Database Design

12

presented by:
Tim Haithcoat
University of Missouri
Columbia

GIS Database Design

DESIGN (Dictionary Definition)

A scheme in which means to an end are laid down.

DATABASE DESIGN

A scheme in which database goals are defined and database alternatives are laid out, analyzed, and an alternative is selected or recommended. Implies:

- That goals are identified and set.
- A structured decision making process:
  - identify goals of database
  - identify issues or problems in the way of goal achievement
  - identify alternatives to resolve issues
  - select or recommend an alternative
  - test alternative
  - act or implement

Special Aspects of GIS Database Design

- Data model is hybrid
- Location is a special kind of key
- Issues connected with coordinates
- Issues connected with cartography
- Standard database literature and theory do not take account of these special characteristics

Spatial element normally not dealt within database

- ✗ 1773 Eucalyptus, Perris, CA
- ✗ 117 15 00 33 30 10
- ✗ UTM Zone 11S; N 6487529 m; E 455322 m

- All refer to the same location on the earth. In a regular database, there is no way to relate these 3 ways of describing location.
- GIS forces you to deal with this issue so that location truly becomes a special kind of key.
- The explicitly definition of location allows diverse data sets to be related that ordinarily we could not related.
- Georelational Model

Why Database Design?

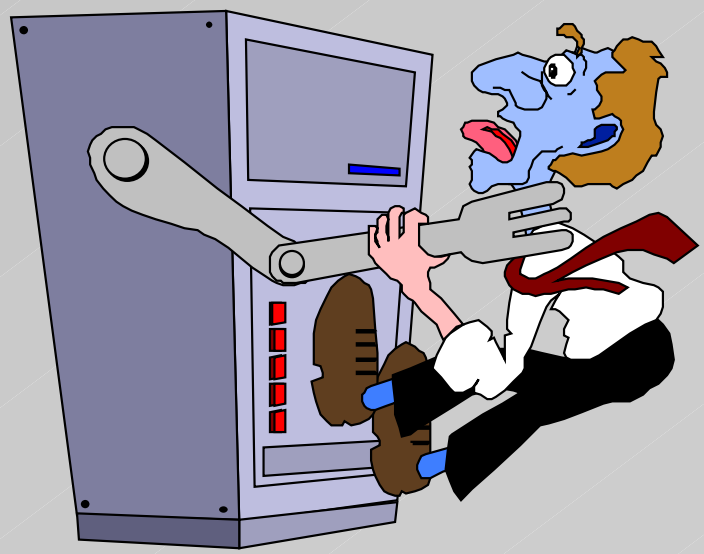
- The design provides the comprehensive mental framework & organization of the database.
  - It allows the database to be viewed in its entirety and evaluate how the various aspects of the database need to interact.
  - It allows for the early identification of major issues, potential problems, and design alternatives.
- To determine answers that define GIS efficiency constraints.
  - about users
  - about data
  - about input
  - security
  - management
  - hardware

GIS components

What happens if you do NOT conduct a database design?

Not conducting a comprehensive design can lead to a poorly constructed and functionally unsatisfactory database. The lack of a design can result in:

- unsupported applications
- an uncorrectable database
- lack of coherency within the database



Build the Database

This is the most critical and time consuming part of the project.

The completeness and accuracy of the database will determine the quality of the analysis and final products.

STEPS:

- Designing the database
- Automating the data
- Getting the spatial data in
- Making the spatial data usable
- Getting the attribute data in
- Managing the database

Stages in Database Design

- **CONCEPTUAL**
  - Software & hardware independent
  - Describes & defines included entities
  - Identifies how entities will be represented in the database
  - Requires decisions about how real-world dimensionality & relationships will be represented
- **LOGICAL**
  - Software specific but hardware independent
  - Sets out the logical structure of the database elements, determined by the database management system used by the software
- **PHYSICAL**
  - Both hardware & software specific
  - Requires consideration of how files will be structured for access

Determining Objectives

Questions to Consider:

- What is the problem to be solved?
- How is it solved now?
- Are there alternative methods for solving the problem?
- What are the final products of the project - reports, working maps, presentation quality maps?
- How frequently will products be generated?
- Who is the intended audience for the final products - technicians, planners, managers, the general public?
- Are there, or will there be, other uses for this data?
- If there are other uses, what will be the specific requirements?

Assess User Needs

To design a GIS database, you must know:

Users: Who will use the database?

Data Elements: What kinds of data will they need to have in the database?

Purpose: Why do they need the data and how will they use it?

Sources: Where are the data elements they need?

Quality: What is the resolution, accuracy, and reliability of the data available?

Precision: Does the input data warrant the storage of double precision coordinates?

Building the Database (Database Design)

Determine study area bounds

Determine what coordinate system

Identify the required spatial data layers

Determine the required feature attributes

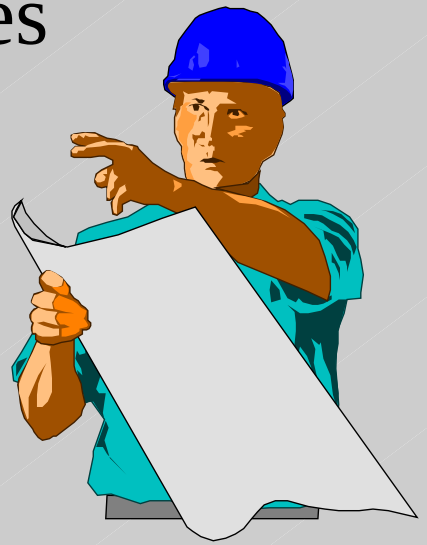
Define each attribute and its codes

Coordinate registration

Code schemes

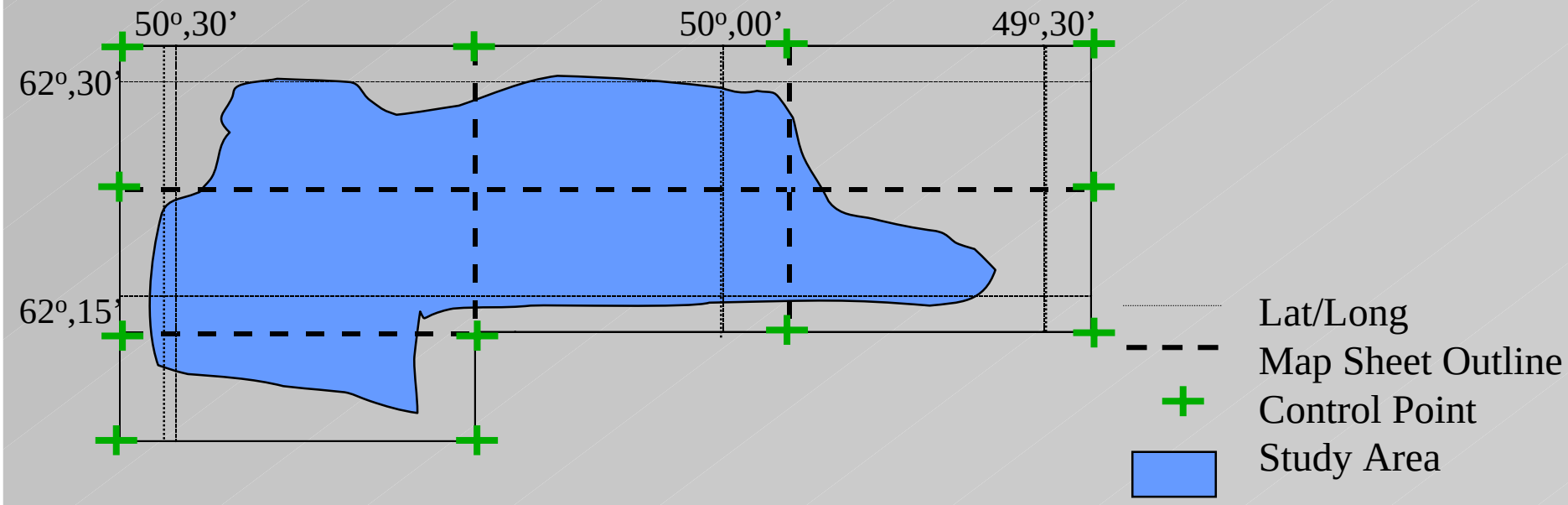
Allocate storage

Project organization



Define Study Area

- To define the GIS database study area, you must determine:
  - The geographic extent of the data elements required to meet user needs (boundaries)
  - Several points within the extent of the database for which the coordinate value are known in a common Cartesian coordinate system (Geographic Reference/Control Points)



Organize & Describe Data Elements

- Data elements should be organized into “layers” according to the logical relationships among them.

Terrain Layer	Infrastructure Layer
Soil	Roads
Slope	Railroads
Geology	Transmission Line
Aspect	Pipelines
Vegetation	

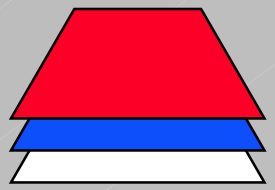
- Data elements should be described according to a well-defined classification and coding schema.

Vegetation	Description	Code
Forest	Pine	(P) or (100)
	Spruce	(SP) or (101)
	Cedar	(c) or (200)
Wetland	Emergent Marsh	(EM) or (402)
	Spring Bog	(SB) or (403)

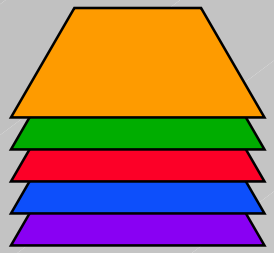
Data Organization

DATA ELEMENTS

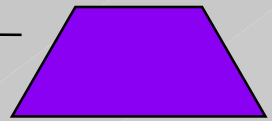
- Parcels
  - Easements
  - Zoning
  - Blocks



- Special Uses
  - Wildlife Habitat
  - Archaeological Sites
  - Flood Zones
  - Drainage Districts
  - Sewer & Water Facilities

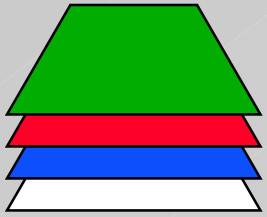


- Administrative
  - Census Tracts
  - Census Blocks
  - Fire Response Zones
  - Council Districts
  - Precincts
  - Zip Codes



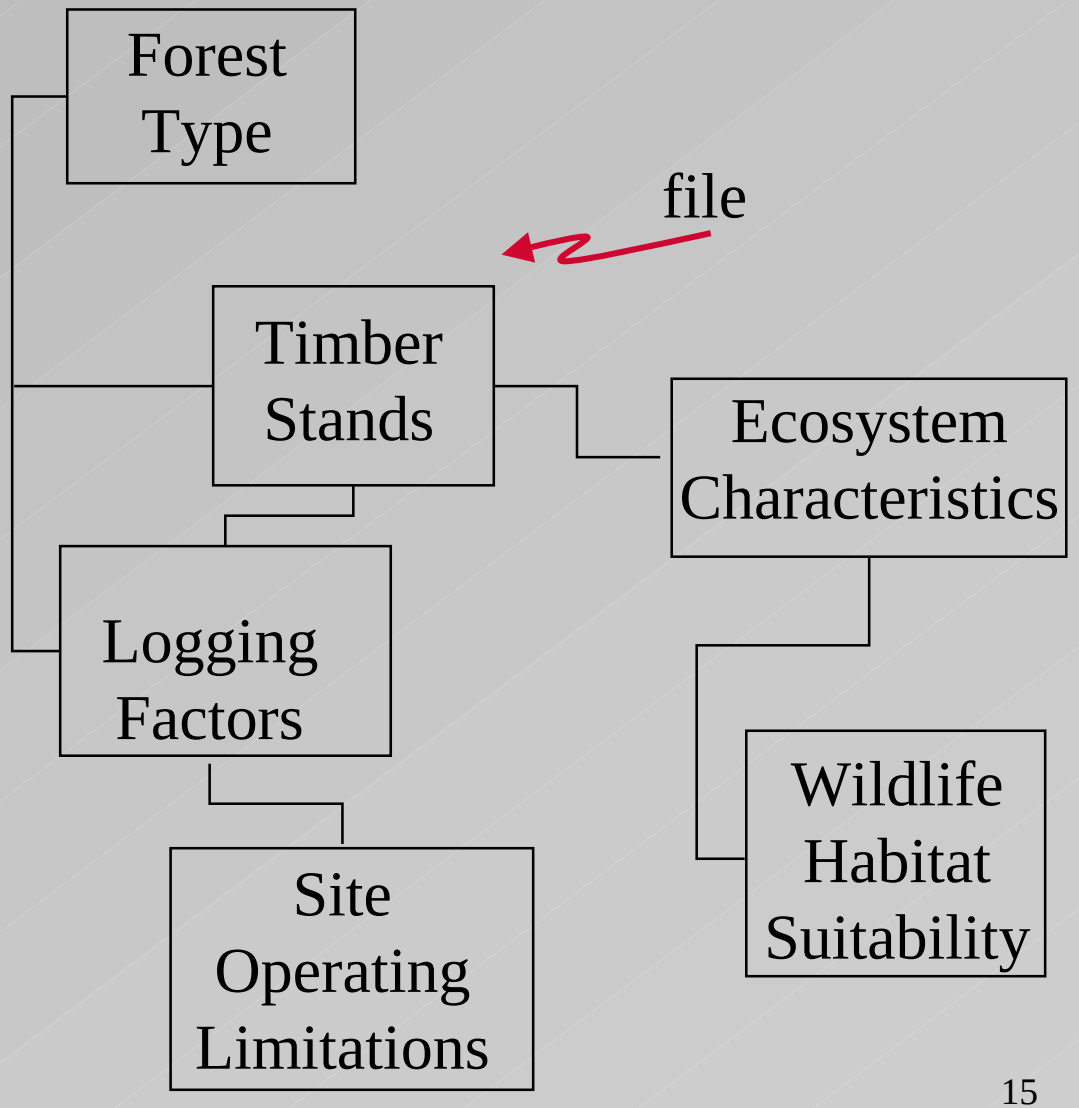
- Natural Resources
  - Geology
  - Soils
  - Vegetation

  - Hydrology
  - Slope
  - Ecosystems
  - Watersheds



Define Database Structure

- The Physical Structure of the Database
  - Data Elements Assigned to Files
  - Relationships Among Files

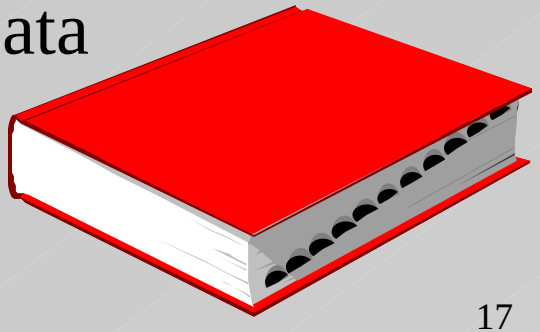


Item Types

EXAMPLE	TYPE	ABBREVIATION
Main Street	Character	C
10/15/90 10/15/1990	Date	D
23675	Integer	I
347.22	Numeric	N

Data Dictionary

- **Definition**
  - Detailed description of data within a database (documentation)
  - "Users Guide" to the Database
- **Functions**
  - Focal Point for Data Standardization
  - Facilitates Development of Application Programs
  - Assists with Database Design Process
  - Improves Understanding of the Data
  - Reduces Data Redundancy
  - Improves Data Integrity



Common Design Problems

You have one table with a large number of fields that don’t all relate to the same subject.

You have fields that are intentionally left blank in many records because they aren’t applicable to those records.

You have a large number of tables which contain the same fields.

Refining the Design

Questions to Ask:

- ? Did you forget any fields? Is there information you need that isn't included?
- ? Did you choose a good primary key for each table? Is it easy to type and remember?
- ? Are you repeatedly entering duplicate information in one of your tables?
- ? Do you have tables with many fields, a limited number of records, and many empty fields in individual records?

Issues in Database Design

- What storage media should be used?
  - How big is the database?
  - How will database grow over time?
  - What are the required access speeds?
- Should data be partitioned by location or theme?
- Distributed or centralized? If distributed, on what host does what data go?
- Who is responsible for making sure the data standards are adhered to?
- Who is responsible for enforcing the data standards?