M.Tech Geoinformatics

Geographic Information System

(24CC03)

Unit II: Data Management

Prakash. K

Guest Faculty

Department of Geography

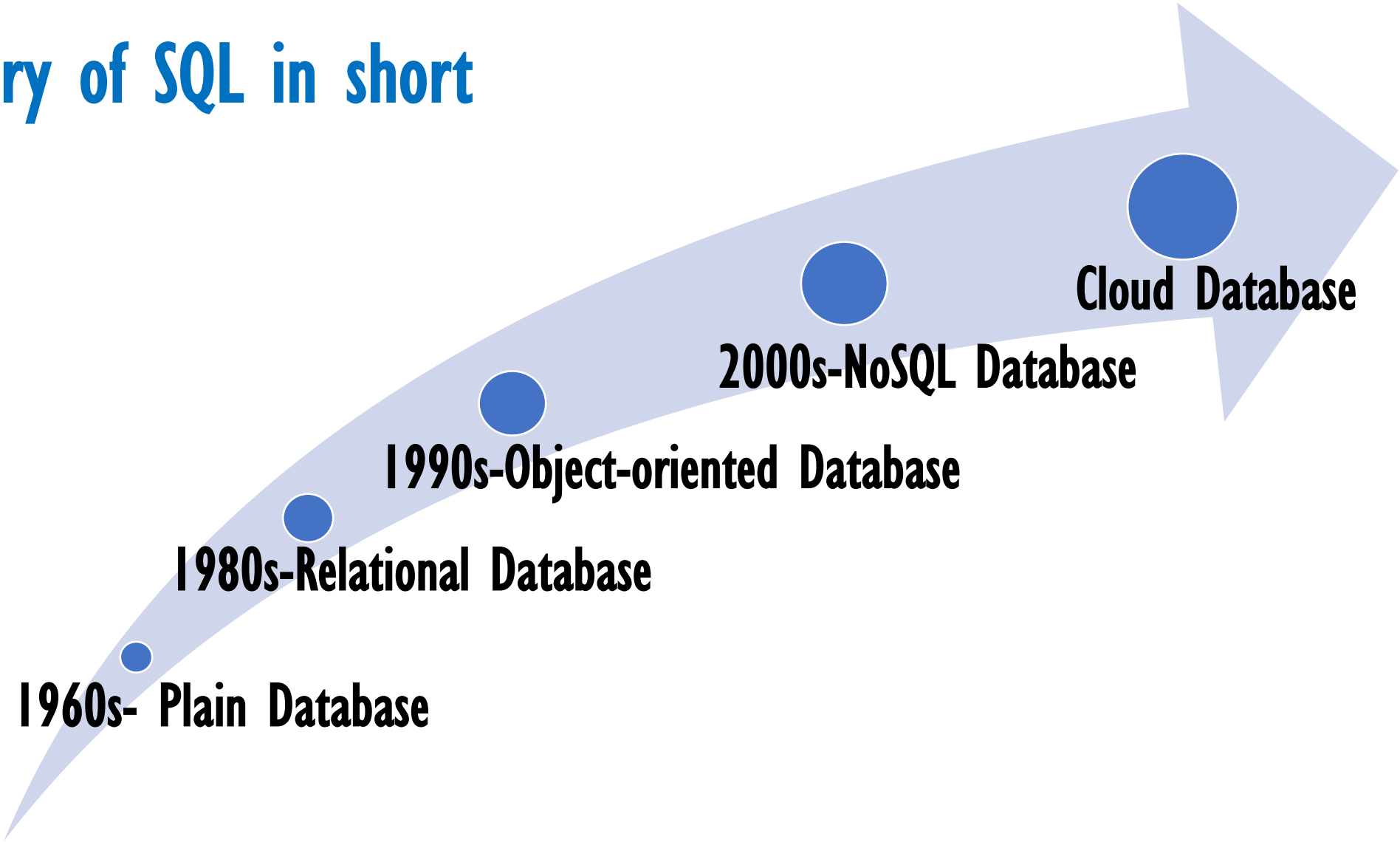
Bharathidasan University, Tiruchirappalli.

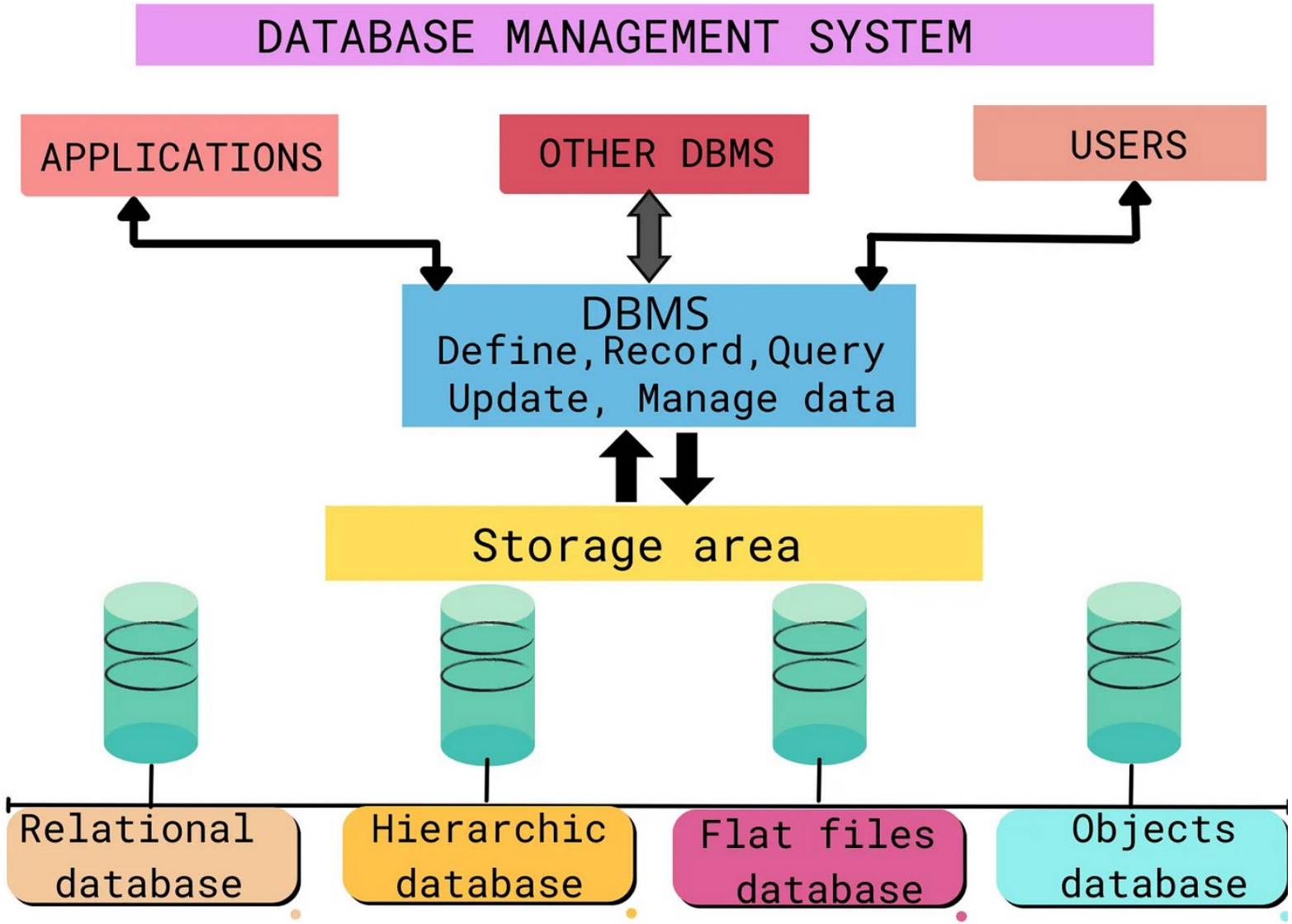
Spatial Database

Database

- A database is an organized collection of structured information, or data, typically stored electronically in a computer system.
- A database is usually controlled by a Database Management System (DBMS).
- Data within the most common types of databases in operation today is typically modeled in rows and columns in a series of tables to make processing and data querying efficient.
- The data can then be easily accessed, managed, modified, updated, and organized.
- Most databases use Structured Query Language (SQL) to perform this.

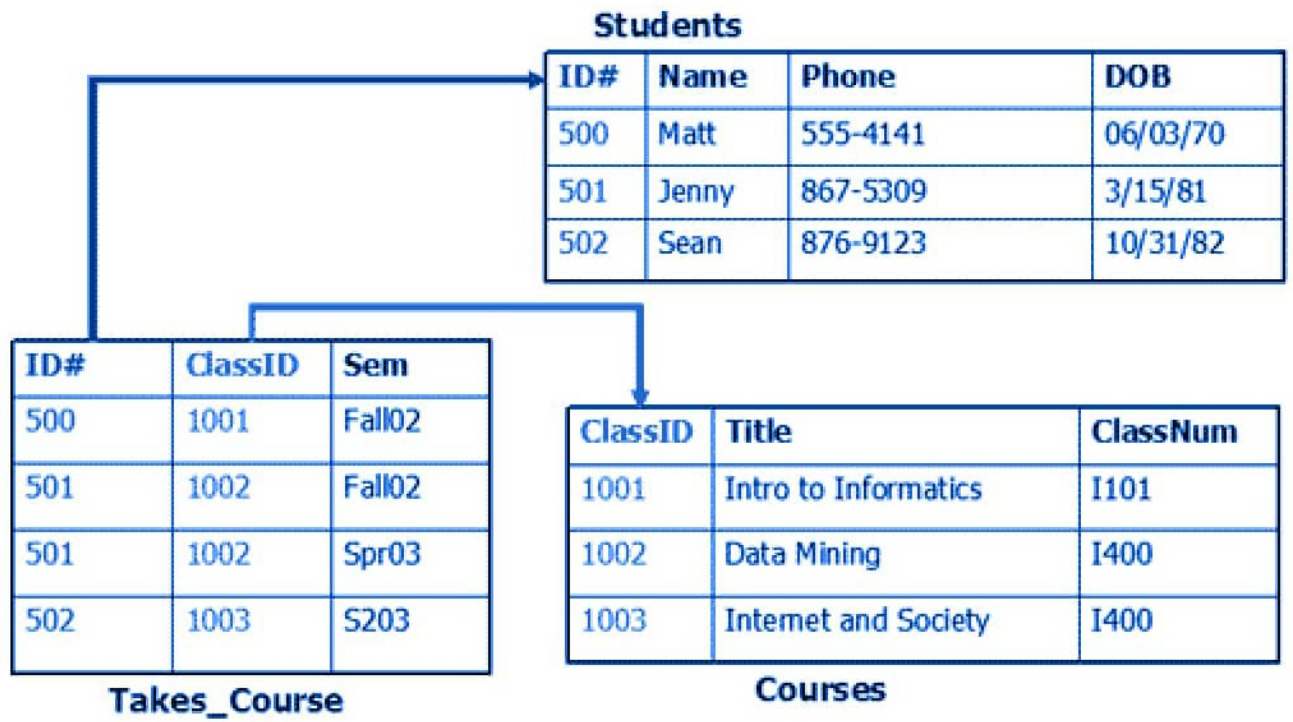
- History of SQL in short




DATABASE MANAGEMENT SYSTEM
APPLICATIONS
OTHER DBMS
USERS
DBMS
Define,Record,Query
Update, Manage data
Storage area
Relational database
Hierarchic database
Flat files database
Objects database

- SQL is a programming language used by nearly all relational databases to
    - define, query, manipulate, and to provide access control to data
- SQL was first developed at IBM in the 1970s with Oracle as a major contributor
- SQL has spurred many extensions from companies such as IBM, Oracle, and Microsoft.

Relational Database Management System



Students

ID#	Name	Phone	DOB
500	Matt	555-4141	06/03/70
501	Jenny	867-5309	3/15/81
502	Sean	876-9123	10/31/82

Takes_Course

ID#	ClassID	Sem
500	1001	Fall02
501	1002	Fall02
501	1002	Spr03
502	1003	S203

Courses

ClassID	Title	ClassNum
1001	Intro to Informatics	I101
1002	Data Mining	I400
1003	Internet and Society	I400

Database vs File System

- Sharing of data
- Data Abstraction
- Security and Protection
- Recovery Mechanism
- Manipulation Techniques
- Concurrency Problems
- Data Redundancy and Inconsistency
- Integrity Constraints

Database System:

Oracle, SQL Server, Sybase etc

File System:

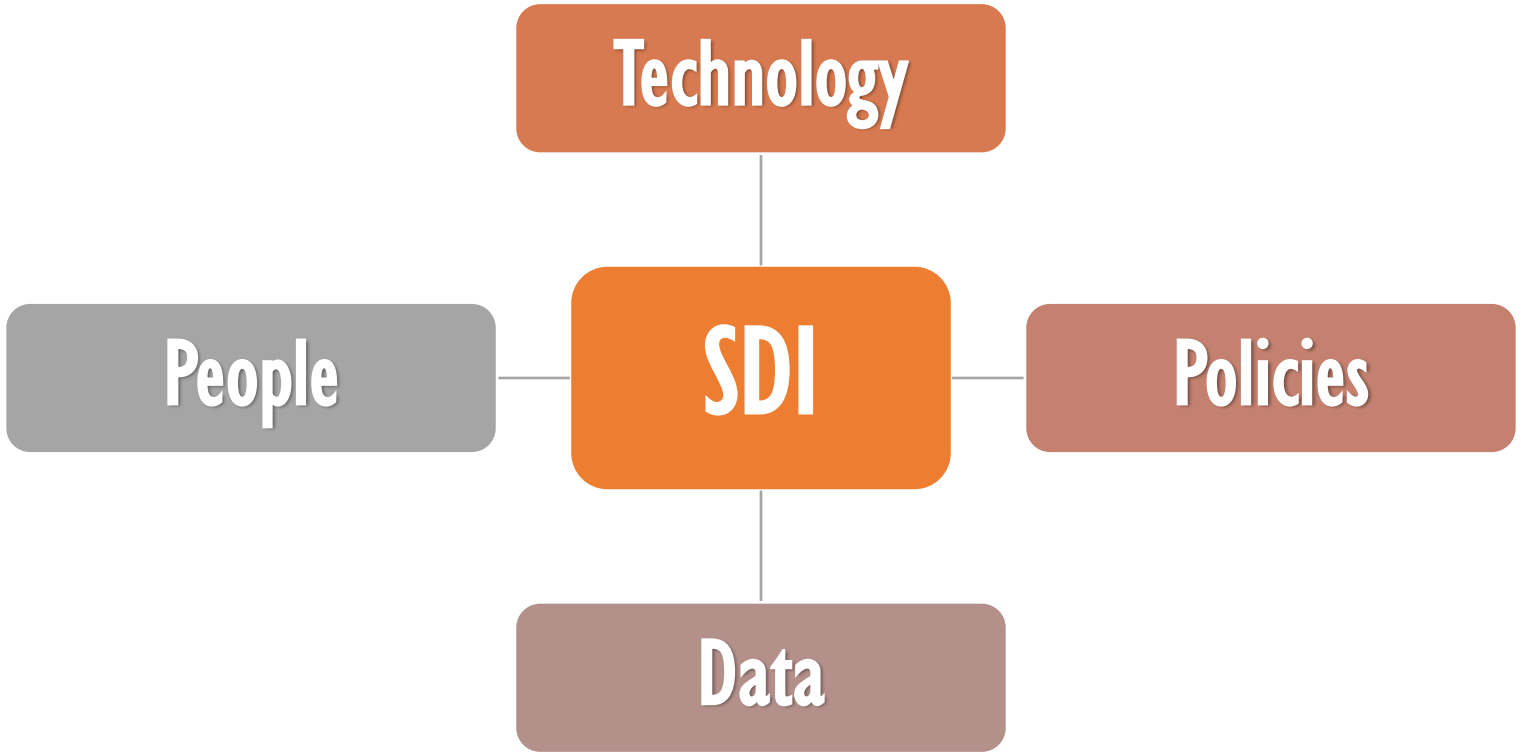
FAT (File Allocation Table)
NTFS (New Technology File System)
APFS (Apple File System)

Spatial Data Infrastructure:
Concepts and Components

What is a Spatial Data Infrastructure (SDI)?

- "The SDI is a basis for spatial data **discovery**, **evaluation**, and **application** for users and providers within all levels of government, commercial sector, non-profit sector, academia and by citizens in general" – GSDI Cookbook v2

Components of SDI



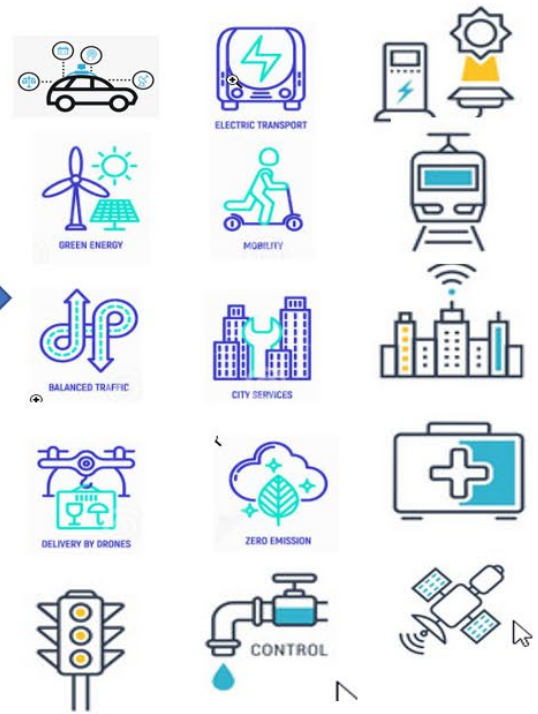
Strategy / Policies | Data Sources and Automation | Applying Technology Intelligence | Business value / Solutions

Policies & Governance

Standards

Collaboration

Capacity Planning



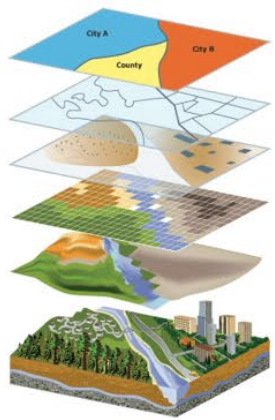
Data Catalog

Master Data

Metadata

Geospatial Datasets

Standards



Generative AI Models

Scalable Cloud Infrastructure

Prediction / Forecasting

Corrective Actions

Location Intelligence

Spatial Data Infrastructure Framework

Consortiums

- Besides these software components, a range of (international) technical standards are necessary that allow interaction between the different software components.
- Among those are geospatial standards defined by the **Open Geospatial Consortium** (e.g., OGC WMS, WFS, GML, etc.) and **International Organization for Standardization (ISO)** for the delivery of maps, vector and raster data, but also data format and internet transfer standards by W3C consortium.

Why SDI?

- Build data once and use it many times
  - for many applications
  - for many users
- Integrate distributed providers of data
- Share costs of data creation and maintenance
- Support sustainable economic, social, and environmental development

If SDI were developed…

- Improved decision
  - Providing decision makers what they really need: indicators, models, trends, patterns
- Business opportunities
  - Development of a private sector involved with data sales and added value
- Increased globalization
  - A chance for developing countries to participate in the knowledge economy

Data Capturing Methods

- Data are vital
- Data sources:

Primary

- Positioning, GPR Surveys
- Satellite Images
- Classical Surveys
  - Polls
  - Questionnaire

Secondary

- Satellite Images
- Maps
- Governments Reports
  - Census
  - NFHS

Tertiary

- References / Bibliographs
- Encyclopaedia
- Dictionaries

Methods of data capturing

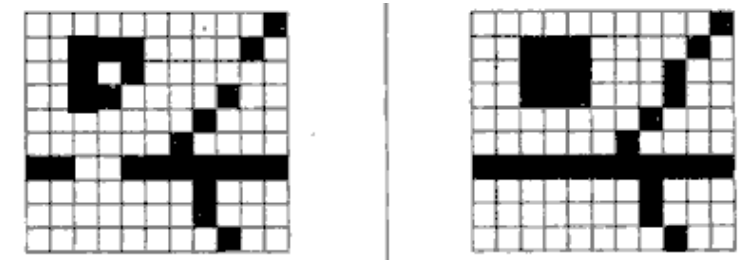
- Collect / Generate: in-person methods of data collection
- Off-the-Shelf (OTS): use existing data
- Digitize: convert into editable vector format
- Import: convert into specific file type
  - Spatial data
  - Attribute data

Spatial Data Editing

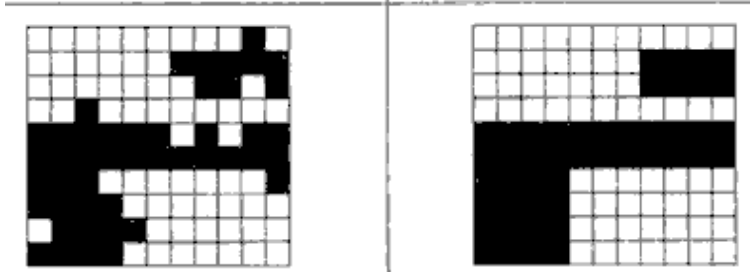
- What is the nature of spatial data?
- Temporal data often needs to be updated / edited
- Create/Edit spatial data may often produce errors
  - Raster Data Editing
  - Spatial Data Editing

Raster Data Editing

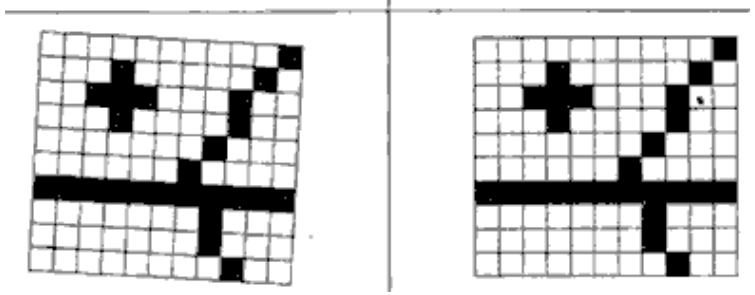
Hole



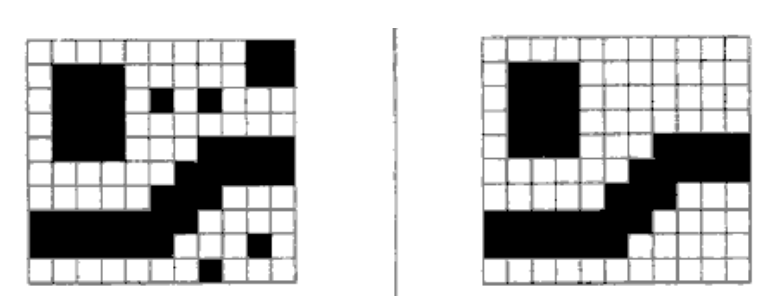
Smoothening



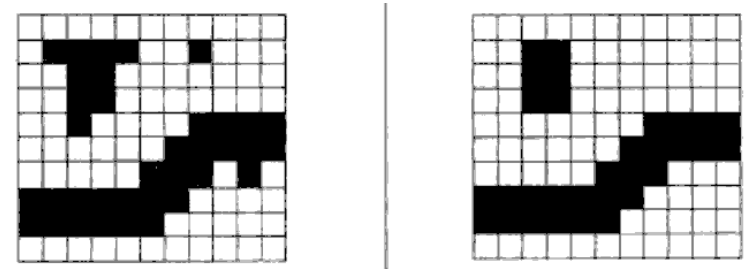
De-skewing



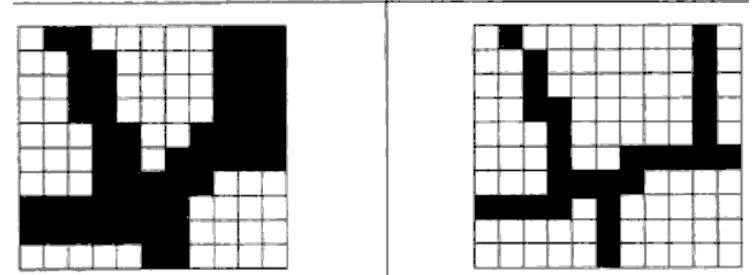
Speckle



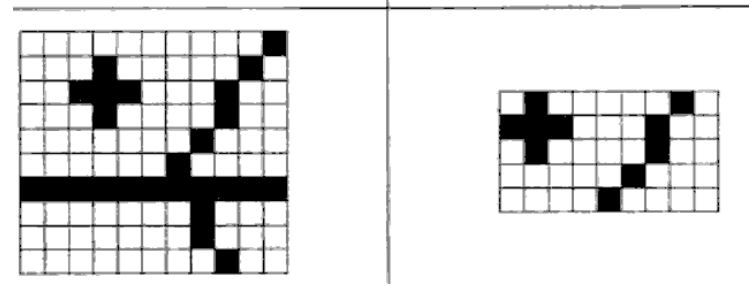
Deletion



Thinning

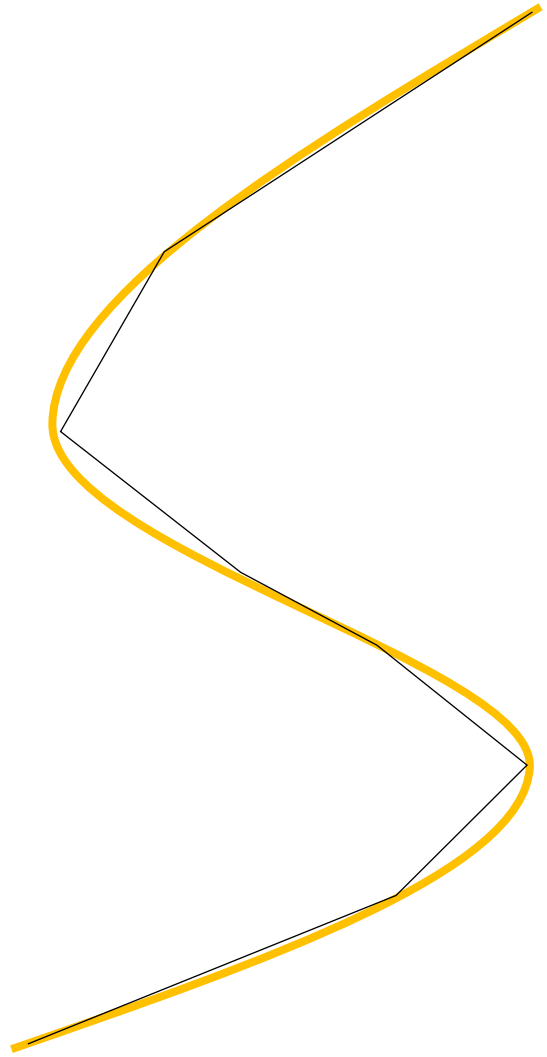


Clipping



Vector Data Editing

- Modification
- Generalization
  - Smoothening — Arc, Bezier curves
  - Densification / De-densification
- Edge Matching
- Merge / Split



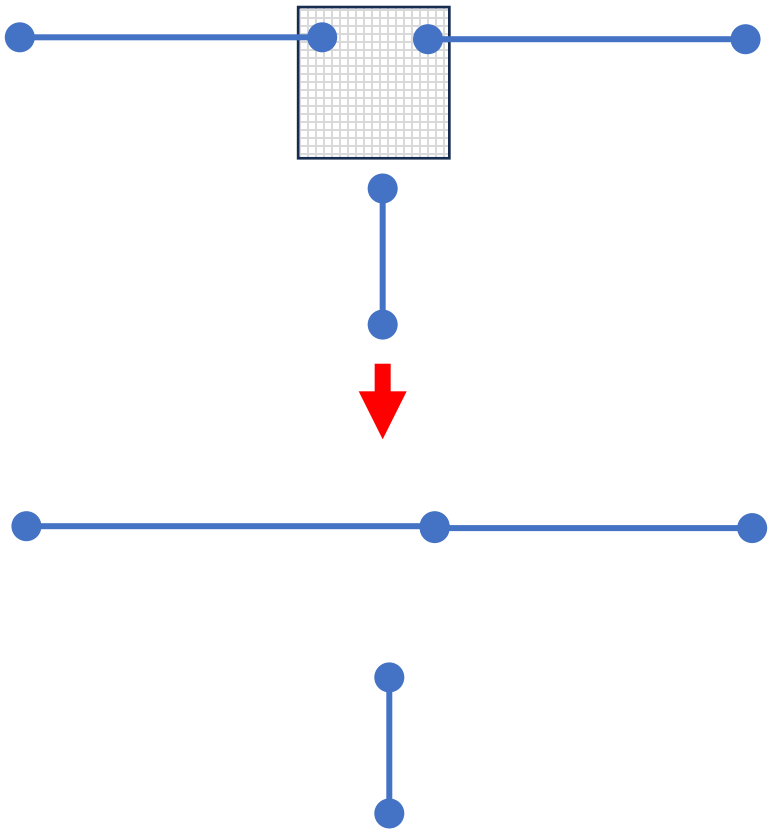
Vector Data Editing

- Geometric / Spatial Tolerance (x, y, z)

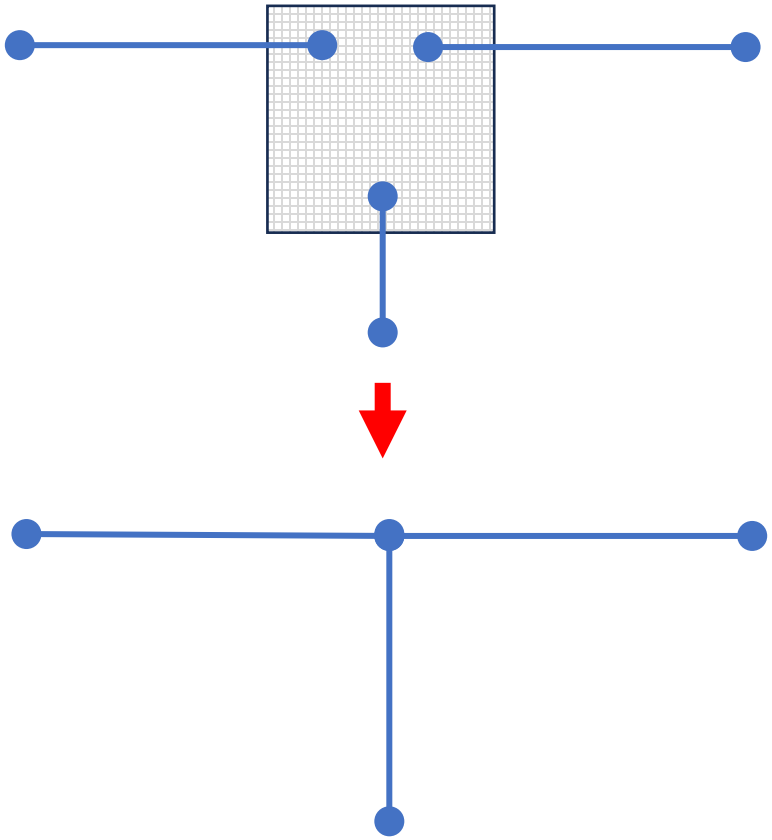
Cluster tolerances is used to integrate vertices:

- An x, y tolerance to find vertices within the horizontal distance of one another.
- A z-tolerance to distinguish whether or not the z-heights or elevations of vertices are within the tolerance of one another and should be clustered.

Tolerance: 0.2

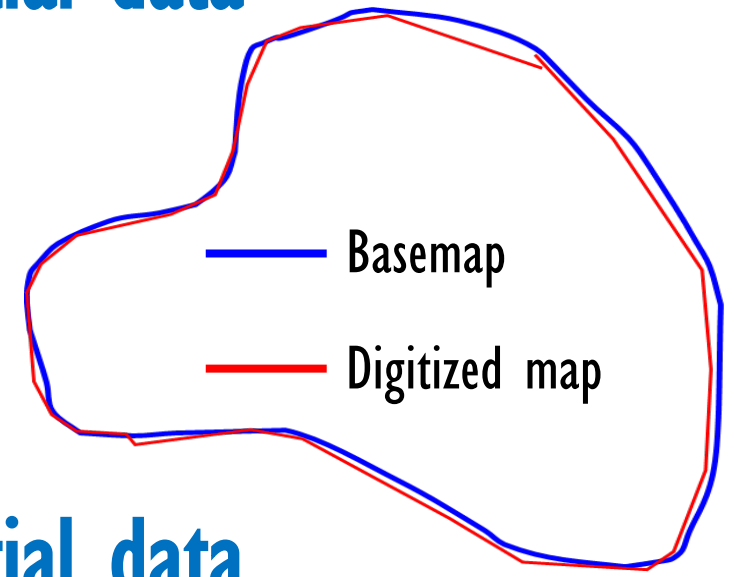


Tolerance: 0.3



Types of Errors and Types of Accuracies

- Spatial / Geometrical / Positional Error / Accuracy
  - Occurs at the time of capturing / digitizing spatial data

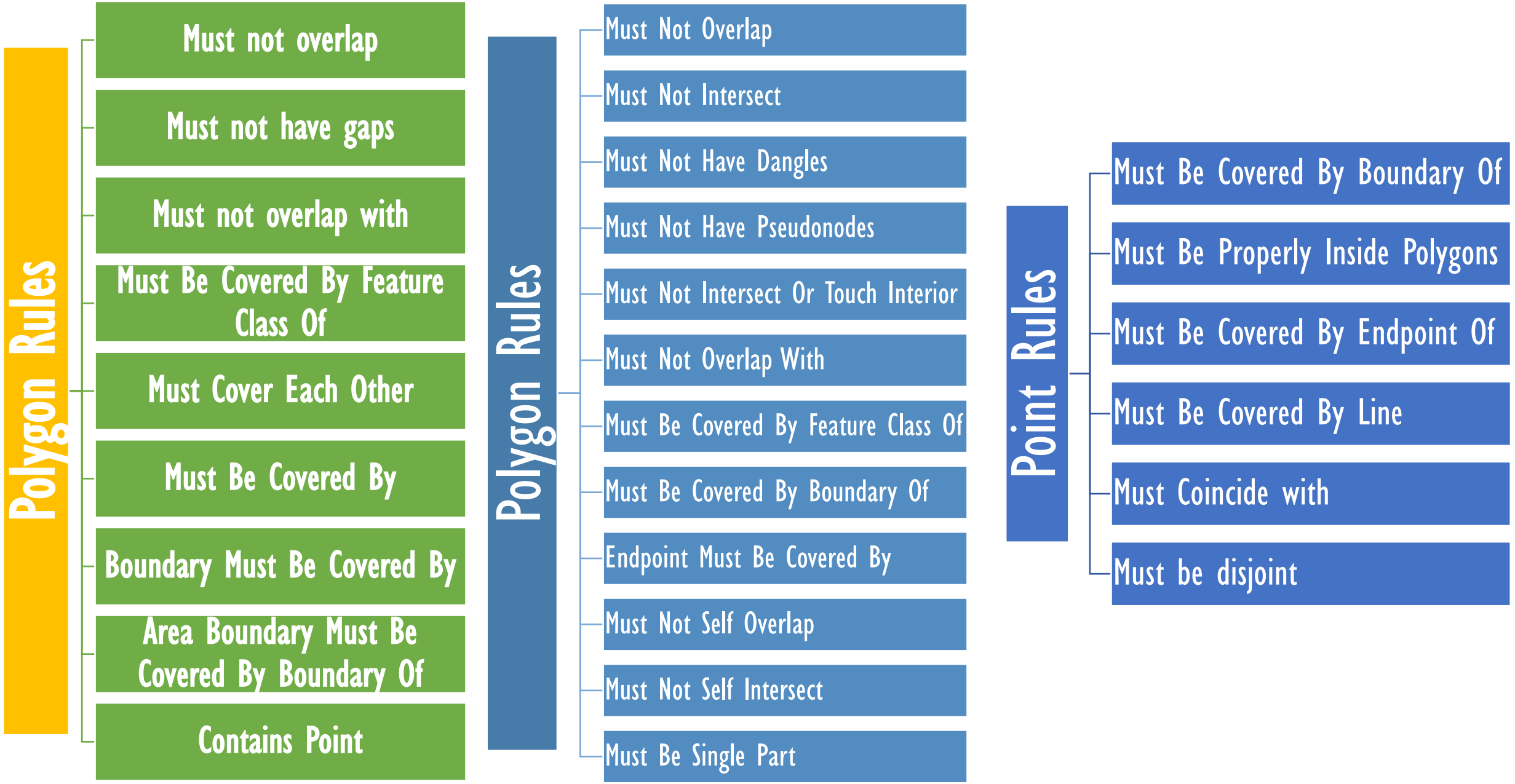


- Topological Error / Accuracy
  - Occurs at the time of capturing / digitizing spatial data
  - Occurs at the time of Editing

• • •

- **Few topological errors**

Point	Overlap
	Misplacement
	…
Line	Overshoot
	Undershoot
	Overlap
	Intersect
	Misplacement
	…
Polygon	Overlap
	Sliver
	Gap
	…


Polygon Rules
Must not overlap
Must not have gaps
Must not overlap with
Must Be Covered By Feature Class Of
Must Cover Each Other
Must Be Covered By
Boundary Must Be Covered By
Area Boundary Must Be Covered By Boundary Of
Contains Point
Polygon Rules
Must Not Overlap
Must Not Intersect
Must Not Have Dangles
Must Not Have Pseudonodes
Must Not Intersect Or Touch Interior
Must Not Overlap With
Must Be Covered By Feature Class Of
Must Be Covered By Boundary Of
Endpoint Must Be Covered By
Must Not Self Overlap
Must Not Self Intersect
Must Be Single Part
Point Rules
Must Be Covered By Boundary Of
Must Be Properly Inside Polygons
Must Be Covered By Endpoint Of
Must Be Covered By Line
Must Coincide with
Must be disjoint

Topology in ArcGIS System

Polygon

Must not overlap

Polygons must not overlap within a feature class or subtype. Polygons can be disconnected or touch at a point or touch along an edge.

Polygon errors are created from areas where polygons overlap.

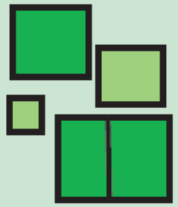
A voting district map cannot have any overlaps in its coverage.

Use this rule to make sure that no polygon overlaps another polygon in the same feature class or subtype.

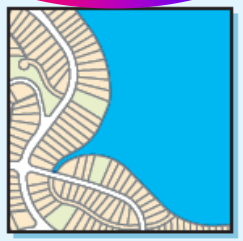
Polygon

Must not overlap with

Polygons of the first feature class or subtype must not overlap polygons of the second feature class or subtype.



Polygon errors are created where polygons from the two feature classes or subtypes overlap.



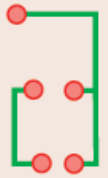
Lakes and land parcels from two different feature classes must not overlap.

Use this rule when polygons from one feature class or subtype should not overlap polygons of another feature class or subtype.

Line

Must not have dangles

The end of a line must touch any part of one other line or any part of itself within a feature class or subtype.



Point errors are created at the end of a line that does not touch at least one other line or itself.

A street network has line segments that connect. If segments end for dead-end roads or cul-de-sacs, you could choose to set as exceptions during an edit session.

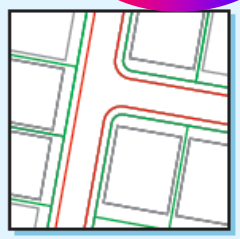
Use this rule when you want lines in a feature class or subtype to connect to one another.

Line

Must not overlap

Lines must not overlap any part of another line within a feature class or subtype. Lines can touch, intersect, and overlap themselves.

Line errors are created where lines overlap.



Lot lines cannot overlap one another.

Use this rule with lines that should never occupy the same space with other lines.

C:\Program Files (x86)\ArcGIS\Desktop10.X\Documentation

Thank You!