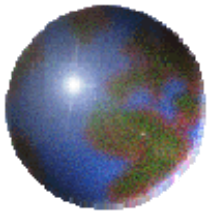


9. GIS Data Collection

Geographic Information Systems and Science

SECOND EDITION

Paul A. Longley, Michael F. Goodchild, David J. Maguire, David W. Rhind

© 2005 John Wiley and Sons, Ltd

Overview

- Introduction
- Primary data capture
- Secondary data capture
- Data transfer
- Capturing attribute data
- Managing a data capture project

© 2005 John Wiley & Sons, Ltd

Data Collection

- One of most expensive GIS activities
- Many diverse sources
- Two broad types of collection
  - Data capture (direct collection)
  - Data transfer
- Two broad capture methods
  - Primary (direct measurement)
  - Secondary (indirect derivation)

© 2005 John Wiley & Sons, Ltd

Data Collection Techniques

	Raster	Vector
Primary	Digital remote sensing images	GPS measurements
	Digital aerial photographs	Survey measurements
Secondary	Scanned maps	Topographic surveys
	DEMs from maps	Toponymy data sets from atlases

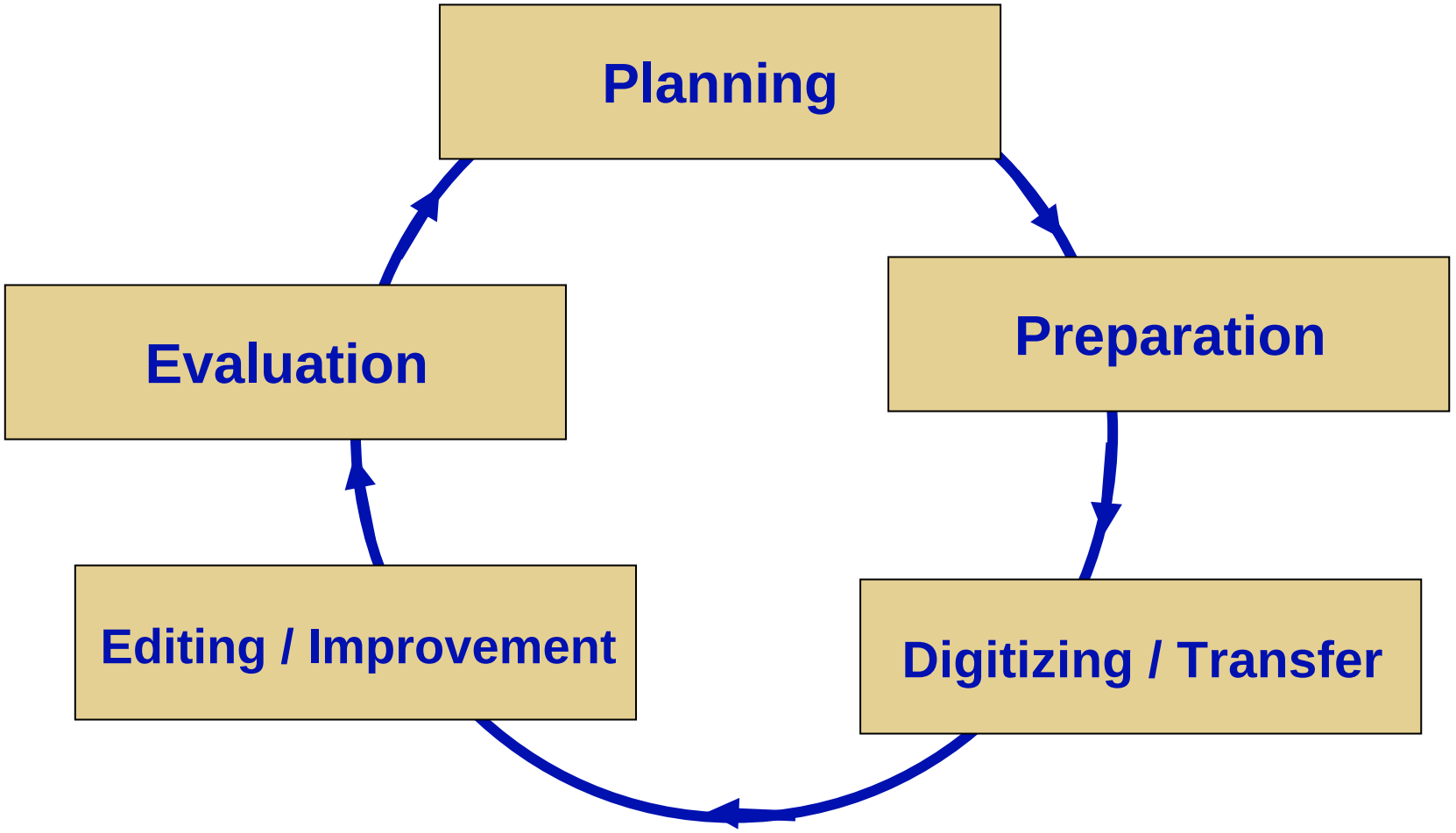
© 2005 John Wiley & Sons, Ltd

Client-server GIS Costs

	10 Seats		100 Seats	
	$	%	$	%
Hardware	30	3.4	250	8.6
Software	25	2.8	200	6.9
Data	400	44.7	450	15.5
Staff	440	49.1	2000	69.0
Total	895	100	2900	100

© 2005 John Wiley & Sons, Ltd

Stages in Data Collection Projects

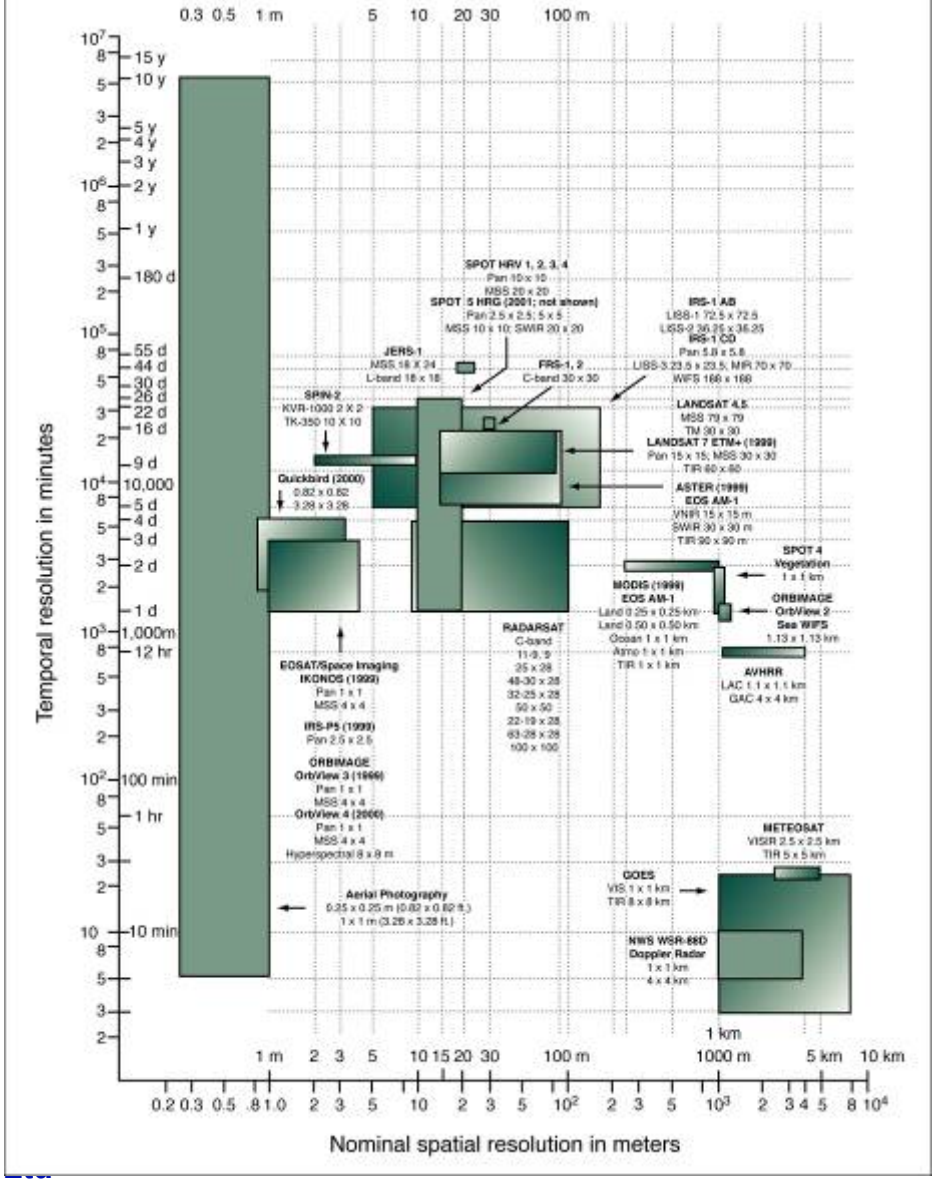


© 2005 John Wiley & Sons, Ltd

Primary Data Capture

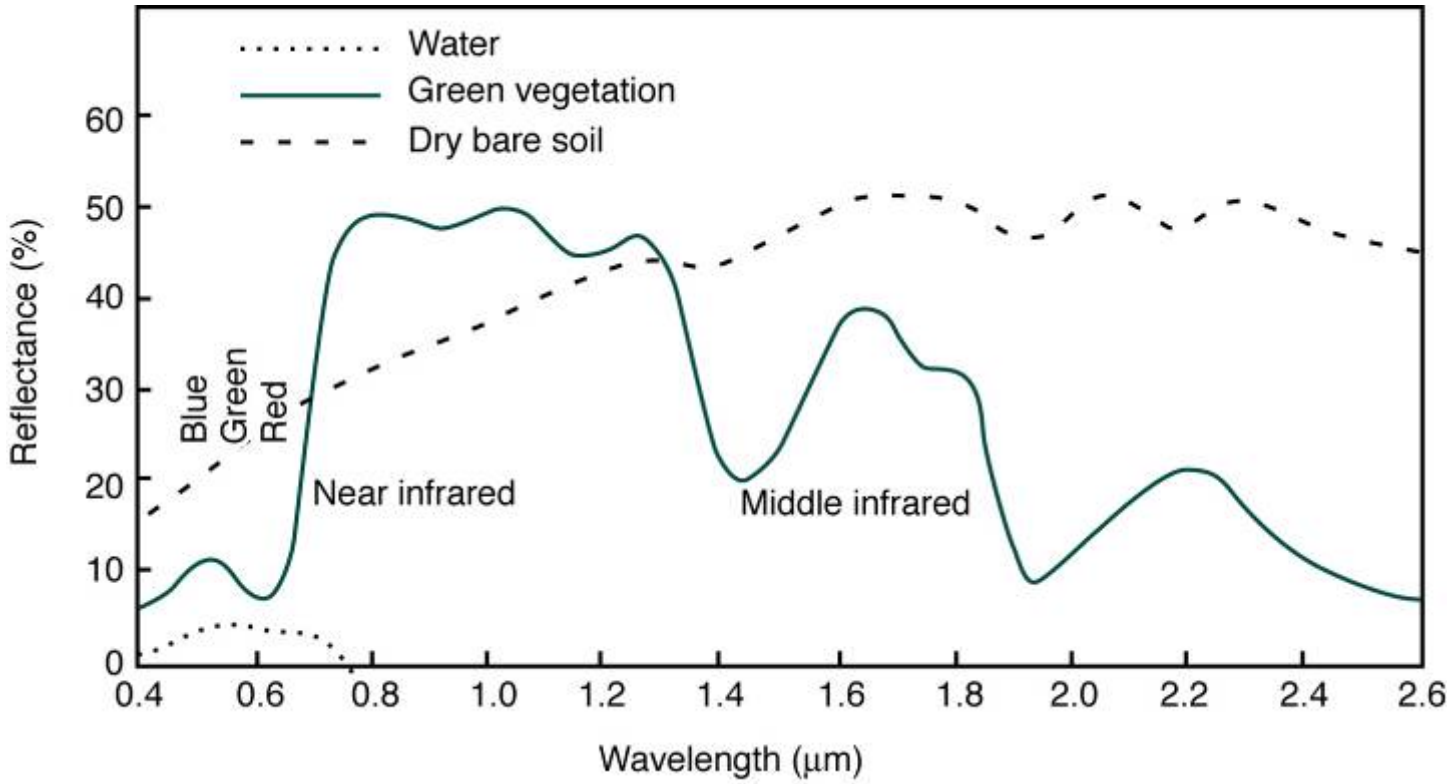
- Capture specifically for GIS use
- Raster – remote sensing
  - e.g. SPOT and IKONOS satellites and aerial photography
  - Passive and active sensors
- Resolution is key consideration
  - Spatial
  - Spectral
  - Temporal

© 2005 John Wiley & Sons, Ltd



© 2005 John Wiley & Sons, Ltd.

Typical Reflectance Signatures



© 2005 John Wiley & Sons, Ltd

Vector Primary Data Capture

- Surveying
  - Locations of objects determines by angle and distance measurements from known locations
  - Uses expensive field equipment and crews
  - Most accurate method for large scale, small areas
- GPS
  - Collection of satellites used to fix locations on Earth's surface
  - Differential GPS used to improve accuracy

© 2005 John Wiley & Sons, Ltd

Total Station

© 2005 John Wiley & Sons, Ltd

Secondary Geographic Data Capture

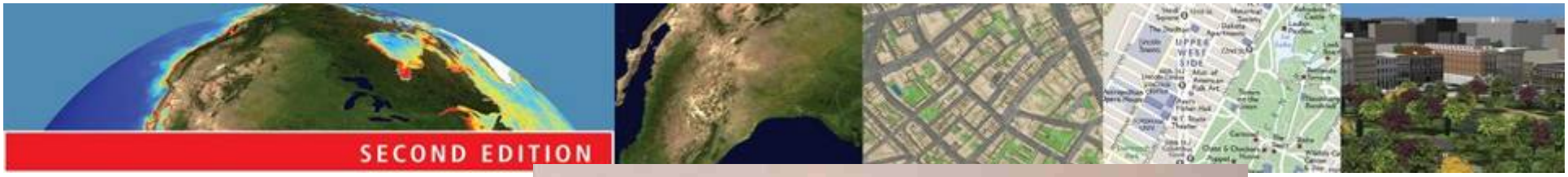
- Data collected for other purposes can be converted for use in GIS
- Raster conversion
  - Scanning of maps, aerial photographs, documents, etc
  - Important scanning parameters are spatial and spectral (bit depth) resolution

© 2005 John Wiley & Sons, Ltd

Vector Secondary Data Capture

- Collection of vector objects from maps, photographs, plans, etc.
- Digitizing
  - Manual (table)
  - Heads-up and vectorization
- Photogrammetry – the science and technology of making measurements from photographs, etc.
- COGO – Coordinate Geometry

© 2005 John Wiley & Sons, Ltd



Scanner

© 2005 John Wiley & Sons, Ltd

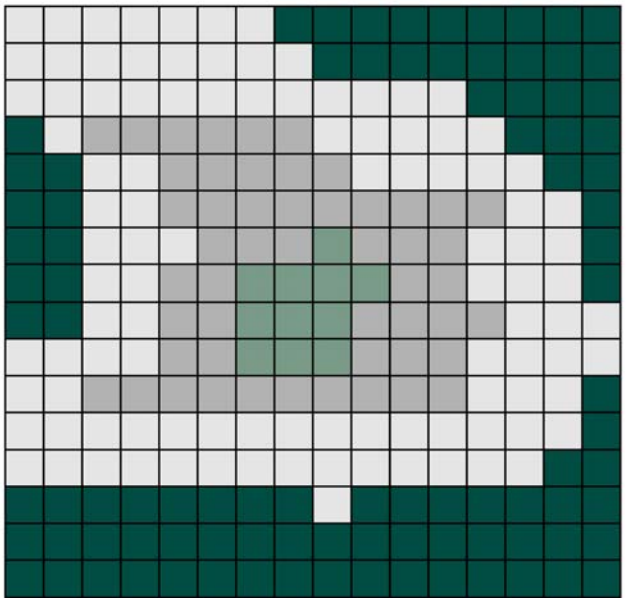
Vector Over Raster

© 2005 John Wiley & Sons, Ltd

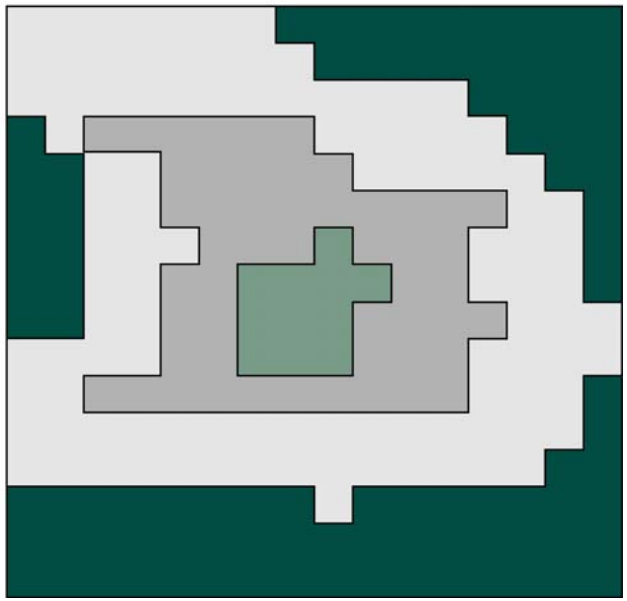
Digitizer

Batch Vectorization

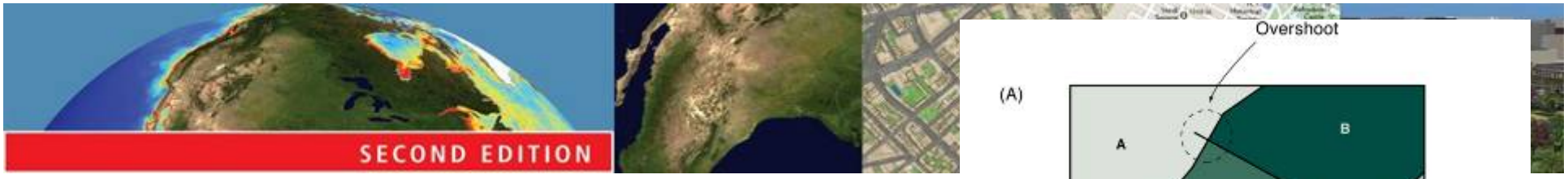
(A)



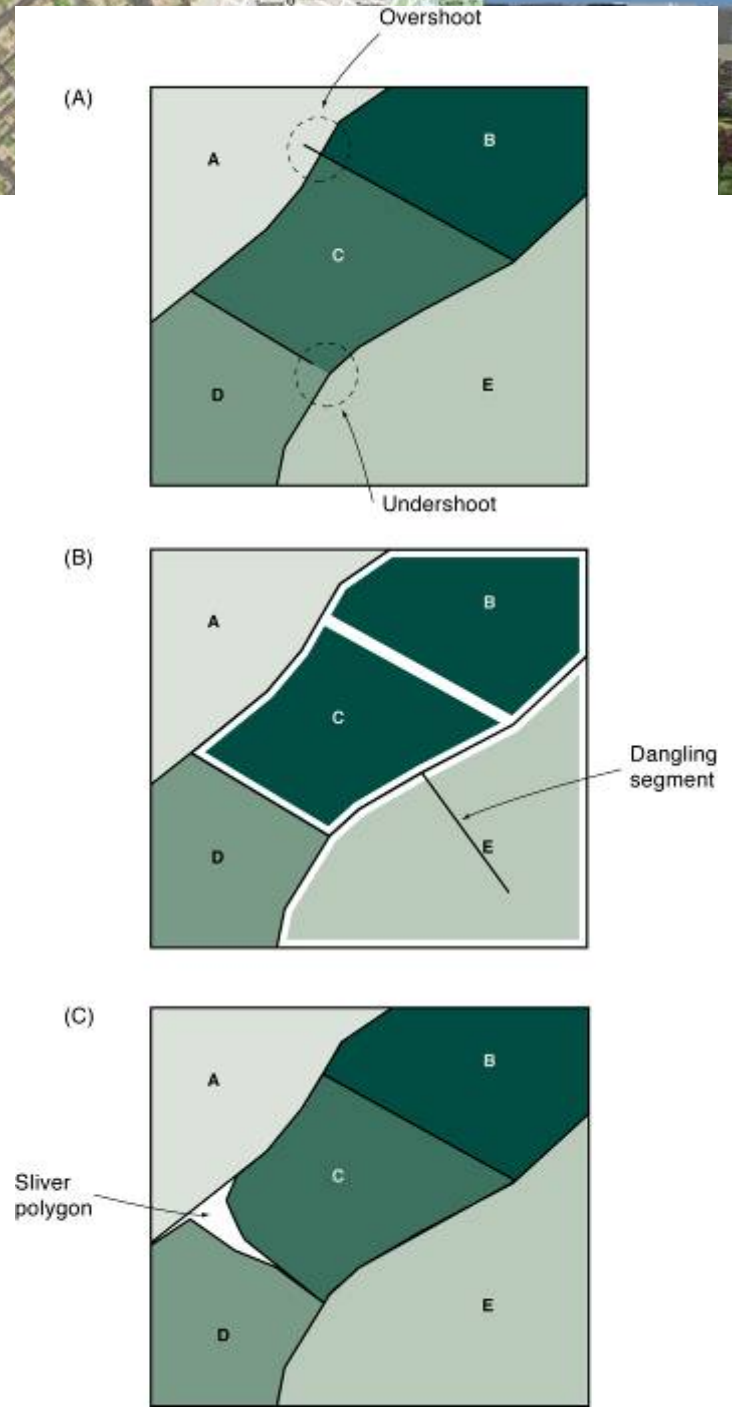
(B)



© 2005 John Wiley & Sons, Ltd

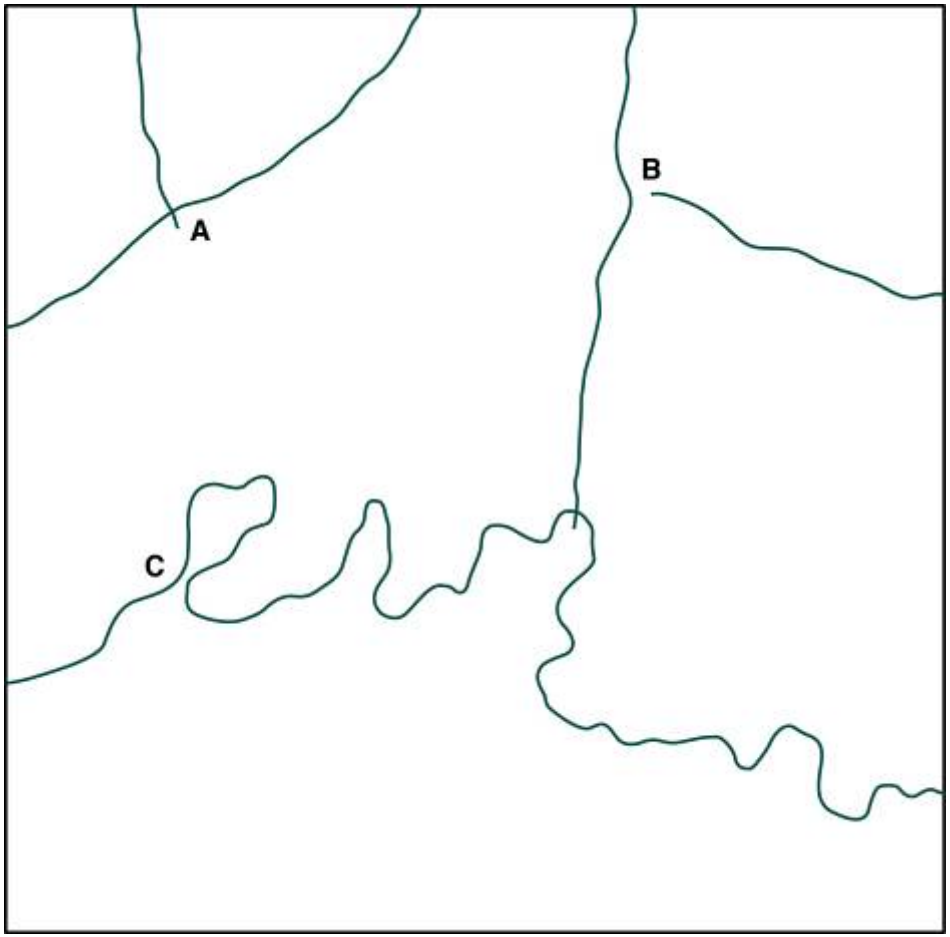


Typology of human errors in digitizing: (A) undershoots and overshoots; (B) invalid polygons; and (C) sliver polygons



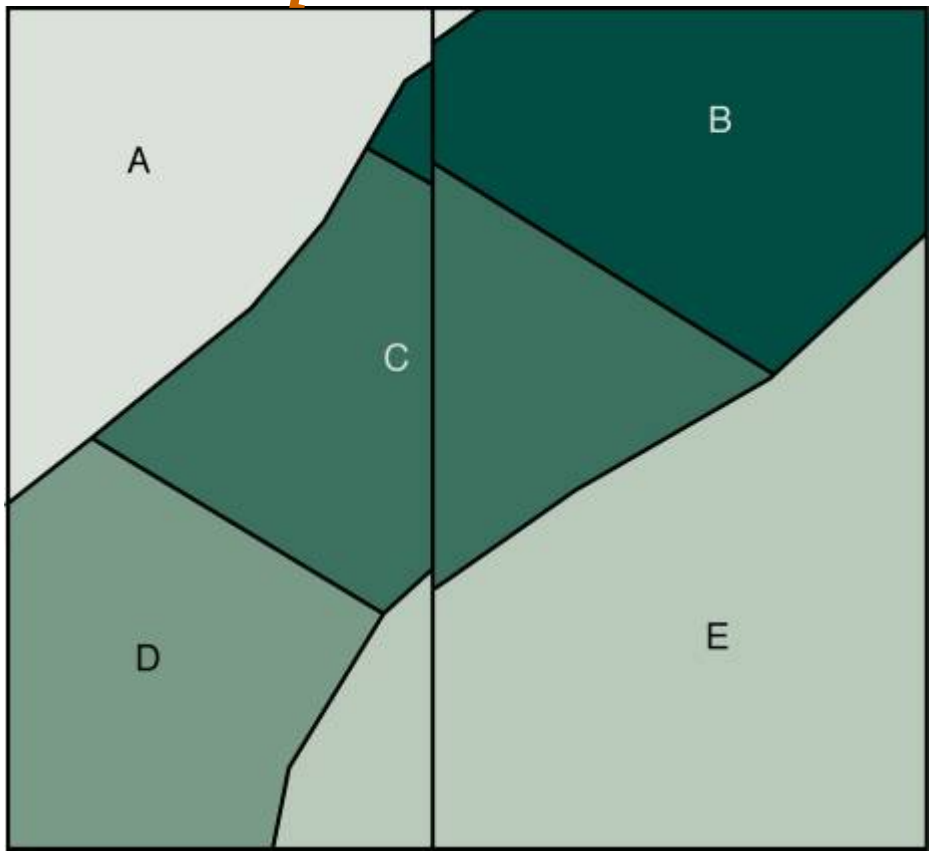
© 2005 John Wiley & Sons, Ltd

Error induced by data cleaning

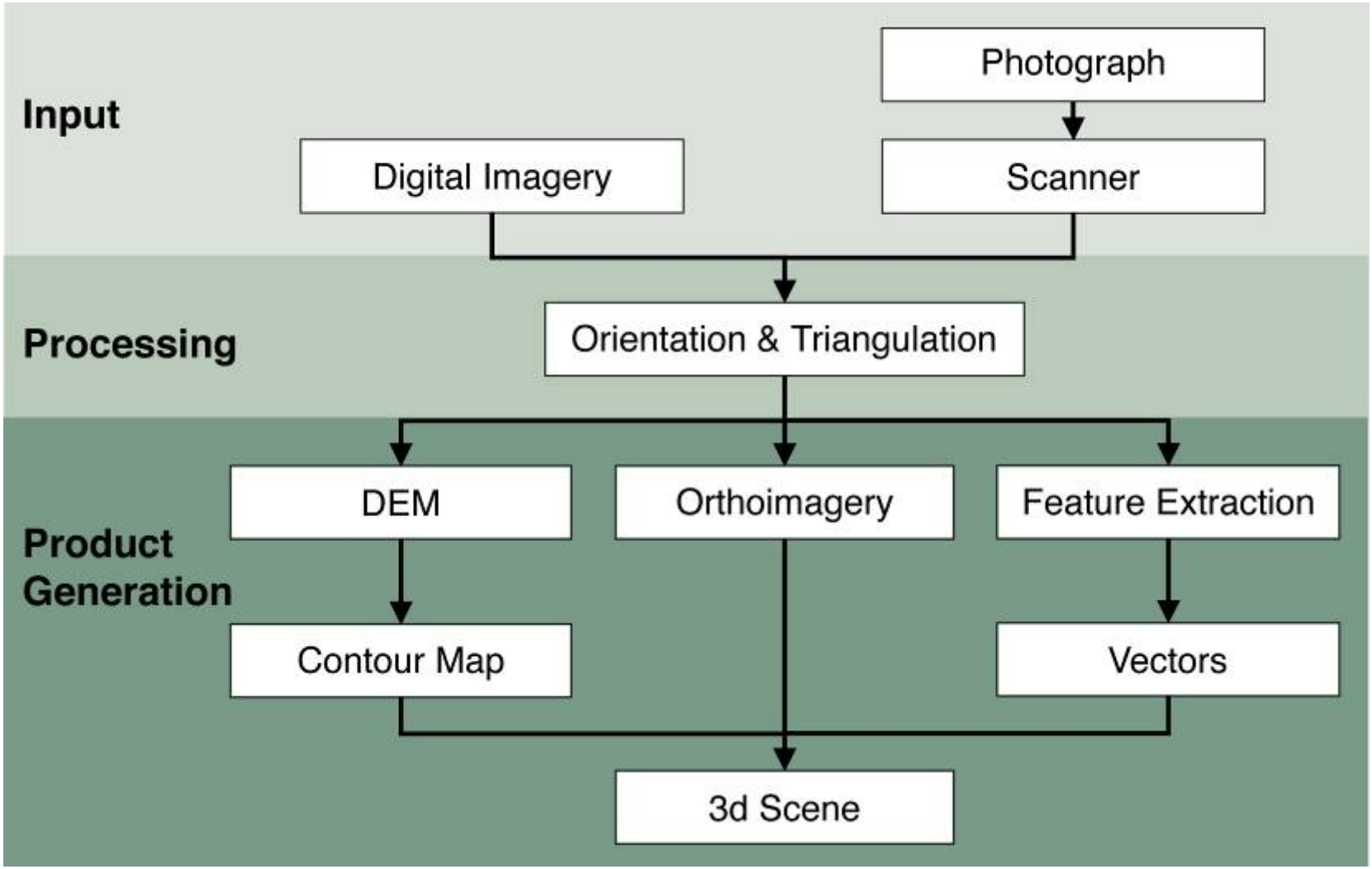


© 2005 John Wiley & Sons, Ltd

Mismatches of adjacent spatial data sources that require rubber-sheeting



© 2005 John Wiley & Sons, Ltd

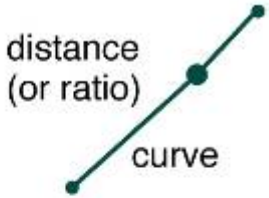


© 2005 John Wiley & Sons, Ltd

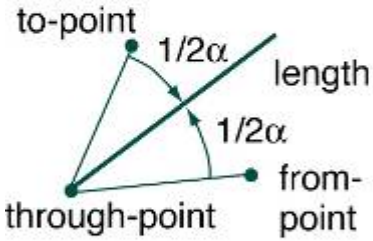
© 2005 John Wiley & Sons, Ltd

COGO construction tools used to represent geographic features

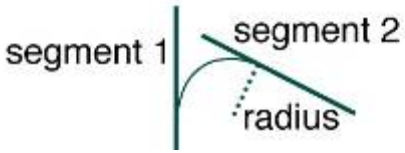
Construct Along



Line Construct Angle Bisector



Construct Fillet



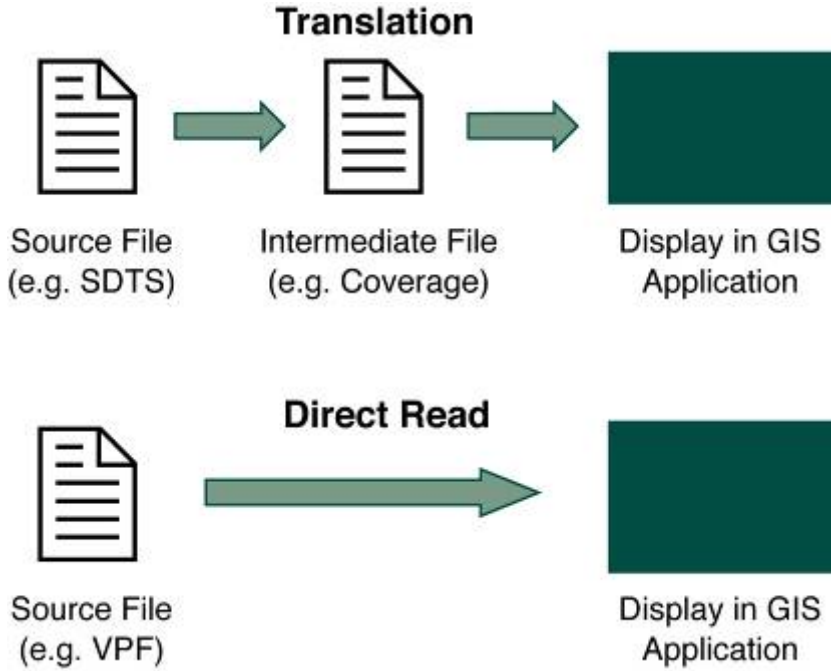
© 2005 John Wiley & Sons, Ltd

Data Transfer

- Buy v build is an important question
- Many widely distributed sources of GI
- Key catalogs include
  - US NSDI Clearinghouse network
  - Geography Network
- Access technologies
  - Translation
  - Direct read

© 2005 John Wiley & Sons, Ltd

Comparison of data access by translation and direct read



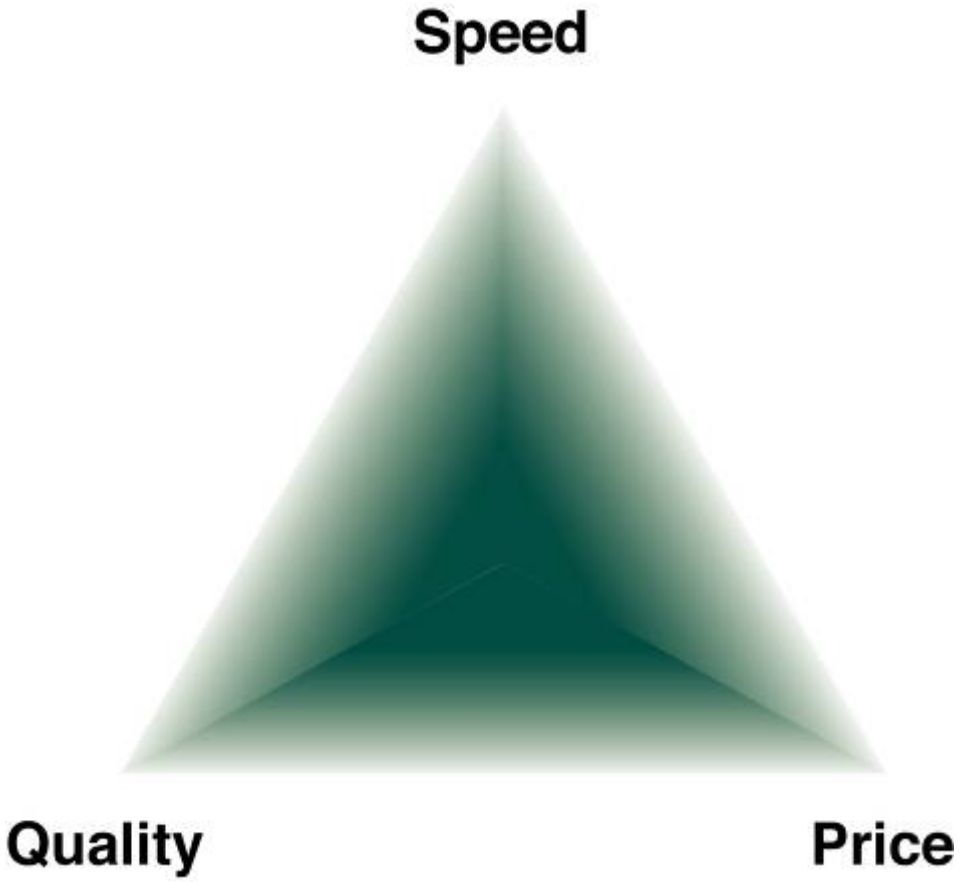
© 2005 John Wiley & Sons, Ltd

Managing Data Capture Projects

- Key principles
  - Clear plan, adequate resources, appropriate funding, and sufficient time
- Fundamental tradeoff between
  - Quality, speed and price
- Two strategies
  - Incremental
  - ‘Blitzkrieg’ (all at once)
- Alternative resource options
  - In house
  - Specialist external agency

© 2005 John Wiley & Sons, Ltd

Figure 9.17 Relationship between quality, speed, and price in data capture (Source: after Hohl 1998).



© 2005 John Wiley & Sons, Ltd

Summary

- Data collection is very expensive, time-consuming, tedious and error prone
- Good procedures required for large scale collection projects
- Main techniques
  - Primary
    - Raster – e.g. remote sensing
    - Vector – e.g. field survey
  - Secondary
    - Raster – e.g. scanning
    - Vector – e.g. table digitizing

© 2005 John Wiley & Sons, Ltd