consistent.

The entity-relationship diagram (page 5) is mainly useful in the next phase of the GIS design - Conceptual Design, where the data model for the entire system will be defined. If entity-relationship diagrams are prepared for individual applications, they will than be available for the Conceptual Design phase. Otherwise, these diagram can be prepared during the Conceptual Design phase.

Figure 9 is a diagrammetric representation of the flow of information from the elements of the application description to the master data list and the list of GIS functions.

Appendix Table of Contents

Appendix A - GIS Application Description Forms

Appendix B

Appendix C - Sample GIS Application Descriptions

Appendix D

Appendix E

GIS Application Description

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Purpose and Description:

Type of Application:

Display	Display/Map Scale:
Query	Query Key:
Query & Display	Response Time:
Map Analysis	Frequency:
Spatial Model	

Data Required:

Features (entities):	Attributes:

Prepared by: Approved by: Date:

Map Display

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Graphical Output Sample: *Screen:* *Hard Copy:*

Symbols/Legend

Prepared by: Approved by: Date:

Table Display

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Report Layout/Format: *Screen:* *Hard Copy:*

HEADINGS				
SUB-HEADINGS				
SUB-TOTALS/TOTALS:				

Prepared by: Approved by: Date:

Data Flow Diagram

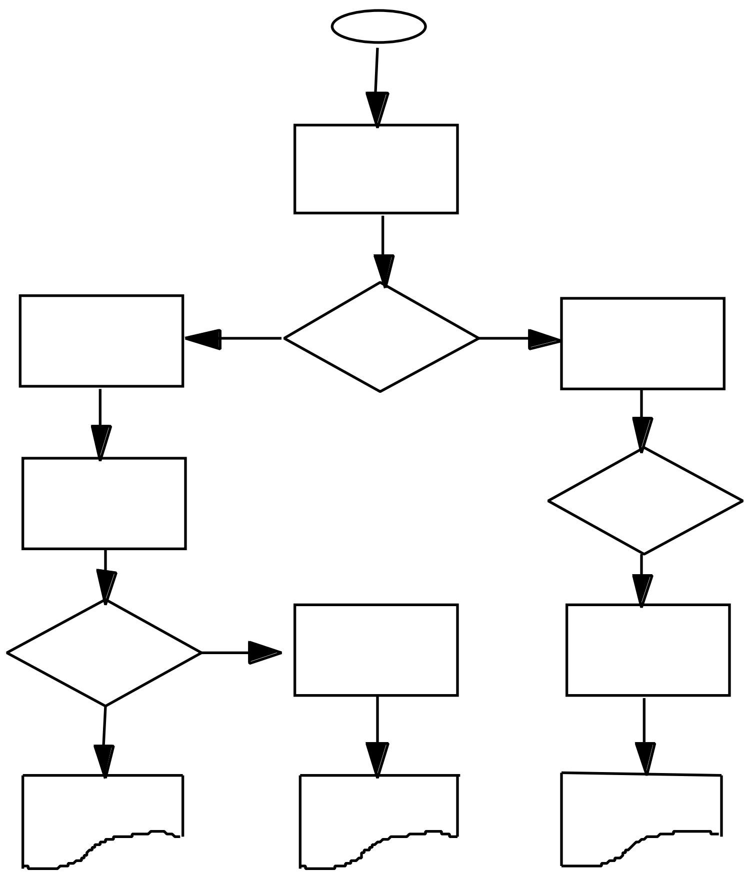
Name of Government

Geographic Information System

Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Process Description: Data Flow Diagram or Flow Chart



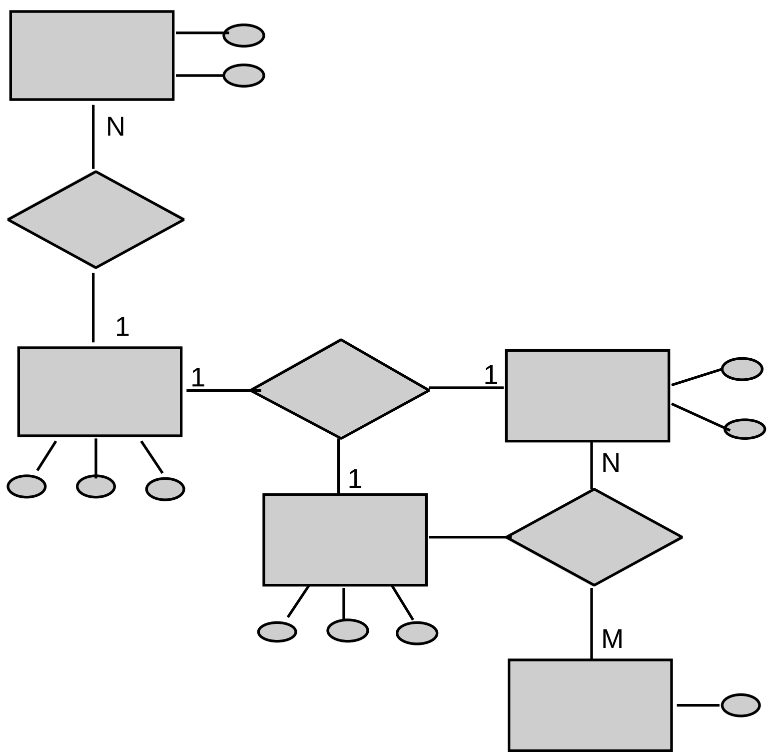
Prepared by: Approved by: Date:

Entity-Relationship Diagram

Name of Government
Geographic Information System
Requirements Analysis

Application Identification #:
Application Name:
Department:
Defined by:

Data Description: Entity - Relationship Diagram



Prepared by: Approved by: Date:

Master Data List

Entity	Attributes	Spatial Object
Street_segment	name, address_range	Line
Street_intersection	street,_names	Line
Parcel	section_block_lot#, owner_name, owner_address, site_address, area, depth, front_footage, assessed_value, last_sale_date, last_sale_price, size (owner_name, owner_address, assessed_value as of previous January 1st)	Polygon
Building	building_ID, date_built, building_material, building_assessed_value	Footprint
Occupancy	occupant_name, occupant_address, occupancy_type_code	None
Street_segment	name, type, width, length, pavement_type	Polygon
Street_intersection	length, width, traffic_flow_conditions intersecting_streets	Polygon
Water_main	type, size, material, installation_date	Line
Valve	type, installation_date	Node
Hydrant	type, installation_date, pressure last_pressure_test_date	Node
Service	name, address, type, invalid_indicator	None
Soil	Soil_code, area	Polygon
Wetland	wetland_code, area	Polygon
Floodplain	flood_code, area	Polygon
Traffic_zone	zone_ID#, area	Polygon
Census_tract	tract#, population	
Polygon		
Water_district	name, ID_number	Polygon
Zoning	zoning_code, area	Polygon

Erie County Water Authority
Geographic Information System
Requirements Analysis

Application Identification #: 19
Application Name: Customer Phone Inquiry
Department: Dispatch
Defined by: T. May

Purpose and Description:

To respond to phone inquires of: 1) "no water;" or 2) to take requests for service. (Reference Donohue #18)

Type of Application:

Display	Display/Map Scale: 1" = 100'
Query	Query Key: Address
Query & Display XX	Response Time: 10 seconds
Map Analysis	Frequency: xx/day
Spatial Model	

Data Required:

Features (entities):	Attributes:
ROW	Location (boundary), street name, street_address_range
Pipe	Location (line), size
Parcel	Location (boundary), address
Building	Location (footprint)
Services	Location(parcel), address, name, status
Projects	Location (boundary), current_in_progress
Work_orders	Location (parcel_by_address), current, date work_order_number

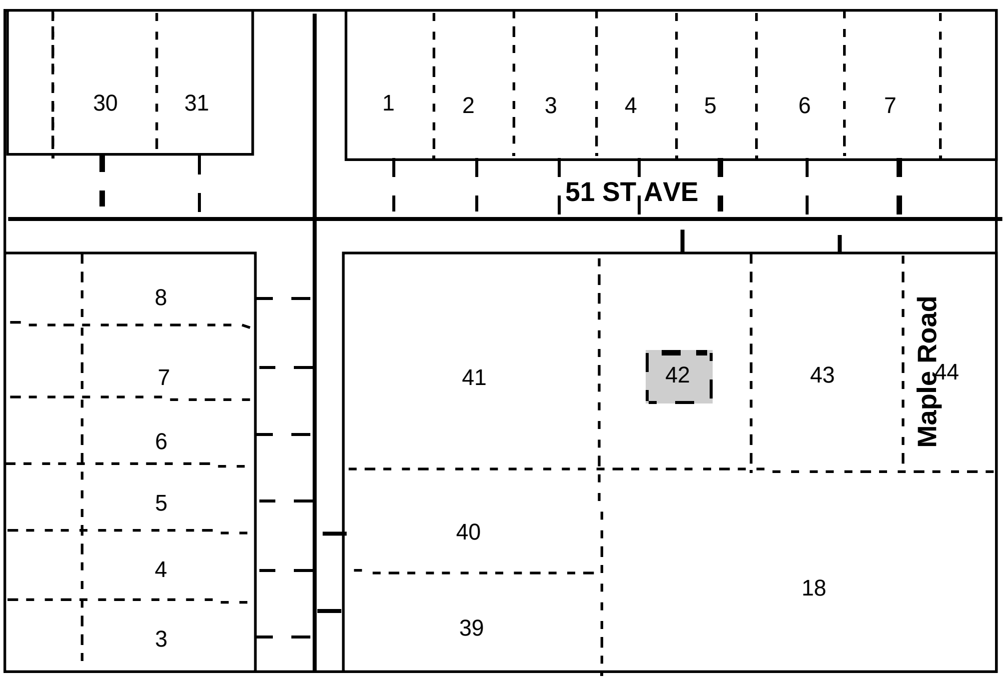
Prepared by: Approved by: Date:

ECWA Geographic Information System

Erie County Water Authority
Geographic Information System
Requirements Analysis

Application Identification #:	19
Application Name:	Customer Phone Inquiry
Department:	Dispatch
Defined by:	T. May

Graphical Output Sample: *Screen: XX* *Hard Copy:*



Symbols/Legend

Parcel
Water Main
Building
ROW
Service Connection

Prepared by: Approved by: Date:

Erie County Water Authority
Geographic Information System
Requirements Analysis

Application Identification #:	19
Application Name:	Customer Phone Inquiry
Department:	Dispatch
Defined by:	T. May

Report Layout/Format: *Screen: XX* *Hard Copy:*

Customer Service Inquiry

Name	Address	Status	Pipe Size	Project	Work Order
J.J. Jones	1551 51st Ave.	Active	4 in.	None	None

Prepared by: Approved by: Date:

County of Erie
Geographic Information System
Requirements Analysis

Application Identification #: 1
Application Name: Erie County Map Guide
Department: Public Works
Defined by: Roger Fik

Purpose and Description:
To provide a general multi-purpose map of Erie County for public use.

Type of Application:

Display	X	Display/Map Scale: 1"=2 miles
Query		Query Key: N/A
Query & Display		Response Time: 5 to 10 minutes
Map Analysis		Frequency: Yearly publication
Spatial Model		

Data Required:

Features (entities):	Attributes:
County Boundary	Location, Name (line)
Townships	Location, Name (Polygon)
Cities	Location, Name (Polygon)
Villages	Location, Name (Polygon)
Communities	Location, Identifier (Node)
County Roads	Location, Name (Line)
City, Town, & Village Roads	Location, Name (Line)
State Highways	Location, Name (Line)
Interstate, State Thruway, & Expressways	Location, Name (Line)
Interstate Route Numbers	Route Identifier, Location
State Route Numbers	Route Identifier, Location
US Route Numbers	Route Identifier, Location
Reservations	Location, Name (Polygon)
State Parks	Location, Name (Polygon)
County Parks	Location, Name (Polygon)
County Forests	Location, Name (Polygon)
Streams, Rivers, & Creeks	Location, Name (Line)
Water Bodies	Location, Name (Polygon
Airports	Location, Name (Node)
County Jurisdiction Designation	Location, Name (Node)

Prepared by J. Volpe: Approved by: Date: 3/15/94

County of Erie
Geographic Information System
Requirements Analysis

Application Identification #: 48
Application Name: Job Training Site Location
Department: Social Services
Defined by: Jeff Embury (C/O Jim Kubacki)

Purpose and Description: To provide trainees with an adequate training site while minimizing the distance they must travel to reach that site.

Type of Application:

Display	Display/Map Scale: Multiple
Query	Query Key: Trainee Address
Query & Display X	Response Time: < 1 Minute
Map Analysis	Frequency: Daily
Spatial Model	

Data Required:

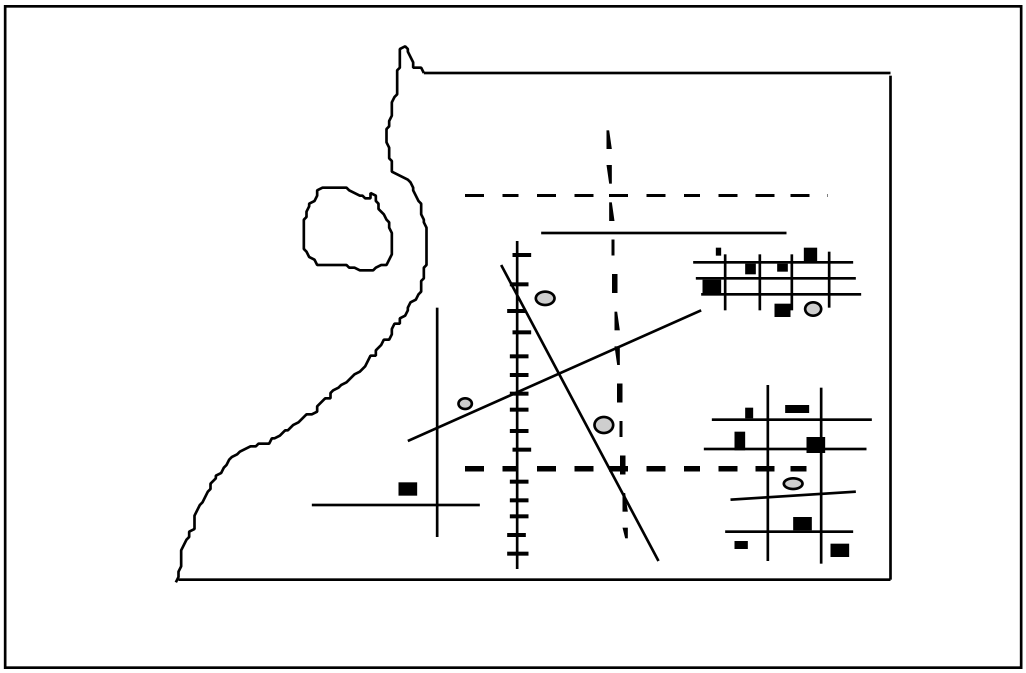
Features (entities):	Attributes:
ROADS	XY_Location, Name, Address_Range
TRAINEE	Trainee_Address, Trainee_Name
TRAINING SITE	Site_Name, Site_Address, Site_Phone #
SUBWAY	XY_Location, Subway_Stop_Name, Subway_Stop_Location
BUS ROUTES	XY_Location, Busroute_Number, Bus_Stop_Name, Bus_Stop_Location

Prepared by: Approved by: Date:

County of Erie
Geographic Information System
Requirements Analysis

Application Identification #: 48
Application Name: Job Training Site Location
Department: Social Services
Defined by: Jeff Embury/Jim Kubacki

Graphical Output Sample: *Screen:* *X* *Hard Copy:*



Symbols/Legend

■ Trainee ● Job Training Site — Road – – Bus ┼┼┼ Subway

Prepared by: Eric Covino Approved by: Date: 5/5/94

Data Flow Diagram

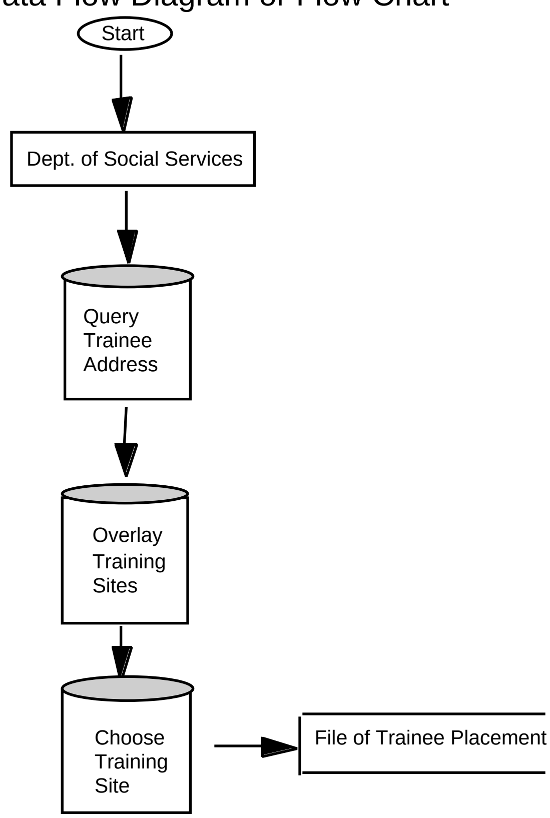
Name of Government

Geographic Information System

Requirements Analysis

Application Identification #: 48
Application Name: Job Training Site Location
Department: Social Services
Defined by: Jeff Embury (C/O JimKubacki)

Process Description: Data Flow Diagram or Flow Chart



Prepared by: Approved by: Date:

Data Flow Diagram

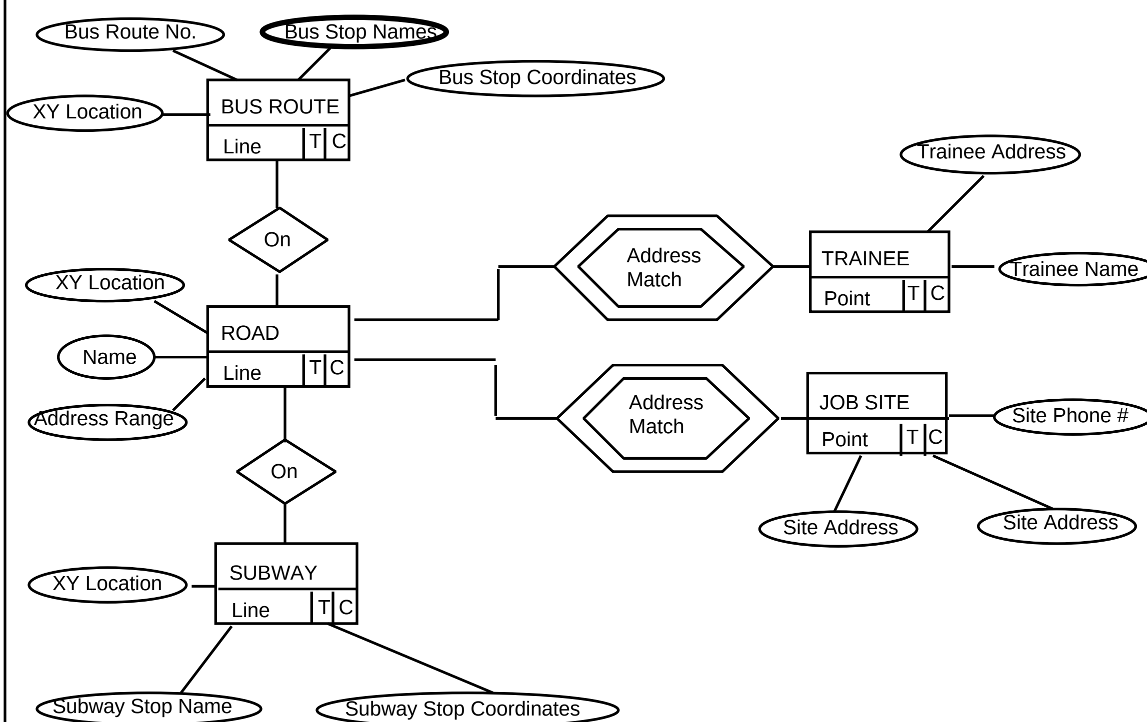
Name of Government

Geographic Information System

Requirements Analysis

Application Identification #:	48
Application Name:	Job Training Site Location
Department:	Social Services
Defined by:	Jeff Embury (C/O Jim Kubacki)

Process Description: Data Flow Diagram or Flow Chart



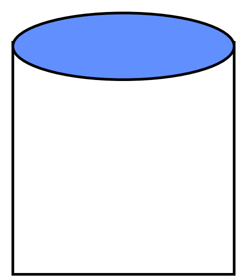
Prepared by: Approved by: Date:

Data Flow Diagraming

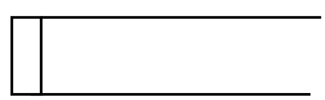
Data flow diagrams offer a standardized method of portraying processes, data stores, and participants that make up a logical activity potential GIS application). Four symbols are used in a data flow diagram:

A square represents people, organizations, things, or sources or destinations of data or information

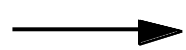
A cylindrical shape to represent a process or activity



An open rectangle to represent a data stored from which data can be added or removed



An arrow to represent data flows. Arrow can be annotated as necessary to describe nature or content flow.



List of Application Name, Type, and Frequency

Appl#	Application Name	Type	Frequency	
11	Subdivision Development Map	Query & Display	1	per month
12	Counter Query Map	Query	50	per day
13	Land Use/Land Value Map	Display	1	per year
14	Assessed Value Map	Query & Display	3	per year
15	Grievance Map	Query & Display	1650	per year
16	Comparable Property Map	Query & Display	1	per month
17	Built/Vacant Map	Display	1	per year
19	Sanitary Sewer Line Map	Query & Display	2	per week
28	Public Improvement Map	Query & Display	10	per week
29	Total Committed Flow Map	Spatial Model	20	per week
36	Storm Sewer Map	Display	10	per day
37	Youth League and Residency Check Map	Query	1500	per year
41	Optimal Snow Removal Route Map	Spatial Model	10	per month
63	Population Density Map	Browse	50	per year
70	Population Projection	Spatial Model	4	per year

Geographic Information System
Requirements Analysis

Application Identification #: 11

Application Name: Subdivision Development Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To monitor the progress of development of an approved subdivision (how many lots are built and the rate of the building).

Type of Application:
Query & Display

Display/Map Scale: 100;200;400
Response Time:
Frequency: 1 per month

Data Required:

Feature	Spatial Object	Attribute
Parcel	Polygon	
		location
Street (double line)	Polygon	
		name
		location
Subdivision	Polygon	
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 12

Application Name: Counter Query Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To provide a quick query of one or more parcels and the associated parcel data (mostly ARLM file data) for answering inquiries at the counter or over the telephone.

Type of Application:
Query

Display/Map Scale: 50;100;200;400
Response Time:
Frequency: 1 per month

Data Required:

Feature	Spatial Object	Attribute
Building	Polygon	
		assessed value
		building #
Parcel	Polygon	
		subdivision lot
		SBL #
		location
Street (center line)	Line	
		length
		address range
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 13

Application Name: Land Use/Land Value Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To produce a display of the value of land per square foot and/or front footage by land use type.

Type of Application:
Display

Display/Map Scale: 200;400
Response Time: 1 week
Frequency: 1 per year

Data Required:

Feature	Spatial Object	Attribute
Parcel	Polygon	
		depth
		front footage
		size
		last sale price
		land use code
		location
Street (center line)	Line	
		name
		location

Prepared by:

Approved by:

Date: 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 14

Application Name: Assessed Value Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To produce a map showing the assessed values (by range) for a small area; or for designated neighborhoods.

Type of Application:
Query & Display

Display/Map Scale: 400
Response Time: Interactive
Frequency: 3 per year

Data Required:

Feature	Spatial Object	Attribute
Neighborhood	Polygon	
		name
		location
Parcel	Polygon	
		assessed value
		location
Street (double line)	Polygon	
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 15

Application Name: Grievance Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To show assessed values of properties in the same area as a parcel where a grievance is filed.

Type of Application:
Query & Display

Display/Map Scale: 100;200;400
Response Time: interactive
Frequency: 1650 per year

Data Required:

Feature	Spatial Object / Attribute
Parcel	Polygon
	assessed value
	location
Street (double line)	Polygon
	name
	location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 16

Application Name: Comparable Property Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

Show the comparable properties selected to determine the assessed value of a given property.

Type of Application:
Query & Display

Display/Map Scale: 100;200;400
Response Time:
Frequency: 1 per month

Data Required:

Feature	Spatial Object	Attribute
Parcel	Polygon	
		address
		assessed value
		location
Street (double line)	Polygon	
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 17

Application Name: Built/Vacant Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To display the built and vacant parcels

Type of Application:
Display

Display/Map Scale: 400
Response Time: Interactive
Frequency: 1 per year

Data Required:

Feature	Spatial Object	Attribute
Occupancy	Node	
		occupant type code
		occupant address
		Occupant name
Parcel	Polygon	
		built/vacant code
		location
Street (double line)	Polygon	
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 19

Application Name: Sanitary Sewer Line Map

Department: Assessor

Defined by: H. Williams

Purpose and Description:

To show the location of sanitary sewer lines for the purpose of approving digging activities.

Type of Application:
Query & Display

Display/Map Scale: 50;1000
Response Time: 5 min.
Frequency: 2 per week

Data Required:

Feature	**Spatial Object**	**Attribute**
Building footprint	Polygon	
		business name
		building name
		address
Manhole	Node	
		depth
		location
Sanitary sewer line	Line	
		location
Sidewalk	Line	
		location
Storm sewer line	Line	
		location
Street (double line)	Polygon	
		name
		location
		address range
Wye hook ups (new only)	Node	
		distance from manholes
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 28

Application Name: Public Improvement Map

Department: Engineering

Defined by: P. Bowers

Purpose and Description:

To show facilities near a certain parcel for review of a public improvement permit or site plan.

Type of Application:
Query & Display

Display/Map Scale: 100;200
Response Time: 30 sec
Frequency: 10 per week

Data Required:

Feature	Spatial Object / Attribute
Parcel	Polygon
	location
Sanitary sewer	Line
	location
Storm drainage	Line
	location
Street (double line)	Polygon
	curb location
	pavement type
	location
Water main	Line
	installation date
	material
	size
	type
	location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 29

Application Name: Total Committed Flow Map

Department: Engineer

Defined by: P. Bowers

Purpose and Description:

To keep track of the total committed flow of sanitary and storm sewers.

Type of Application:
Spatial Model

Display/Map Scale: 1000
Response Time: 1 min
Frequency: 20 per week

Data Required:

Feature	Spatial Object	Attribute
Detention pond	Polygon	
		capacity
		size
		location
Ditches	Polygon	
		capacity
		size
		location
Monitoring point	Node	
		location
Sanitary sewer line	Line	
		capacity
		size
		location
Storm sewer line	Line	
		capacity
		size
		location
Street (center line)	Line	
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 36

Application Name: Storm Sewer Map

Department: Building Department

Defined by: T. Ketchum

Purpose and Description:

To display the location of storm sewers.

Type of Application:
Display

Display/Map Scale: 100;200
Response Time: 12 sec
Frequency: 10 per day

Data Required:

Feature	Spatial Object	Attribute
Contours	Line	
		location
		elevation
Easement	Polygon	
		location
		type
Manhole	Node	
		location
		invert elevation
		rim/surface elevation
Parcel	Polygon	
		location
Storm sewer line	Line	
		location
Street (double line)	Polygon	
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 37

Application Name: Youth League and Residency Check Map

Department: Recreation Department

Defined by: J. Bloom

Purpose and Description:

To determine the appropriate league for a resident (by parcel) and discover non-resident applications.

Type of Application:
Query

Display/Map Scale: 1000
Response Time: 30 sec
Frequency: 1500 per year

Data Required:

Feature	Spatial Object	Attribute
League	Polygon	
		type
		location
Parcel	Polygon	
		owner address
		owner name
		land use
		address
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 41

Application Name: Optimal Snow Removal Route Map

Department: Highway Department

Defined by: F. Jurgens

Purpose and Description:

To calculate the most efficient routes for snow removal and salting.

Type of Application:
Spatial Model

Display/Map Scale: 1000
Response Time: 1 week
Frequency: 10 per month

Data Required:

Feature	Spatial Object / Attribute
Street (center line)	Line
	length
	class
	width
	location
Street intersections	Node
	traffic flow conditions
	street names
Traffic zone	Polygon
	area
	zone code

Prepared by:

Approved by:

Date: 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 63

Application Name: Population Density Map

Department: Planning

Defined by: G. Black

Purpose and Description:

To browse population density by census tract, block group, or block.

Type of Application:
Browse

Display/Map Scale: variable: 200 to 1000
Response Time: Interactive
Frequency: 1 per year

Data Required:

Feature	Spatial Object	Attribute
Census Block	Polygon	
		block #
		population total
		location
Census tract	Polygon	
		tract #
		location
Parcel	Polygon	
		area
		land use code
		location
Street (center line)	Line	
		name
		location

Prepared by: **Approved by:** **Date:** 03-June-96

Geographic Information System
Requirements Analysis

Application Identification #: 70

Application Name: Population Projection

Department: Planning

Defined by: C. Brown

Purpose and Description:

To estimate future population of the Town, by small area (census tract, block group, and possibly block).

Type of Application:
Spatial Model

Display/Map Scale: 1000
Response Time: 1 day
Frequency: 4 per year

Data Required:

Feature	Spatial Object	Attribute
Census Block	Polygon	
		block #
		size
		location
Census tract	Polygon	
		tract #
		size
		location
Net migration	None	
		application #69
Wetland	Polygon	
		area
		wetland code
Zoning	Polygon	
		area
		zoning code

Prepared by: **Approved by:** **Date:** 03-June-96

Master Data List

Feature	Attribute	Spatial Object
Building		Polygon
	assessed value	
	building #	
Building footprint		Polygon
	address	
	building name	
	business name	
Census Block		Polygon
	block #	
	location	
	popultion total	
	size	
Census tract		Polygon
	location	
	size	
	tract #	
Contours		Line
	elevation	
	location	
Detention pond		Polygon
	capacity	
	location	
	size	
Ditches		Polygon
	capacity	
	location	
	size	
Easement		Polygon
	location	
	type	
League		Polygon
	location	
	type	
Manhole		Node
	depth	
	invert elevation	
	location	
	rim/surface elevation	
Monitoring point		Node
	location	

Master Data List Cont'd

Neighborhood		Polygon
	location	
	name	
Net migration		None
	application # 69	
Occupancy		Node
	occupant address	
	occupant name	
	occupant type code	
Parcel		Polygon
	address	
	area	
	assessed value	
	built/vacant code	
	depth	
	front footage	
	land use	
	land use code	
	last sale price	
	location	
	owner address	
	owner name	
	SBL #	
	size	
	subdivision lot	
Sanitary sewer line		Line
	capacity	
	location	
	size	
Sidewalk		Line
	location	
Storm drainage		Line
	location	
Storm sewer line		Line
	capacity	
	location	
	size	

Master Data List Cont'd

Street (center line)		Line
	class	
	length	
	location	
	name	
	width	
	address range	
Street (double line)		Polygon
	address range	
	curb location	
	location	
	name	
	pavement type	
Street intersections		Node
	street names	
	traffic flow conditions	
Subdivision		Polygon
	boundary	
	name	
Traffic zone		Polygon
	area	
	zone code	
Water main		Line
	installation date	
	location	
	material	
	size	
	type	
Wetland		Polygon
	area	
	wetland code	
Wye hook ups (new only)		
	distance from manholes	node
	location	
Zoning		Polygon
	area	
	zoning code	

GIS Application by Department by Type

Department	Browse	Display	Query	Query & Display	Spatial Model	Total
Assessor	0	2	1	4	0	7
Building Dept.	0	1	0	0	0	1
Engineering	0	0	0	2	1	3
Highway Dept.	0	0	0	0	1	1
Planning	1	0	0	0	1	2
Recreation Dept.	0	0	1	0	0	1
Total	1	3	2	6	3	15

GIS Application by Department by Frequency

Department	Browse	Display	Query	Query & Display	Spatial Model	Total
Assessor	0	2	18250	1677	0	19929
Building Dept.	0	3650	0	0	0	3650
Engineering	0	0	0	624	1040	1664
Highway Dept.	0	0	0	0	120	120
Planning	50	0	0	0	4	54
Recreation Dept.	0	0	1500	0	0	1500
Total	50	3652	19750	2301	1164	26917

PART 1 - DATA MODELING

1 INTRODUCTION

This guide describes data modeling in general, spatial data modeling in specific, the setting of GISspecifications, and an introduction to spatial data and metadata standards. These activities arecollectively called ***conceptual design of the GIS system*** (Figure 1). This activity takes theinformation developed during the Needs Assessment and places it a structured format. The resultof this activity will be a ***GIS data model and functional specifications for the GIS system.***

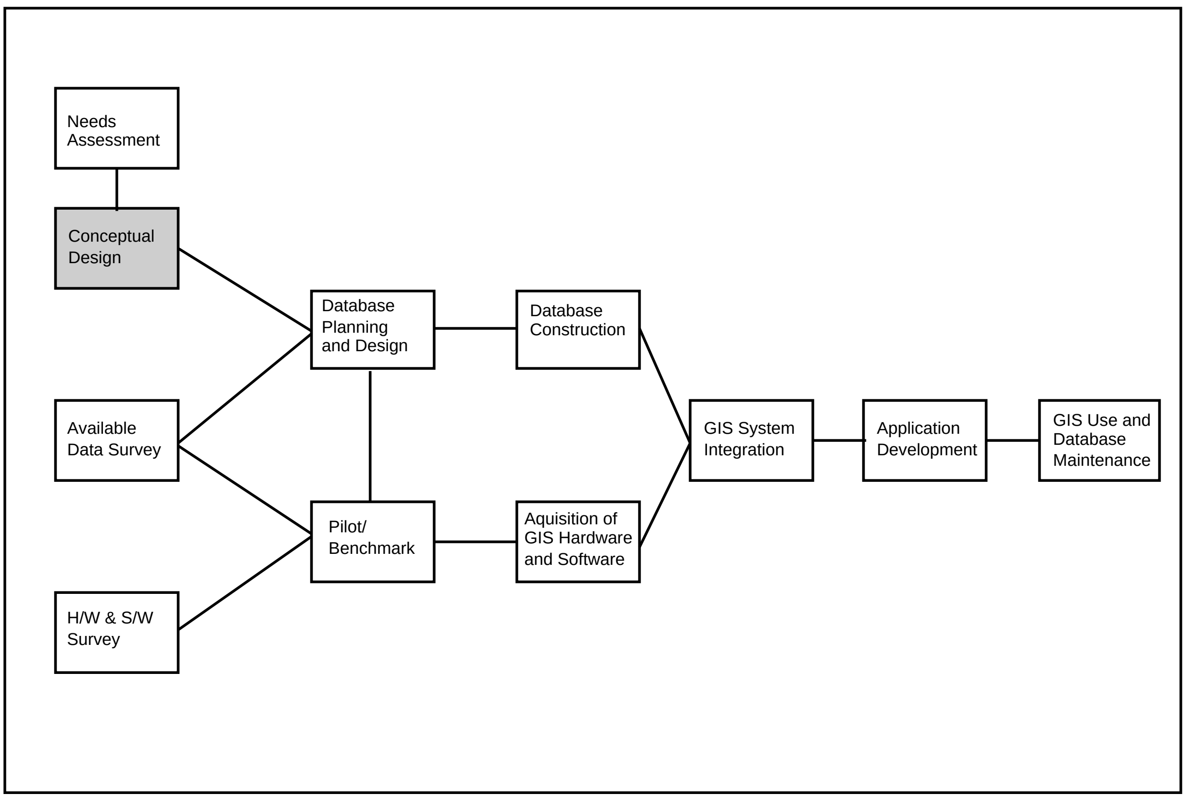


Figure 1 - GIS Development Process

Conceptual design is the first step in ***database design*** where the contents of the intended database are identified and described. Database design is usually divided into three major activities

- Conceptual data modeling: identify data content and describe data at an abstract, or conceptual, level. This step is intended to describe ***what the GIS must do*** and does not deal with ***how the GIS will be implemented*** - the "how" question is the subject of logical and physical database design;
- Logical database design: translation of the conceptual database model into the data model of a specific software system; and

- Physical database design: representation of the logical data model in the schema of the software.

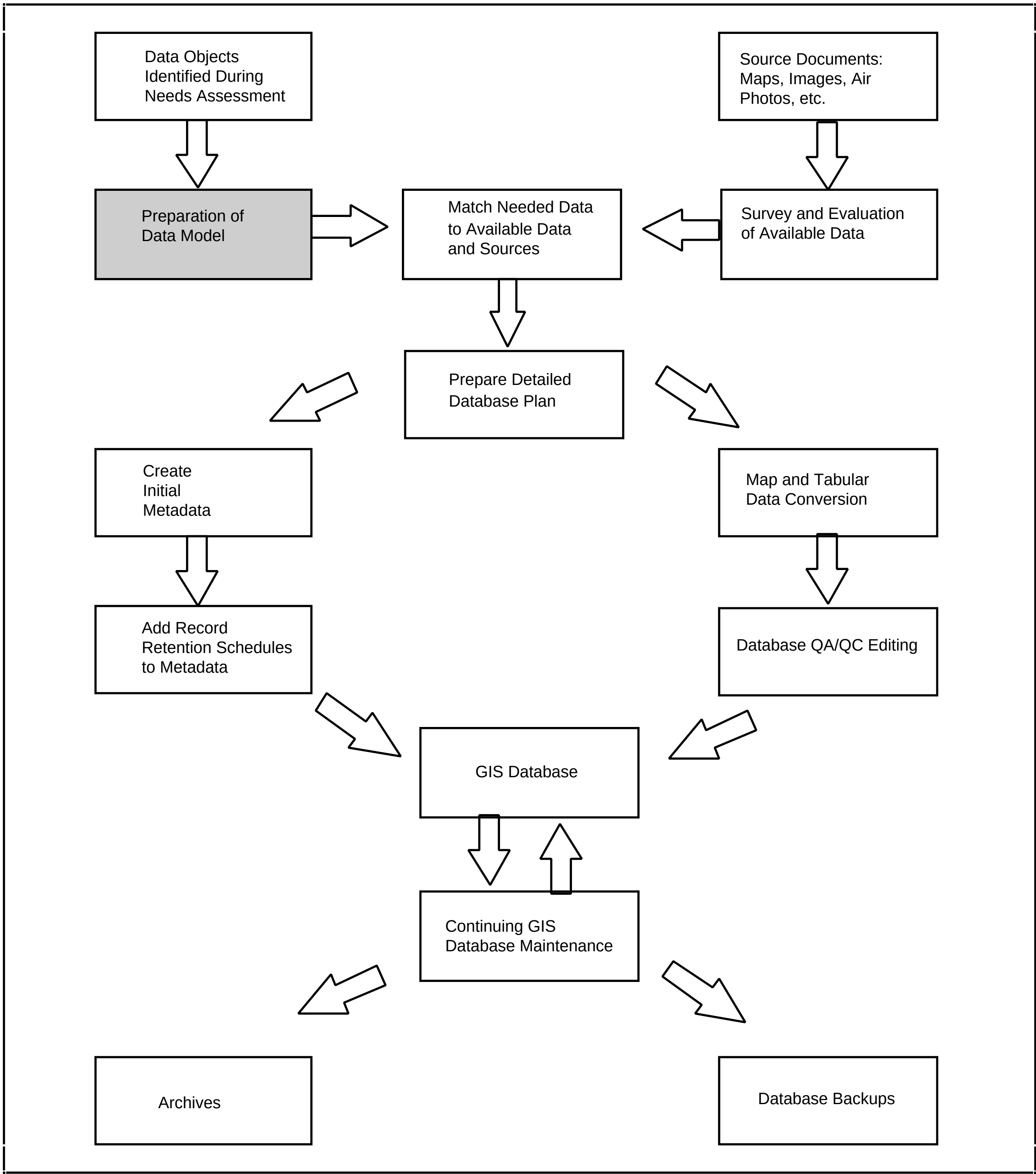


Figure 2 - Life Cycle of a GIS Database

The conceptual design of the GIS system is primarily an exercise in database design. ***Database planning is the single most important activity in GIS development.*** It begins with the identification of the needed data and goes on to cover several other activities collectively termed the

data life cycle - identification of data in the needs assessment, inclusion of the data in the data model, creation of the metadata, collection and entry into the database, updating and maintenance, and, finally, retained according to the appropriate record retention schedule (Figure 2). A complete data plan facilitates all phases of data collection, maintenance and retention and as everything is considered in advance, data issues do not become major problems that must be addressed after the fact with considerable difficulty and aggravation.

The conceptual design of the GIS also includes identification of the basic GIS architecture (functions ofhardware and GIS software), estimates of usage (derived from the needs assessment), and scopingthe size of the GIS system. All of this is done with reference to the existing data processingenvironments (legacy systems) that must interface with the GIS.

Preparing A GIS Data Model

A ***data model*** is a formal definition of the data required in a GIS. The data model can take one of several forms, the two used in this guideline are ***a structured list*** and ***an entity-relationship diagram***. The purpose of the data model, and the process of specifying the model, is to ensure that the data has been identified and described in ***a completely rigorous and unambiguous fashion and that both the user and GIS analyst agree on the data definitions.*** The data model is then the ***formal specification*** for the ***entities***, their ***attributes*** and all ***relationships*** between the entities for the GIS.

Building a data model is not necessarily an easy task. Most professionals in local government will not have had experience in this task. The GIS analyst of the project is the individual who either should build the data model or acquire assistance, such as a qualified consultant, to complete this task. If the opportunity exists for the GIS analyst to attend a database design course or seminar, this would enhance this person's ability to build the model but, more importantly, provide the knowledge for using the final data model in building the GIS. To the extent that data models prepared for other local governments match the needs of a particular GIS development program, or can be easily adapted, they can be modified for use as the data model. However, the GIS analyst must have a good understanding of the resulting model and how it is used to build and manage the GIS database.

The next sections of the guideline first discuss the nature of geographic data, then present themethodology used for data modeling, and lastly describe the development of a GIS data modelfrom the information collected during the Needs Assessment. The example provided in the last section is actually a sample local government GIS data model and is suitable for direct use, withappropriate modification to specific situations.

2 NATURE OF GEOGRAPHIC DATA

Geographic data describe entities which have a location. The geographic data includes the locationinformation and other information about the entity of interest. This other information will bereferred to as attributes of the entities. Historically several terms have been used to describe thedata in a GIS database, among them ***features, objects, or entities***. The term ***feature*** derives fromcartography and is commonly used to identify "features shown on a map," while ***entity*** and ***object*** are terms from computer science used to identify the elements in a database. The normal dictionarydefinitions of these terms are:

> ***Object***: a thing that can be seen or touched; material thing that occupies space
>
> ***Entity***: a thing that has definite, individual existence in reality
>
> ***Feature***: the make, shape, form or appearance of a person or thing

A good GIS database design methodology requires the use of terms in a clear an unambiguousmanner. This guideline will use the term ***entity*** to represent objects or things to be included in thedatabase and ***attribute*** will be the term for representing the characteristics or measurements to berecorded for the entities. Other terms have commonly been used to describe the organization ofentities and attributes in a GIS, such as ***layer, coverage, base map, theme, and others***. Each of thesewill usually refer to a collection of one or more entities organized in some useful way which isspecific to the GIS software in use. These terms will become important during the logical/physicaldatabase design activities where decisions about ***how*** the GIS data are to be stored in the GISdatabase are made. The conceptual database design activity is focused solely on specifying whatis to be included in the GIS database and should provide clear and unambiguous representation ofthe entire GIS database.

In addition to a clear and concise definition of entities and their attributes, data modeling describes ***relationships between entities***. An example of a relationship between an **employee** and a **company** would be "**works for**."

Employee - Works For - Company

Relationships may be bi-directional, thus:

Company - Has - Employees

An important aspect of a relationship is "cardinality," that is if the relationship is between only one of each entity or if either entity may be more than one. For example, one company usually has many employees whereas one employee works for only one company. The possible cardinalities are: one-to-one; one-to-many; and many-to-many. Thus:

```
                          --- Has --->
Company (One)  <---   Work For  ------  (Many) Employees
```

There are many variations of the notation used to express these facts. The notation recommended for local government will be described later.

Geographic, or spatial data, differs from other "regular" data that are included in computerdatabases in how entities are defined and in the relationships between entities. Entity identificationfor spatial data includes the definition of a physical or abstract entity (e.g., a building) and ***thedefinition of a corresponding spatial entity (i.e., a polygon to represent the building footprint).*** This latter, or second entity does not exist for other types of computer databases. The existence ofthe corresponding spatial entity is one of the major factors that distinguishes GIS from other types ofsystems and is what makes it very important to utilize proper planning and design techniques whenbuilding a GIS. An example will be used to illustrate this difference.

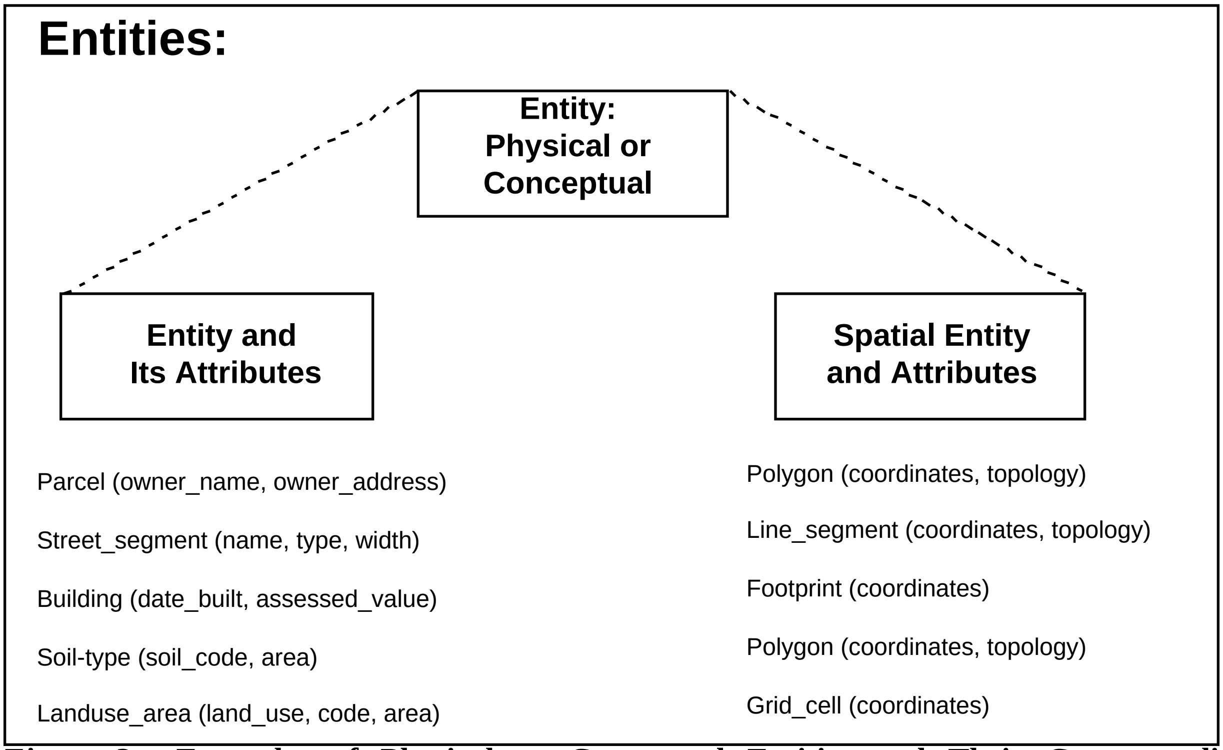


Figure 3 - Examples of Physical or Conceptual Entities and Their Corresponding Spatial Entities

3 ENTITY-RELATIONSHIP (E-R) DATA MODELING

To start the discussion of entity-relationship modeling, two examples will be shown. One, a regular database and the second, a simple GIS database. The personnel database inany local government could have entities of ***employee, dependent and department.*** Relationshipsbetween these entities would be employee " works in" department and dependent "is a member of"the employee's family. Some of the attributes for each entity would be as follows:

Employee (name, age, sex, job title)
Dependent (name, age, relationship_to _employee, i.e., spouse, child, etc.)
Department (department_name, function, size)

A diagrametric representation of the example would be as follows:

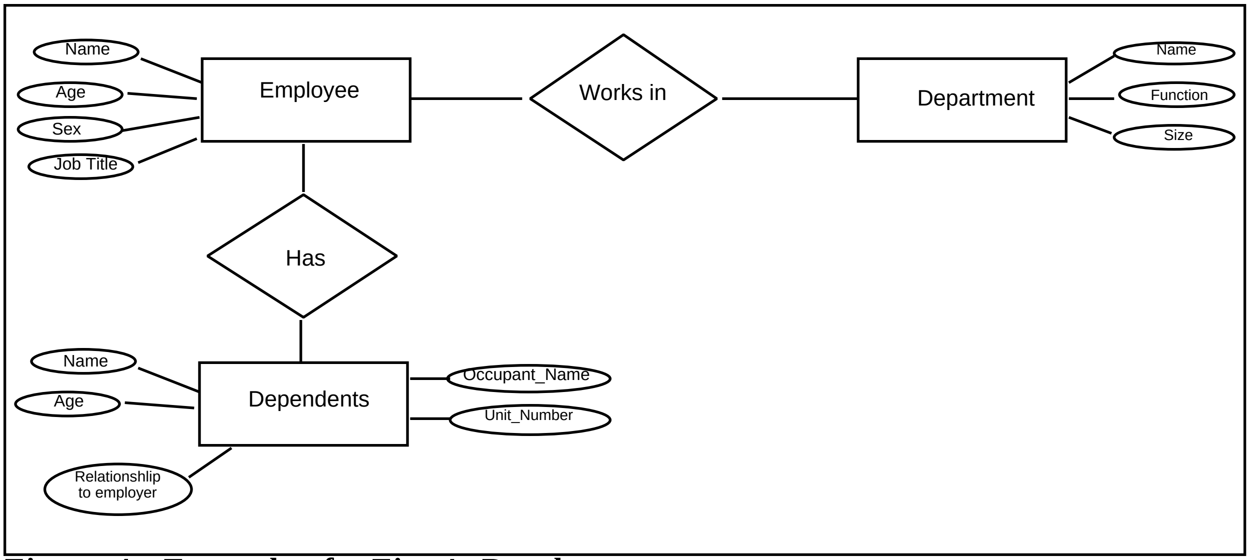


Figure 4 - Example of a Firm's Database

An example of a simple spatial database would be a follows:

Parcel	ID#, owner_name, owner_address, site_address	Polygon
Building	Building_name, height, floor_area	Footprint
Occupant	Occupant_name, unit_number	None

The diagrammatic form of this spatial database would be as follows:

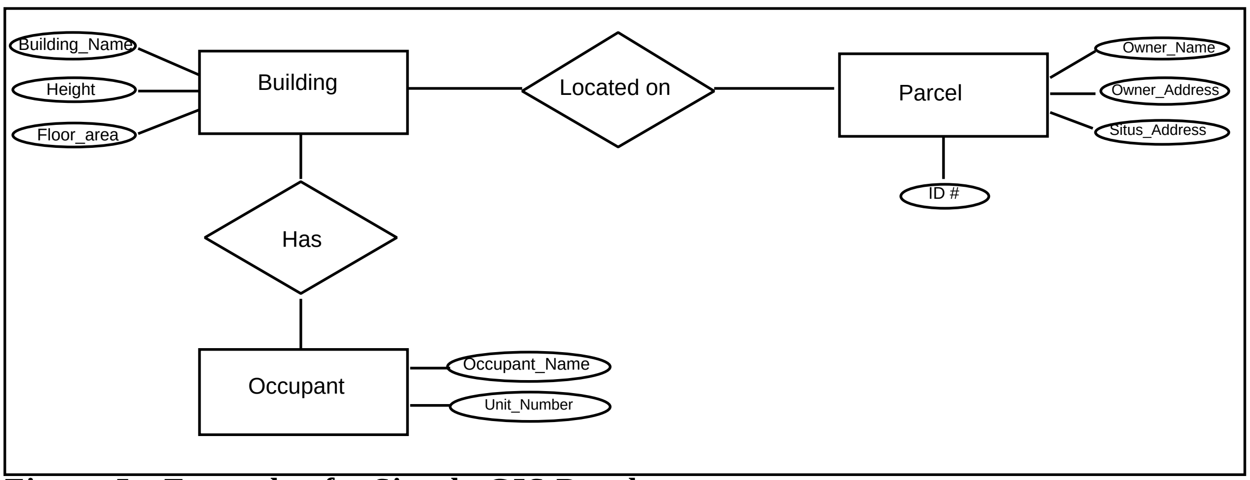


Figure 5 - Example of a Simple GIS Database

This example has been presented using two standard notational forms for conceptual databasedesign: ***a relation***, the entity name followed by a list of attributes; and ***an entity-relationshipdiagram*** showing entities, their attributes, and the relationships between entities. On figure 5, there are two things to notice:

- The standard entity - relationship diagram has no provision for representing the corresponding spatial entity (point, line, polygon) of the data; and
- The representation of the attributes (ellipses) can be somewhat awkward due to different name lengths and the number of attributes to be shown.

The two notational forms modified to accommodate GIS data will be used as the primary tools for conceptual database design in this guideline; however, modifications will be made to adequately represent GIS data.The next section will provide the formal definition of the basic entity-relationship data modelingmethod, the modifications needed to represent GIS data, followed by examples of GIS data entities and attributes typical for local government andthe by a description of how to model these data using the modified entity-relationship data modelingtechnique.

Basic Entity-Relationship Modeling

The basic entity-relationship modeling approach is based on describing data in terms of the three parts noted above (Chen 1976):

- Entities
- Relationships between entitles
- Attributes of entitles or relationships

Each component has a graphic symbol and there exists a set of rules for building a graph (i.e., an E-R model) of a database using the three basic symbols. Entities are represented as ***rectangles,*** relationships as ***diamonds*** and ***attributes*** as ellipses.

The normal relationships included in a E-R model are basically those of:

1. Belonging to;
2. Setand subset relationships;
3. Parent-child relationships; and
4. Component parts of an object.

Theimplementation rules for identifying entities, relationships, and attributes include an Englishlanguage sentence structure analogy where the nouns in a descriptive sentences identify entities,verbs identify relationships, and adjectives identify attributes. These rules have been defined byChen (1983) as follows:

> Rule 1: A common noun (such as person, chair), in English corresponds to an entity type on an E-R diagram.
>
> Rule 2: A transitive verb in English corresponds to a relationship type in an E-R diagram.
>
> Rule 3: An adjective in English corresponds to an attribute of an entity in an E-R diagram.

English statement: Mr. Joe Jones resides in the Park Avenue Apartments which is located on land parcel #01-857-34 owned by the Apex Company.

Analysis: .. "Joe Jones"," "Park Avenue Apartments," "land parcel" and "Apex Company" are nouns and therefore can be represented as entities "occupant," "building," "parcel," and "owner." "resides," "located on" and "owned by" are transitive verbs (or verb phrases) and therefore define relationships.

Example of Simple E-R diagrams

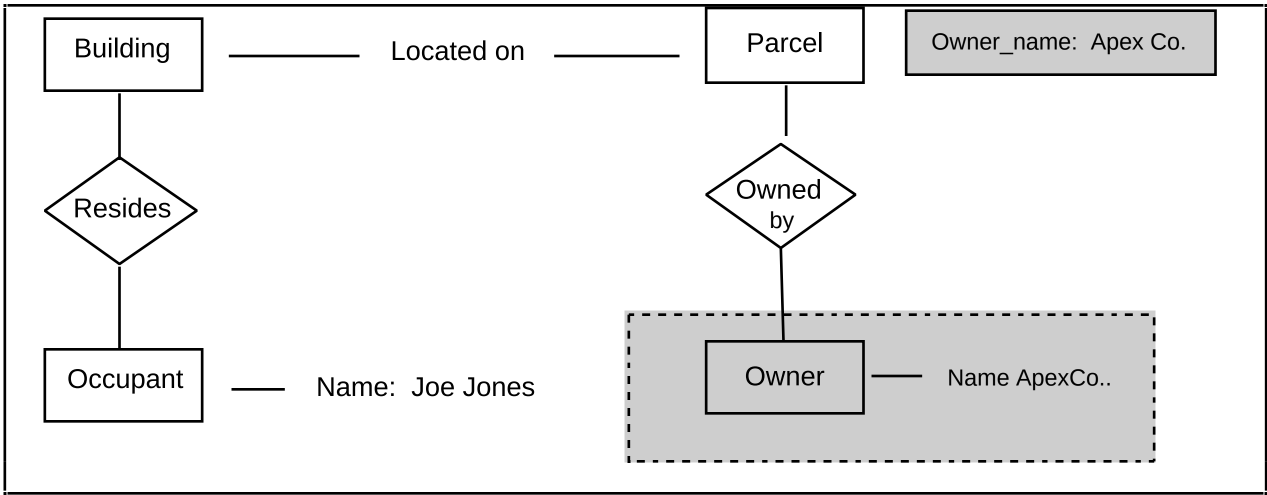


Figure 6 Shows a Simple E-R Diagrams of the Previous Example

Many times it is possible to build different E-R diagrams for the same data. For example, instead of creating the entity "owner" in the above example, the owner's name could be an attribute of parcel (shaded areas of figure 6). During the process of building an E-R diagram (i.e., the conceptual model) for a database, the analyst must make decisions as to whether something is best represented as an entity or as an attribute of some other entity.

The process of constructing an E-R diagram uncovers many inconsistencies or contradictions in the definition of entities, relationships, and attributes. Many of these are resolved as the initial E-R diagram is constructed while others are resolved by performing a series of transformations on the diagram after its initial construction. The final E-R diagram should be totally free from definitional inconsistencies and contradictions. If properly constructed, an E-R diagram can be directly converted to the logical and physical database schema of the relational, hierarchical or network type database for implementation.

Unique Aspects of Geographic Data

In the simplest terms, we think of geographic data as existing on maps as points, lines and areas. Early GIS systems were designed to digitally encode these spatial objects and associate one or morefeature codes with each spatial feature. Examples would be a map of land use polygons, a set ofpoints showing well locations, a map of a stream shown as line segments. For the purposes ofplotting (redrawing the map) a simple data structure linking (x,y) coordinates to afeature code was sufficient.

Topology

A distinguishing feature of a modern GIS is that some spatial relationships between spatial entities will be coded in the database. This coding is termed ***topologically coding.*** Topologyis based on graph theory, where a diagram can be expressed as a set of nodes and links in a mannerthat shows logical relationships. Applied to a map, this concept is used to abstract the featuresshown on the map and to represent these features as nodes and arcs (point and lines). Nodes are the end-points of arcs and areas are formed by a set of arcs. If the concept and definitions of topologic data structures are not familiar to the reader, the following readings are recommended:

- **Geographic Information Systems: A Guide to the Technology,** by John Antenucci, et. al, pages 98-99.

- **Fundamentals of Spatial Information Systems**, by Robert Lauring and Derek Thompson, pages 206-211.

- **ARC/INFO Data Model, Concepts, & Key Terms**, by Environmental Systems Research Institute, Inc., pages 1-12 to 1-15.

Coordinate strings without topology with associated feature codes were called "spaghetti" files because there was not any relationship between any two coordinate strings formally encoded in the database. For example, the "GIS system" would not "know" if two lines intersect or not or whether they had common end points. These relationships could be seen by the human eye if a plot were to be made or alternatively could be calculated (often a time consuming process). Typical of this type of geographic data file are those produced by computer-aided drafting systems (CAD), or known as .dxf, .dwg, or .dgn files.

4 GEOGRAPHIC DATA MODELS

The data models in most contemporary GISs are still based on the cartographic view. Other data models have begin to evolve, but are still very limited. Current and potential geographic data models include:

- The cartographic data model: points, lines and polygons (topologically encoded) with one, or only a few, attached attributes, such as a land use layer represented as polygons with associated land used code

- Extended attribute geographic data mode: geometric objects as above but with many attributes, such as census tract data sets;

- Conceptual object/spatial data model: explicit recognition of user defined objects, zero or more associated spatial objects, and sets of attributes for reach defined object (example: user objects of land parcel, building, and occupant, each having its own set of attributes but with different associated spatial objects: polygon for land parcel, footprint for building, and no spatial object of occupant).

- Conceptual objects/complex spatial objects: multiple objects and multiple associated spatial objects (example: a street network with street segments having spatial representations of both line and polygon type and street intersections having spatial representations of both point and polygon type).

Current GIS are based on the cartographic and extended attribute data models. The trend to object-oriented computer systems and databases will require that GIS planners view their databases from an "object viewpoint."

Spatial Relationships

GISs also differ from other systems in that they include spatial relationships. These relationships are included in the GIS either by the topologic coding or by means of calculations based on the (x,y) coordinates. One common calculation is whether or not two lines intersect. Table 1 shows the spatial relationships, associated descriptive verbs, and the common implementation of each relationship by a GIS.

Spatial Relationships

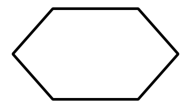
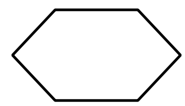
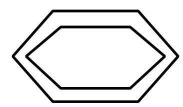
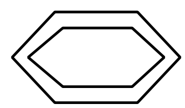
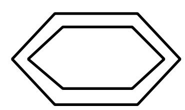
Spatial Relationship	Descriptive Verbs	Common GIS Implementation	E-R Model Symbol
Connectivity	Connect, link	Topology	
Contiguity	Adjacent, abutt	Topology	
Containment	Contained, containing, within	X, Y coord. operation	
Proximity	Closest, nearest	X,Y coord. operation	
Coincidence	Coincident, Coterminous	X,Y coord. operation	

Figure 7 - Spatial Relationships in a GIS

Connectivity and contiguity are implemented through topology: the link-node structure for connectivity through networks and the arc-polygon structure for contiguity. Containment and proximity are implemented through x,y coordinates and related spatial operations: containment is determined using the point-, line-, and polygon-on-polygon overlay spatial operation and proximity is determined by calculating the coordinate distance between two or more x,y coordinate locations. The spatial relationship of coincidence may be complete coincidence or partial coincidence. The polygon-on-polygon overlay operation in ARC/INFO™ calculates partial coincident of polygons in two different coverages. The System 9™ Geographic Information System recognizes coincident features through a "shared primitive" concept (the geometry of a point or line is stored only once and then referenced by all features sharing that piece of geometry). Future versions of commercial GISs will likely implement coincident features through either the "belonging to" database relationship or through x,y coordinates and related spatial operations, whichever is more efficient within the particular GIS.

In summary, there are three types of relationships that will be represented in a geographic database with an "object view" orientation:

- Normal database relationships, which are represented in a relational database by means of keys (primary and secondary)
- Spatial relationships represented in the GIS portion of the database by topology
- Spatial relationships that exist only after a calculation is made on the (x,y) coordinates

5 METHODOLOGY FOR MODELING

Modeling a geographic database using the E-R approach requires an expanded or extended concept for:

- Entity identification and definition; and
- Relationship types and alternate representational forms for spatial relationships.

There are three considerations in the identification and definition of entities in a geographic database:

Correct identification and definition of entities

Entities in a geographic database are defined as either discrete objects (e.g., a building, a bridge, a household, a business, etc.) or as an abstract object defined in terms of the space it occupies (e.g., a land parcel, a timber stand, a wetland, a soil type, a contour, etc.). In each of these cases we are dealing with entities in the sense of "things" which will have attributes and which will have spatial relationships between themselves. These "things" can be thought of as "regular" entities.

Defining a corresponding spatial entity for each "regular" entity

A corresponding spatial entity will be one of the spatial data types normally handled in a GIS, e.g., a point, line, area, volumetric unit, etc. The important distinction here is that we have a single entity, its spatial representation and a set of attributes; we do not have two separate objects (Figure 3 illustrates this concept). A limited and simple set of spatial entitles may be used, or alternatively, depending on the anticipated complexity of the implemented geographic information system, an expanded set of spatial entities may be appropriate. The corresponding spatial entity for the regular entity may be implied in the definition of the regular entity, such as abstract entities like a wetland where the spatial entity would normally be a polygon, or a contour where the spatial entity would be a line. Other regular entitles may have a less obvious corresponding spatial entity. Depending on the GIS requirements, the cartographic display needs, the implicit map scale of the database and other factors, an entity may be reasonably represented by one of several corresponding spatial entities. For example, a city in a small-scale database could have a point as its corresponding spatial entity, while the same city would have a polygon as its corresponding spatial entity in a large-scale geographic database.

Recognize multiple instances of geographic entities, both multiple spatial instances and multiple temporal instances

Multipurpose (or corporate) geographic databases may need to accommodate multiple corresponding spatial entities for some of the regular entities included in the GIS. For example, the representation of an urban street system may require that each street segment (the length of street between two intersecting streets) be held in the GIS as both a single-line street network to support address geocoding, network based transportation modeling, etc., and as a double-line (or polygon) street segment for cartographic display, or to be able to locate other entities within the street segment (such as a water line), etc. In each of these instances the "regular" entity is the street segment, although each instance may have a different set of attributes and different corresponding spatial entities. Also, there may be a need to explicitly recognize multiple temporal instances of regular entities. The simple case of multiple temporal instances will be where the corresponding spatial entity remains the same, however, future GISs will, in all likelihood, have to deal with multiple temporal instances where the corresponding spatial entity changes over time.

Three symbols are defined to represent entities: entity (simple); entity (multiple spatial representations); and entity (multiple time periods). The internal structure of the entity symbol contains the name of the entity and additional information indicating the corresponding spatial entity (point, line or polygon), a code indicating topology, and a code indicating encoding of the spatial entity by coordinates (Figure 8). The coordinate code is, at the present time, redundant in that all contemporary GISs represent spatial entities with x,y coordinates. However, it is possible that future geographic databases may include spatial entities where coordinates are not needed. Similarly, topological encoding is normally of only one type and can, for the present, be indicated by a simple code. However, different spatial topologies have been defined and may require different implementations in a GIS (Armstrong and Densham, 1990). In the future, the topology code may be expanded to represent a specific topologic structure particular to a GIS application.

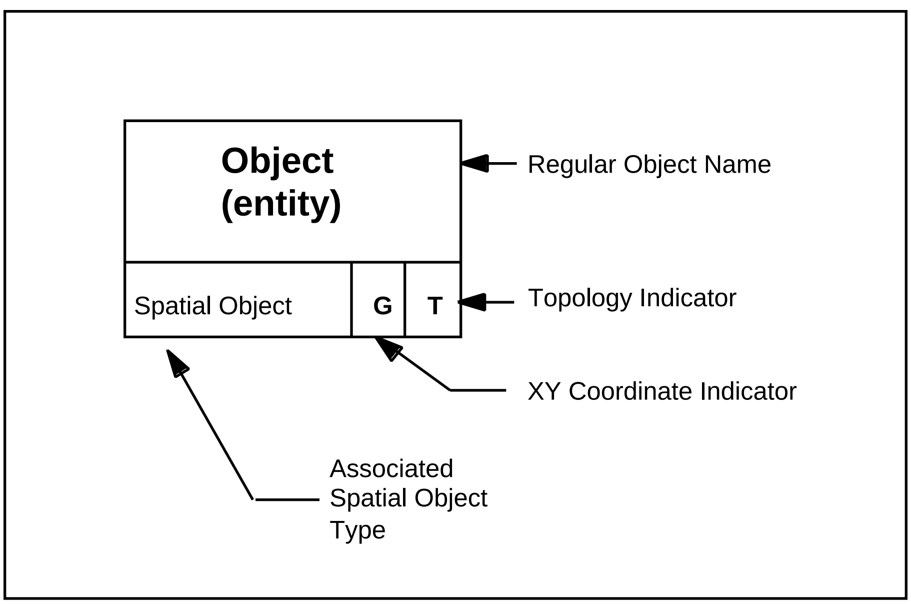


Figure 8 - Entity Symbol for Spatial Objects

Modeling Spatial Relationships

The spatial relationships are defined by three relationship symbols (Figure 9). The traditional diamond symbol can be used for normal database relationships. An elongated hexagon and a double elongated hexagon, are defined to represent spatial relationships. The elongated hexagon represents spatial relationships defined through topology (connectivity and contiguity) and the double elongated hexagon represents spatial relationships defined through x,y coordinates and related spatial operations (coincidence, containment and proximity). The appropriate "verbs" to include in the hexagonal symbols are the descriptors of the spatial relationships (as shown in Figure 7). The spatial operation will be implicitly defined by the relationship symbol (double hexagon), the spatial entity and the topology code. For example, a spatial relationship named "coincident" between entities named "wetlands" and "soils," both of which carry topologic codes and x,y coordinates, indicates the spatial operation of topological overlay. If this does not sufficiently define the spatial operation needed, the name of the spatial operation can be used to describe the relationship, such as shortest path, point-in-polygon, radial search, etc.

Figure 9 shows all symbols needed to construct an Entity-Relationship diagram for a GIS database.

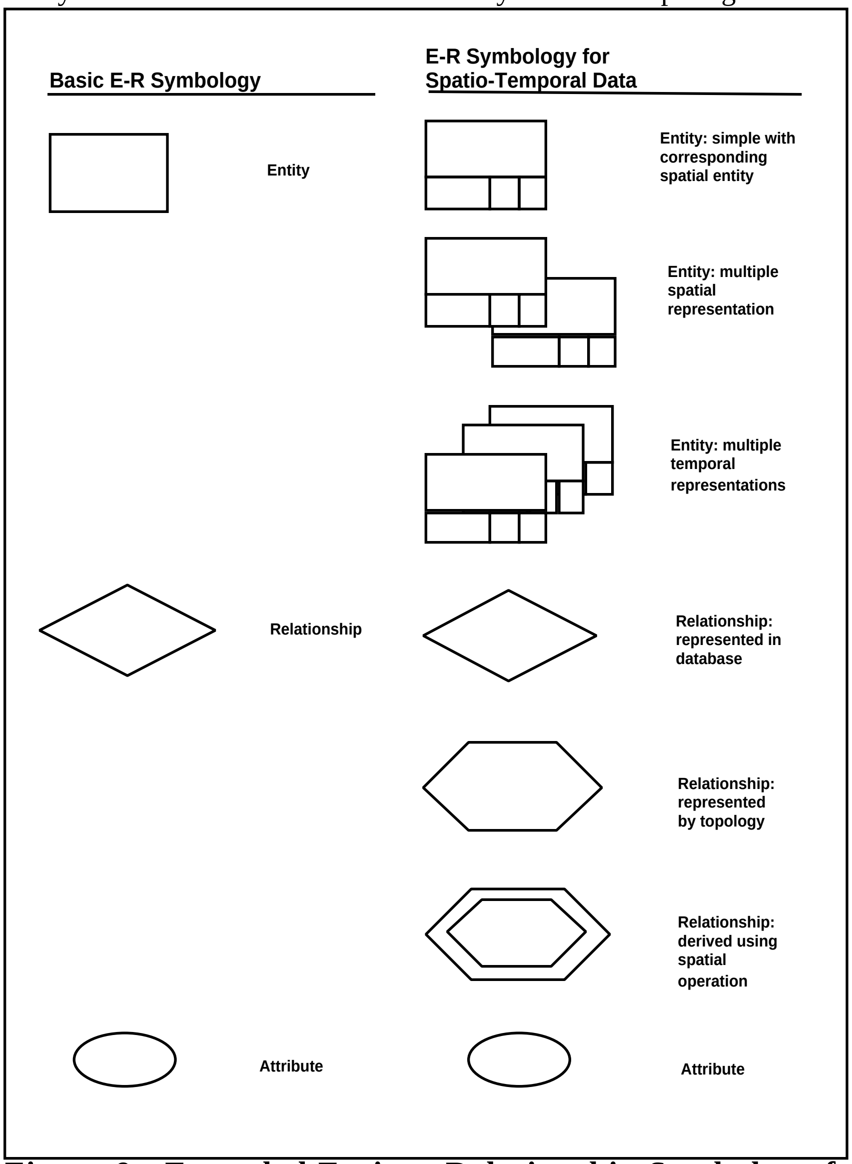


Figure 9 - Extended Entity - Relationship Symbology for Designing GIS Databases, Source: Calkins, 1996

6 DEVELOPING A SPATIAL DATA MODEL (ENTITY-RELATIONSHIP DIAGRAM)

The information needed to develop the E-R diagram representing the spatial data model comes from the Needs Assessment activity as:

- The GIS application descriptions
- The master data list: Lists, entities, corresponding spatial entities and attributes
- The list of functional capabilities (spatial operations)

The process of building the E-R diagram involves taking entities from the master data list one at a time and placing each one on the diagram. For each new entity, any relationship to any previously entered entity should be entered. Relationships are found by examining the Application Descriptions and determining if the GIS processes require a specified operation. For example, if an Application Description indicated that land parcels needed to be compared to a flood plain area, then a spatial relationship of "coincident area" (or topological overlay operation) should be defined between the two entities.

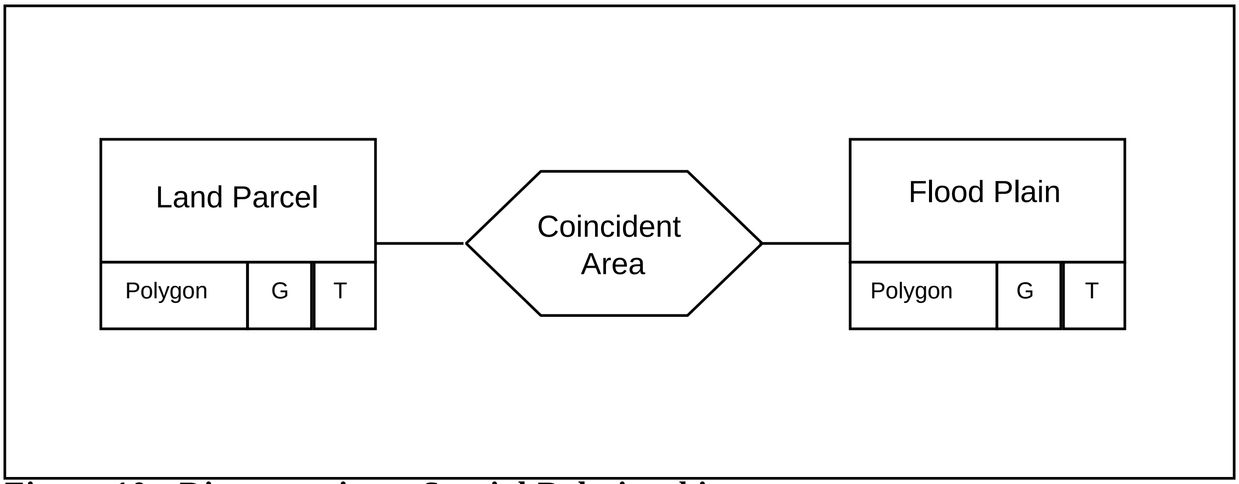


Figure 10 - Diagramming a Spatial Relationship

As each entity is added to the E-R diagram, the list of attributes should be reviewed and checked to determine if the attribute is appropriate for the entity, does not duplicate any other attribute or entity, and can be rigorously defined for entry to create the metadata (metadata is discussed in the next section of this guideline).

Figure 11 is a sample E-R diagram for data commonly used by local government. This example contains 16 entities and 15 relationships. Attributes have not been included in the diagram in order to reduce the size of the diagram for inclusion in this document.

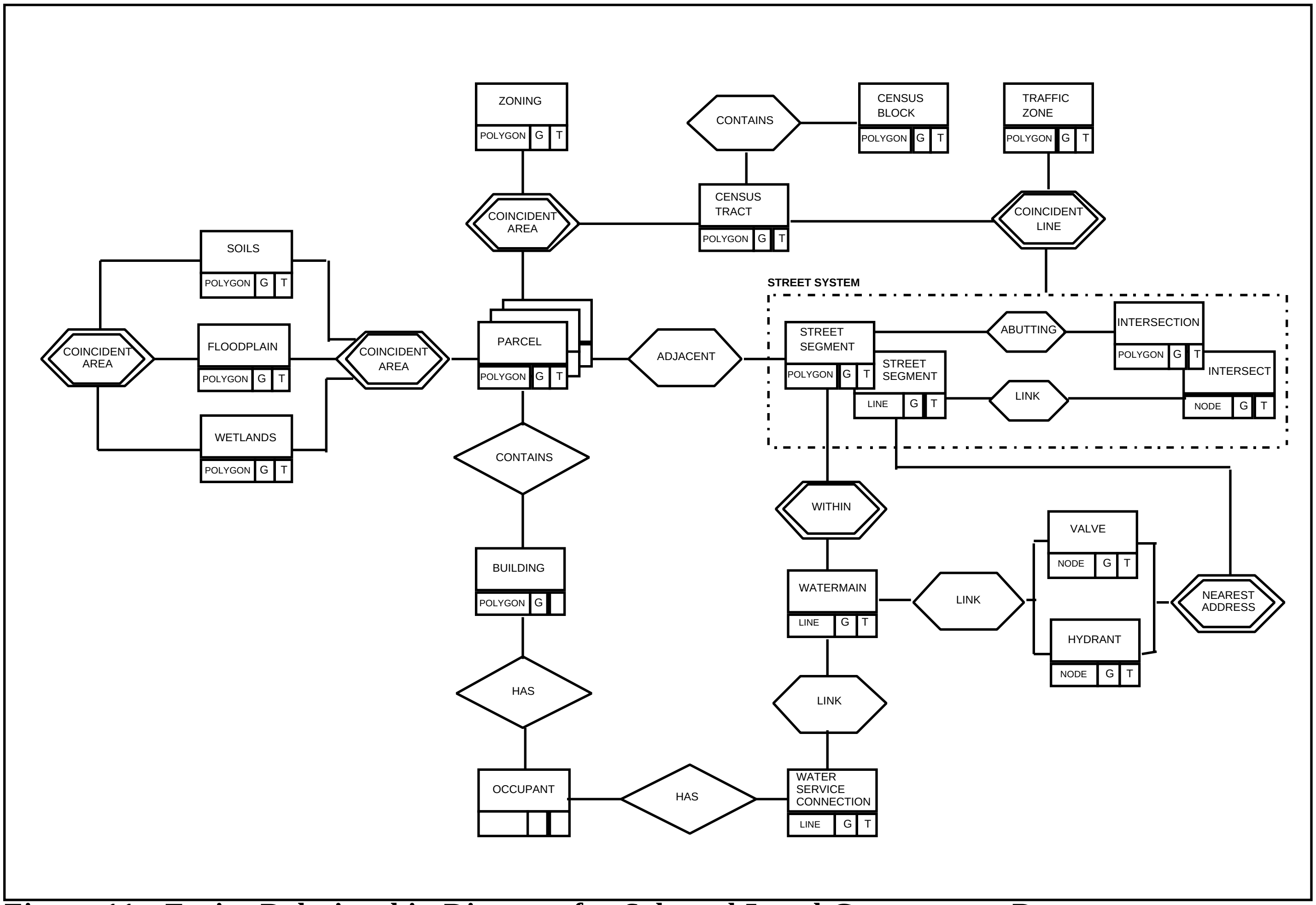


Figure 11 - Entity Relationship Diagram for Selected Local Government Data

7 SUMMARY OF CONCEPTUAL DATA MODELING

The E-R diagram shown in Figure __ will be used to verify with the expected users the data content of the GIS and, by additional reference to the GIS needs analysis, the required spatial operations. Once verified by the users, the E-R representation can be mapped into a detailed database design (as will be described in the Database Planning and Design Guideline)where:

1) Each entity and its attributes map into:
   (a) One or more relational tables with appropriate primary and secondary keys (this assumes the desired level of normalization has been obtained);
   (b) The corresponding spatial entity for the "regular" entity. As most commercial GISs rely on fixed structures for the representation of geometric coordinates and topology, this step is simply reduced to ensuring that each corresponding spatial entity can be handled by the selected GIS package;
2) Each relationship into:
   (a) Regular relationships (diamond) executed by the relational database system's normal query structure. Again, appropriate keys and normalization are required for this mapping.
   (b) Spatial relationships implemented through spatial operations in the GIS. The functionality of each spatial relationship needs to be described, and if not a standard operation of the selected GIS, specifications for the indicated operation need to be written.

PART 2: SPATIAL DATA STANDARDS AND METADATA REQUIREMENTS

Introduction

Spatial data standards cover a variety of topics including the definition of spatial data entities (including a formal data model), methods of representation of the spatial entities in a GIS, specifications for the transfer of spatial data between different organizations, and the definition of the attributes of the spatial entities and the values these attributes may assume. Metadata is "information about data," and should describe the characteristics of the data such as identifying entities and attributes by their standard names and provide information on such items as data accuracy, data sources and lineage, and data archiving provisions.

Much of the work on spatial data standards to date as been done under the auspices of the Federal Geographic Data Committee and only concerns federal spatial data directly. The relationship between the existing federal data standards and state and local spatial data standards have yet to be developed. Appendix A contains a list of current and pending reports on federal spatial data standards. Work towards New York State spatial data standards will be conducted under the proposed GIS Standing Committee of the Information Resources Management Task Force.

Metadata for Local Governments in New York State

Metadata can serve many purposes. Some of the more important functions of metadata are:

- Provide a basic description of a dataset
- Provide information for data transfers to facilitate data sharing
- Provide information for entries into clearinghouses to catalogue the availability of data

The metadata structure and content for local government recommended in this guideline has been prepared according to the following criteria:

- The metadata must first, and primarily, serve as a documentation and data management tool for the data administrator in an agency or department
- Secondly, the metadata must encompass and support the data manager and records management officer in a local agency in all aspects of data management including data definition, source documentation, management and updating, and data archiving and retention requirement.
- Thirdly, the metadata information must be able to generate and supply database descriptions for spatial data clearinghouses such as the prototype New York State Spatial Data Clearinghouse developed under the GIS Demonstration Project conducted by the Center for Technology in Government, SUNY - Albany and any relevant federal spatial data clearinghouses.

The following metadata information is a prototype for a New York State Local Government Spatial Metadata Standard. This metadata is represented in a set of tables listed below and has been implemented in Microsoft Access™. A working copy of this metadata program is available to all local governments in the state. The structure and information on how to use the software are described in a user's guide available with the program. The content of the metadata tables is as shown in the following lists.

Metadata Tables

1. Organization Information
    - Name Of Organization
    - Department
    - Room/Suite #
    - Number And Street Names
    - City
    - State
    - Zip Code
    - Phone Number
    - Fax Number
    - Contact Person
    - Phone Number/Extension
    - Email Address
    - Organization Internet Address
    - Comments

2. Reference Information
    - Filename
    - File Format
    - Availability
    - Cost
    - File Internet Address
    - Metadata Created By
    - Date Metadata Created
    - Metadata Updated By
    - Date Metadata Updated
    - Metadata Standard Name
    - Comments

3. Object/File Name Information
    - Filename
    - Data Object Name

4. Data Object Information
    - Distribution Filename (Same as Filename in Reference Information)
    - Data Object Name
    - Type
    - Data Object Description
    - Spatial Object Type
    - Comments

5. Attribute Information
    - Data Object Name
    - Data Attribute Name
    - Attribute Description
    - Attribute Filename
    - Codeset Name/Description
    - Measurement Units
    - Accuracy Description
    - Comments

6. Data Dictionary Information
    - Data Object Name
    - Data Attribute Name
    - Data Type
    - Field Length
    - Required
    - Comments

7. Spatial Object Information
    - Data Object Name
    - Spatial Object Type
    - Place Name
    - Projection Name/Description
    - HCS Name
    - HCS Datum
    - HCS X-Offset
    - HCS Y-Offset
    - HCS Xmin
    - HCS Xmax
    - HCS Ymin
    - HCS Ymax
    - HCS Units
    - HCS Accuracy Description
    - VCS Name
    - VCS Datum
    - VCS Zmin
    - VCS Zmax
    - VCS Units
    - VCS Accuracy Description
    - Comments

8. Source document information
    - Data Object Name
    - Spatial Object Type
    - Source Document Name
    - Type
    - Scale
    - Date Document Created
    - Date Digitized/Scanned
    - Digitizing/Scanning Method Description
    - Accuracy Description
    - Comments

9. Lineage Information
    - Data Object Name
    - Data Object 1
    - Data Object 2
    - Description of Spatial Operation and Parameters
    - Accuracy Description
    - Comments
10. Update Information
    - Data Object Name
    - Update Frequency
    - Date
    - Updated By
    - Comments

11. Archive Information
    - Data Object Name
    - Retention Class
    - Retention Period
    - Data Archived
    - Archived By
    - Date to be Destroyed

12. Source Documents
    - Source Document Name
    - Source Document ID#
    - Source Organization
    - Type of Document
    - Number of Sheets (map, photo)
    - Source Material (paper, mylar)
    - Projection Name
    - Coordinate System
    - Date Created
    - Last Updated
    - Control/Accuracy (map, photo)
    - Scale
    - Reviewed by
    - Review date
    - Spatial extent
    - File format
    - Comments

13. Entities Contained in Source
    - Source ID#
    - Entity Name
    - Spatial Entity
    - Estimated Volume of Spatial Entity
    - Symbol
    - Accuracy Description of Spatial Entity
    - Reviewed by
    - Review Date
    - Scrub Needed (yes/no)
    - Comments

14. Attributes by Entity
    - Source ID#
    - Entity Name
    - Attribute Description
    - Code Set Name
    - Accuracy Description of Attribute
    - Reviewed By
    - Review Date
    - Comments

Additional Reading

(the following material is quite technical, but a good set of sources on conceptual database design.)

Armstrong, M.P. and P.J. Densham, 1990, Database organization strategies for spatial decision support systems, International Journal of Geographical Information Systems, vol. 4, no. 1, 3-20.

Calkins, Hugh W., Entity Relationship Modeling of Spatial Data for Geographic Information Systems, International Journal of Geographical Information Systems, January 1996.

Chen, P.P., 1976, The entity-relationship model - toward a unified view of data, ACM Transactions on Database Systems, vol. 1, no. 1, March 1976, pp. 9-36

Chen, P.P., 1984, English sentence structure and entity-relationship diagrams, Information Sciences, 29, 127-149

Davis, C., et. al., eds., 1983, Entity-Relationship Approach to Software Engineering, Amsterdam, Netherlands: Elsevier Science Publishers B.V.

Elmasri, R. and S.B. Navathe, 1989, Fundamentals of Database Systems, Redwood City, California: The Benjamin/Cummings Publishing Company, Inc.

Jajodia, S. and P. Ng, 1983, On representation of relational structures by entity-relationship diagrams, Entity-Relationship Approach to Software Engineering, P. Ng and R. Yeh (eds.), Amsterdam, Netherlands: Elsevier Science Publishers B.V., pp. 249-263.

Liskov, B. and S. Zilles, 1977, An introduction to formal specifications of data abstractions, Current Trends in Programming Methodology - Vol. 1: Software Specification and Design, R.T. Yeh (ed), Prentice Hall, pp 1-32.

Loucopoulos, P. and R. Zicari, 1992, Conceptual Modeling, Databases, and CASE: An integrated view of information systems development, New York: John Wiley & Sons, Inc.

Teorey, T.J. and J.P. Fry, 1982, Design of Database Structures, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Ullman, J.D., 1988, Principles of Database and Knowledge-base Systems, 2 vols. (Rockville, Maryland: Computer Science Press, Inc.)

Appendix A

Developing Standards for Spatial Data and Metadata

Spatial data standards are needed in order to facilitate the exchange of spatial data between geographic information systems. We refer to data as "spatial" because the common factor is a geographic reference (a reference in space) which allows the data to be accessed through a GIS. In order to accomplish the goal of facilitating data exchange, spatial data standards should provide:

- Definitions of terms for spatial objects or features included in GIS;
- A structure (or format) for the exchange of spatial data;
- A method for describing the accuracy and lineage of the data; and
- The definition of metadata (the data that describes the spatial data).

The primary purpose for spatial data standards is to facilitate data sharing and exchange, thus the focus only on data issues. The Council concluded that It is not necessary to develop standards for GIS hardware or software at this time. as these standards are expected to evolve from groups such as the Open GIS Consortium, a non-profit trade association formed to implement the Open Geodata Interoperability Specification .

The Current Status of Standards

At present, spatial data standards exist only at the Federal government level. Under the Federal Geographic Data Committee, three standards documents have been prepared:

The Spatial Data Transfer Standard (SDTS - FIPS 173)

This standard defines a method for the exchange of spatial data between different GIS software systems. It also contains definitions of terms for the spatial objects of interest to Federal government agencies.

Content Standards for Digital Geospatial Metadata (proposed)

This standard defines the content for digital geospatial metadata, the information about spatial data that would be entered into a clearinghouse or repository to form a catalog of spatial data available to other users.

Cadastral Standards for the National Spatial Data Infrastructure (draft)

This is a draft standard for cadastral (land ownership) data, one of twelve theme standards documents under preparation.

The Federal Geographic Data Committee has also established a National Spatial Data Infrastructure (NSDI) for the purpose of coordinating geographic data acquisition and access. The mechanism for this will be a National Spatial Data Clearinghouse, a distributed network of geospatial data producers, managers, and users linked electronically. It is envisioned that this network of clearinghouses would contain information about available spatial data. Potential users would search this clearinghouse to find data of interest, access the metadata for a description of data of interest, and could acquire the data from the distributing agency. Spatial data may be deposited directly with a clearinghouse or retained by the originator.

The Federal effort towards standards development started in 1981 and The National Spatial Data Infrastructure and Federal spatial data standards are still evolving at this time. The remaining subject area (theme) standards reports are scheduled for release during the Spring of 1996 (themes are: base cartographic, bathymetric, cultural and demographic, geodetic, geologic, ground transportation, international boundaries, soils, vegetation, water, and wetlands). The table below shows the current status of federal spatial data standards development.
Implementation of the Federal geospatial data standards is through Executive Order 12906 signed by the President on April 11, 1994. The FGDC is directed to " ...seek to involve State, local, and tribal governments in the development and implementation of the initiatives continued in this order." The Order provides that:

"Federal agencies collecting or producing geospatial data, either directly or indirectly ~e.g. through grants, partnerships, or contracts with other entities) shall ensure, prior to obligating funds for such activities, that data will be collected in a manner that meets all relevant standards adopted through the FGDC process. "

Status of Federal Geographic Data Committee Standards

Currently in development:	Completed public review:
National Spatial Data Accuracy Standard	Cadastral Content Standard
Standards for Digital Orthoimagery	Federal Domain of Values for Data Content Standard
Draft Standards for Digital Elevation Data	Cadastral Collection Standard (Cadastral)
Hydrographic and Bathymetric Accuracy Standard	Clearinghouse Metadata Profile (Cadastral)
Standards for Geodetic Control Networks	Classification of Wetlands and Deepwater Habitats of the United States
Transportation Network Profile for Spatial Data Transfer Standard	
Transportation-related Spatial Feature Dictionary	
Soils Data Transfer Standard	
Vegetation Classification Standards	
River Reach Standards and Spatial Feature Dictionary	
Facility ID Code	
Content Standard for Cultural and Demographic Data	
Metadata	

Source: Federal Geographic Data Committee **Newsletter,** November 1995.