

Introduktion till GIS i samhällsbyggnad



Thomas Gumbricht
thomas@karttur.com
www.karttur.com

Dagens föreläsning

- Definitioner
- Geografisk information
- Modellering med geografisk information
- Representation av geografisk information
- Skala och upplösning
- ■ Geografisk generalisering
- Datafångst för GIS
- Analog och digital geografi

Definitioner

- ▶ **Geographic Information Systems**

- ▶ A GIS is a computer system capable of assembling, storing, manipulating, managing, and displaying geographic information.

- ▶ **Geographic Information Science**

- ▶ studies fundamental issues arising from the creation, handling, storage and use of geographic information.

- ▶ **Geographic** – Earth's surface and near-surface

- ▶ **Spatial** – any space (including geographic) e.g. medical imaging,

- ▶ **Geospatial** – synonymous with geographic

Komponenter i GIS

Ett komplett GIS byggs upp av sex komponenter

- Hårdvara
- Mjukvara
- Data
- Människor
- Procedurer
- Nätverk



Syfte med GIS

Syftet med GIS

- att integrera och kombinera geospatial data från **olika källor**
- att **manipulera, analysera och visualisera** den kombinerade datan

Nästan allting händer någonstans.

Att veta vart händelser äger rum är viktigt.



GIS syfte

GIS kombinerar geospatial och icke-geospatial information från olika källor i en geospatial analysis operation för att svara på frågorna om:

- Identifiering → Vad finns där?
- Lokalisering → Var är ... ?
- Trender & förändringar → Vad har förändrats sedan ... ?
- Optimala rutter → Vad är bästa vägen från ... till ... ?
- Mönster → Vad är förhållandet mellan ... ?
- Modeller, planering & prognoser → Ifall att ... ?

Geografisk information

Geografisk information



Position & attribut för ett objekt eller fenomen på eller i närheten av Jordens yta



Geografisk information

Geografisk information länkar en plats, och ofta tid, med någon egenskap för en given plats (och tid)

“Temperaturen vid $28^{\circ}00'53''N$, $86^{\circ}55'35''$ på 8848m över havet klockan 7:00 on 2006-03-29 var $-25^{\circ}C$.”



Verklighet och modell

Världen är oändligt komplex

- Innehållet i en databas representerar en begränsad syn på verkligheten; en rumslig databas är en av oändligt många möjliga representationer av modeller av verkligheten
- Ontologiska aspekter
- Epistemological aspekter
- Användarens tillgång till och tolkning av en rumslig databas är via ett gränssnitt

Geografisk information

Antalet **egenskaper är potentiellt** oändliga:

- i GIS kallar vi dem attribut
- attribut kan vara fysiska, sociala, ekonomiska, demografiska, miljörelaterade etc.

Antalet **platser och tider** är ävenledes oändliga:

- i GIS anger vi dem i ett koordinatsystem

Ju närmare vi betraktar jordytan, ju **mer detaljer** ser vi:

- potentiellt ad infinitum
- geografi är således oändligt komplext

Människan har alltid använt olika metoder för att hantera geografisk komplexitet:

GIS är den digitala erans metod för att hantera geografi och dess komplexitet

“Spatial is special”

Närliggande objekt är ofta mer relaterade än fjärran objekt.

Stora datamängder

Multidimensionella

Spatial autokorrelation

Voluminöst

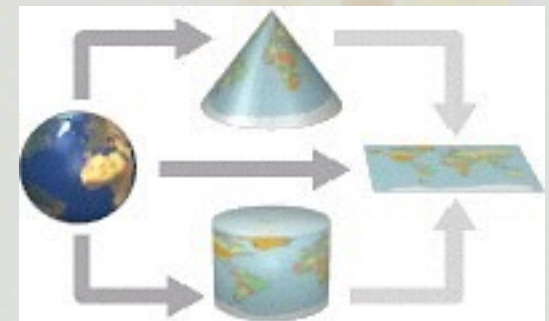
Karaktäristiskt för geografisk information

Olika skalor av spatial upplösning

kräver **projektion** till platt yta

Speciell former för representation i dator

Unika analysmetoder



Verklighet och modell

En rumslig databas kan innehålla

- Digitala abstraktioner av verkliga objekt
 - ex.v. land, vatten, hus, vägar, träd
- Digitala abstraktioner av fiktiva objekt
 - ex.v. politiska gränser, ekosystem

“Spatial is special”

Geografisk information är

multidimensionell

Geometriska attributer

Position

x, y, z

longitude,
latitude,
elevation

Dimensioner

storlek, form

Temporala attribut

Föränderliga över tid

Icke-spatiala attribut

Kvantitativa

fysisk,
social,
ekonomisk,
demografisk,
miljörelaterad,
etc

Kvalitativa

Datatyper i GIS

Mätskalor för icke-spatiala attribute

Nominal skala

Språk:
English
Svenska
Deutch
Espanol

Ordinal skala

Klimat:
hett
varmt
kyligt
kallt

Intervall skala

Temperatur:
°C skala

kvot skala

Befolkningstäthet:
antalet invånare/km²

Cyklisk skala

Riktning (horisontell):
lutningsriktning, vindriktning

Modellering med Geografiska informationssystem

▶ Diskreta objekt

- Delar upp världen i diskreta objekt, såsom byggnader, vägar, politiska gränser etc.
- Representaras av matematiska vektorer

▶ Kontinuerliga fält

- Nästan alla naturliga fenomen har en kontinuerlig utbredning i rummet - ex.v. temperature och lufttryck.
- Representeras av matematiska matriser

Den mest fundamentala uppdelningen i geografisk representation.

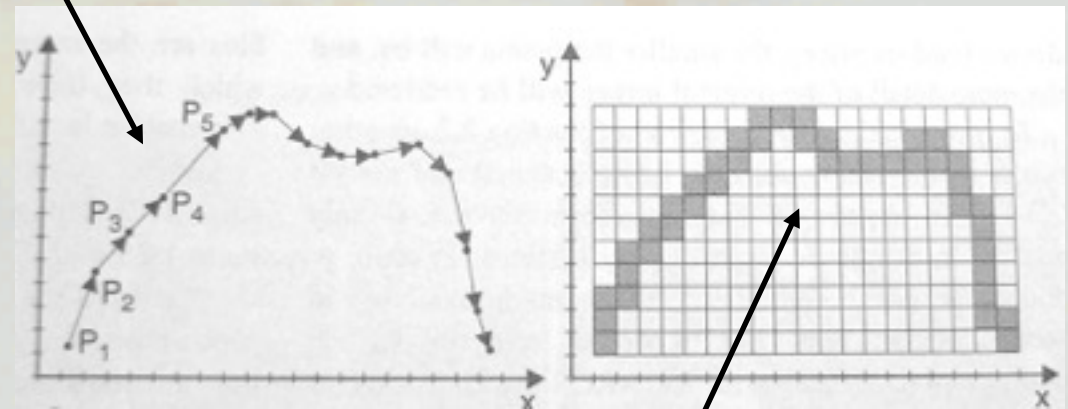
GIS data modeller

Linjer och gränser definieras som en serie kopplade punkter.

Vektor data struktur

Grundläggande enhet:
Geografiska objekt

D i g i t a l
representation
av geografisk
information



Punkter, linjer och gränser definieras som en serie bild-element (**pixels**).

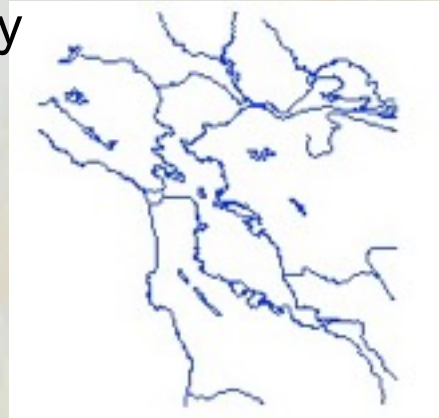
Raster data struktur

Grundläggande enhet: **pixel**

GIS data modeller - exempel

Examples

Hydrology



Road network

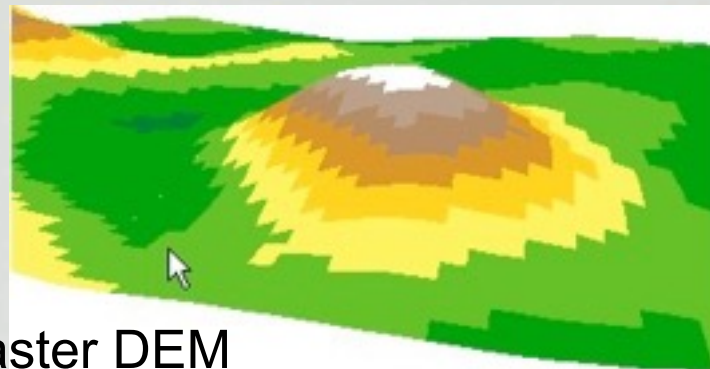


Vector data structure

Remote sensing imagery



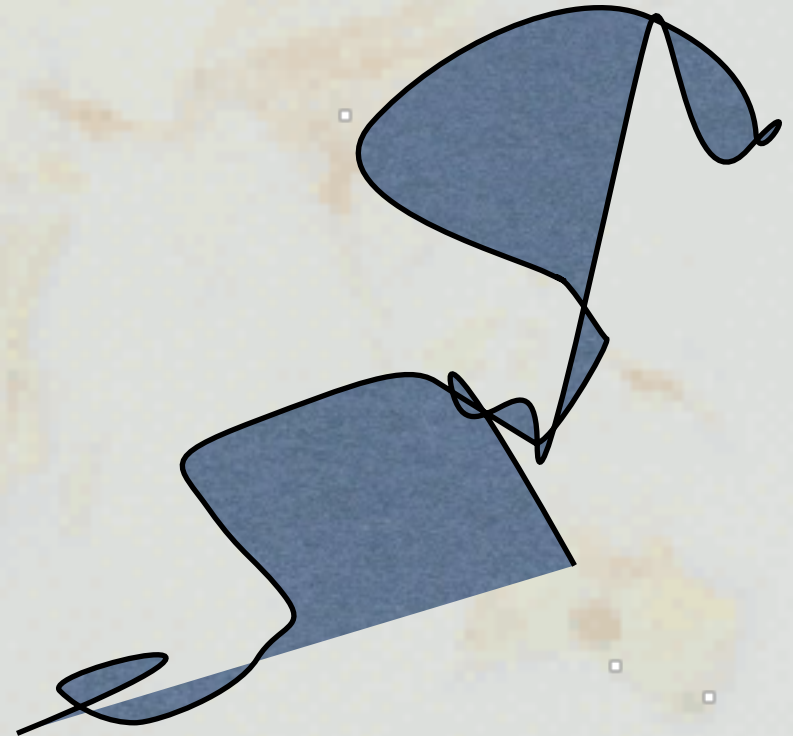
Raster data structure



Raster DEM

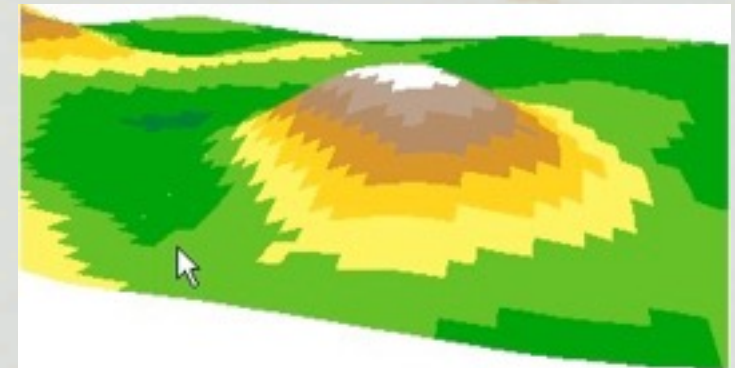
Verklighet och modell

Datorer är bra på att lagra diskret data, men sämre på att lagra kontinuerlig data - till syvende och sist är allt lagrat som 1 eller 0.

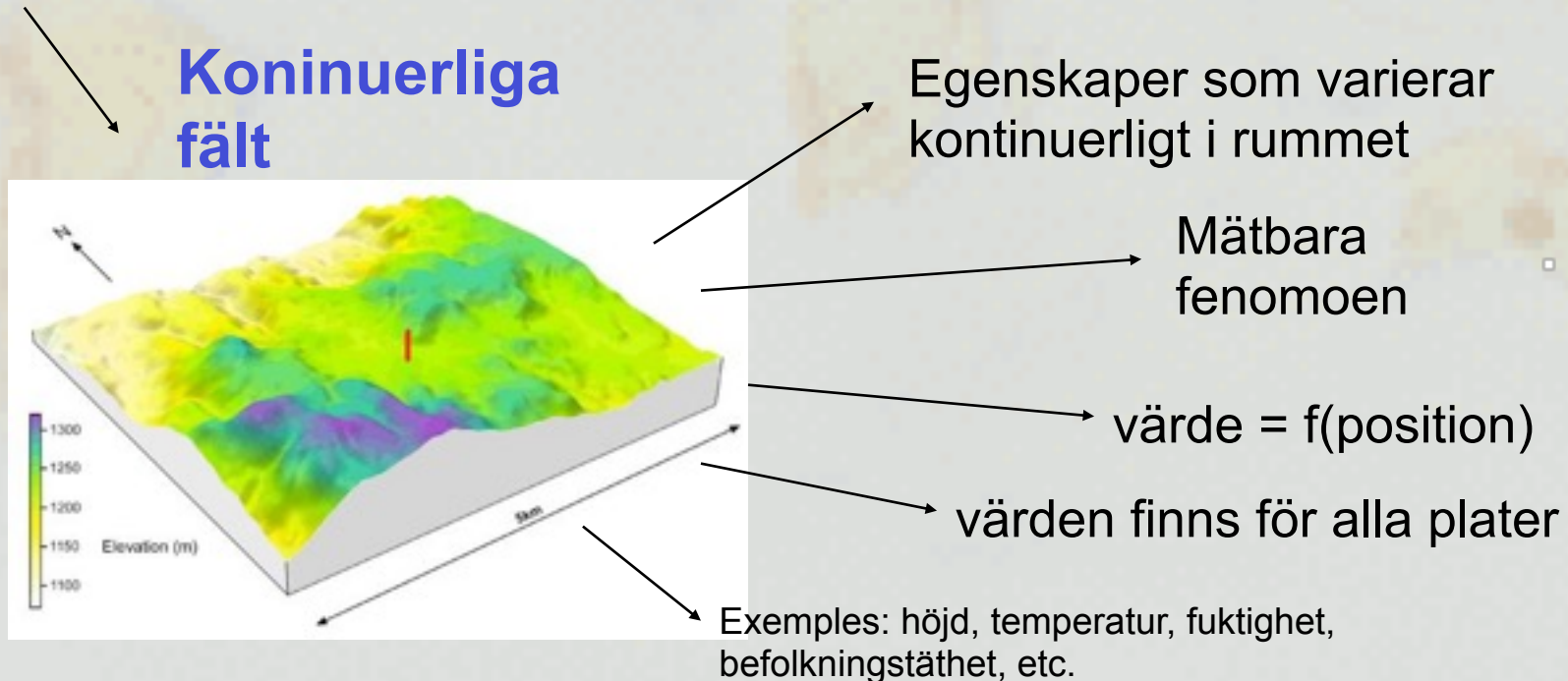
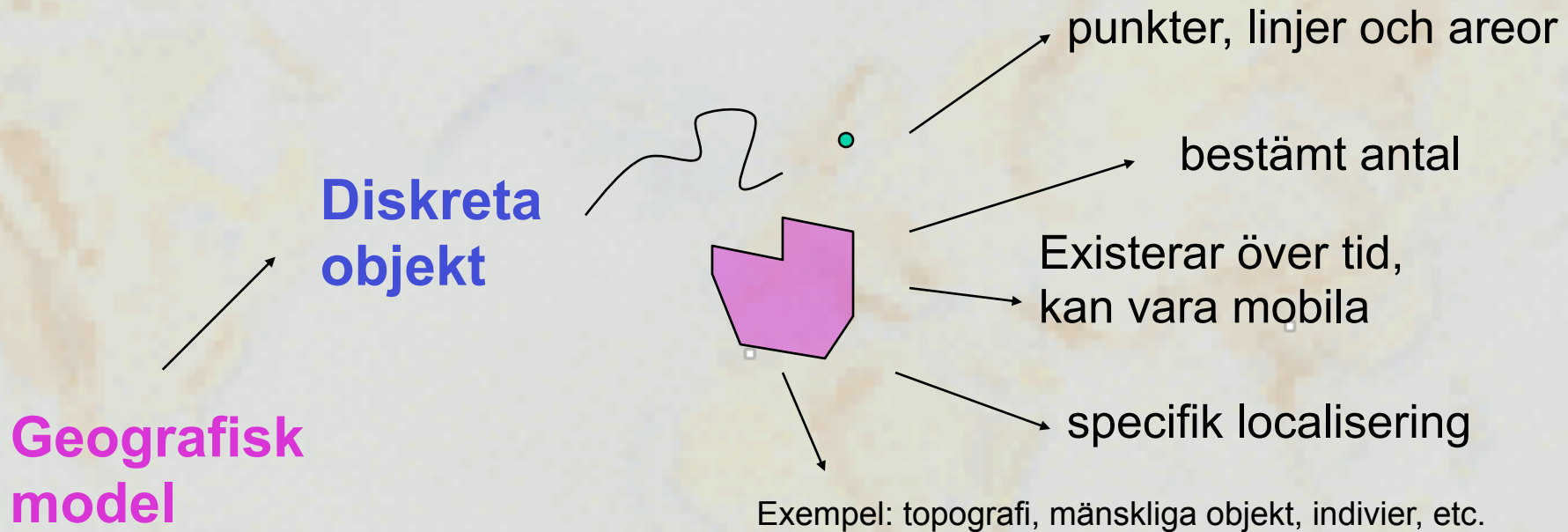


Verklighet och modell

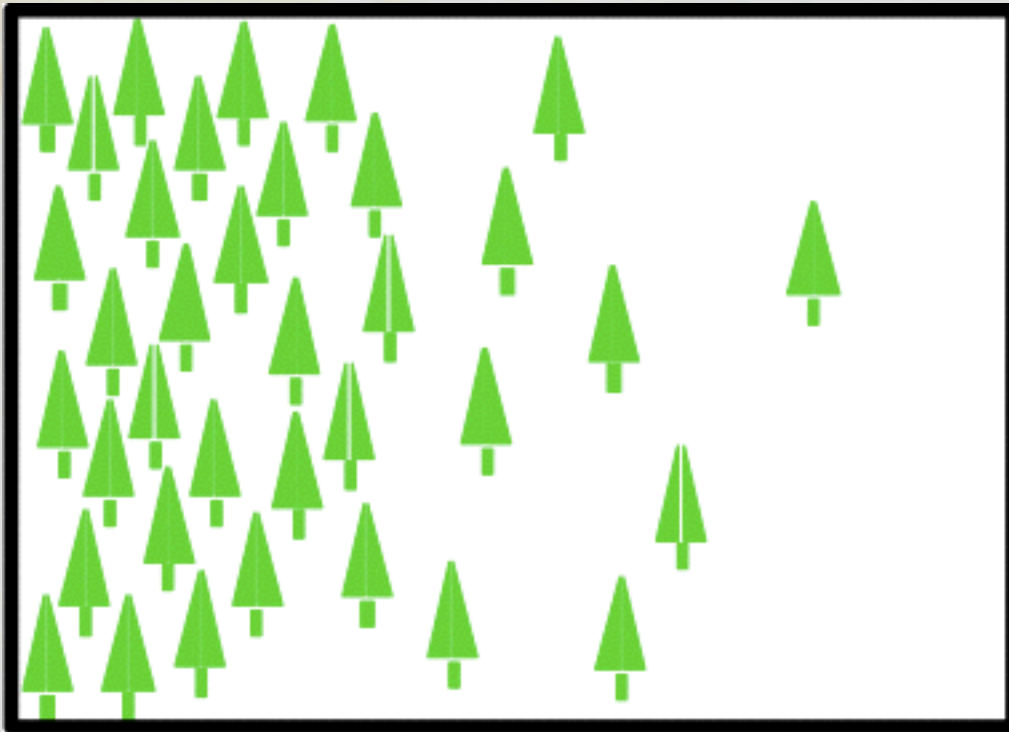
- Objekt som är av diskret natur, hus, vägar, distrikt etc, vållar inga problem att representera som diskreta objekt.
- Egenskaper som finns överallt och som varierar kontinuerligt, elevation, temperatur, lufttryck, måste approximeras till en diskret representation.



Verklighet och modell



Verklighet och modell



Diskreteringen av kontinuerliga fenomen
är ofta godtycklig

Objektmodell och fältmodell

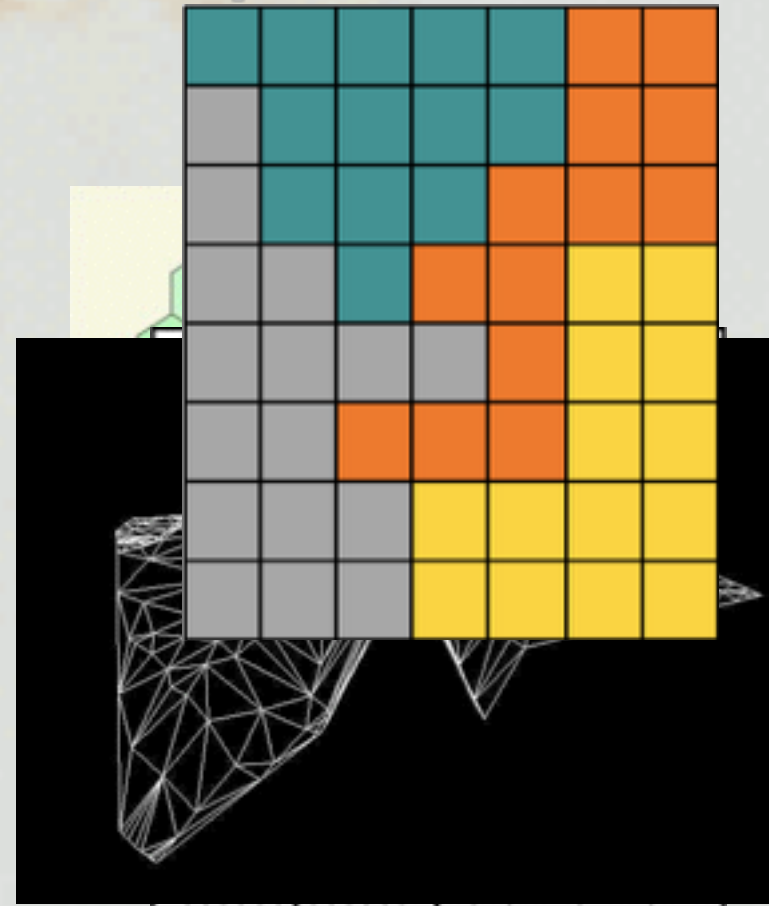
GIS-samhället har utvecklat konceptuella modeller av verkligheten, sprungna ur kartografi snarare än datalogi:

- Objektmodell - punkter, linjer, ytor fyller upp alla delar av rummet
- Fältmodell - Värderna för varje position

Tesseleringsmodeller

Raster data modellen tillhör en större grupp av fältdatamodeller eller tesseleringsmodeller:

- Grid eller raster
- Hexagonaler
- Triangular Irregular Network (TIN)
- Kvadratträd



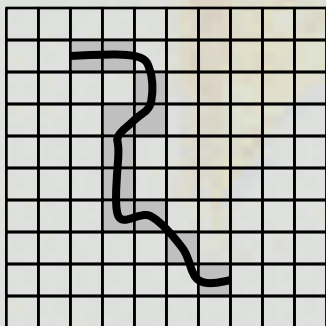
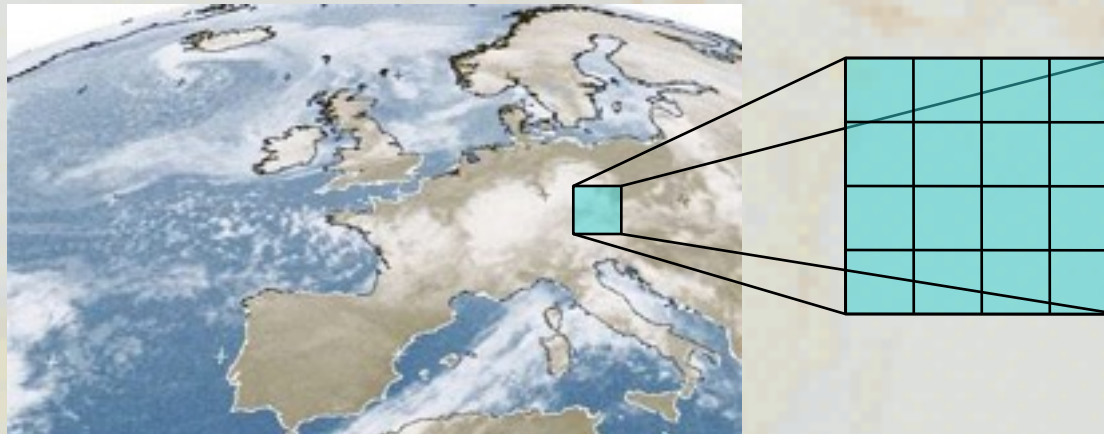
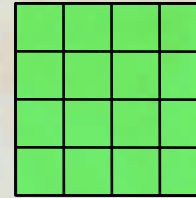
Fältmodell

Raster = regelbunden tesselering

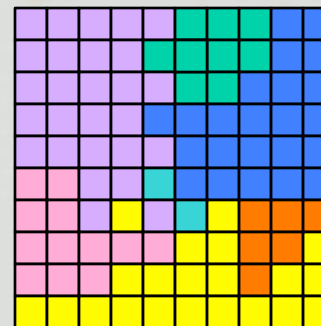
- Delar upp världen i rektangulära celler
- Registrerar grid-hörnen till en geografisk punkt
- Representerar diskreta objekt som grupper av celler med eller utan attributkoppling (koppling via indexnummer)
- Representerar fält som cellvärden (utan attributkoppling)
- Värden för varje cell
- Även celler utan relevant data lagras, som “ingen data”
- Vanligare att använda för fältobjekt
- Lätt att förstå

Raster data struktur

- Delar upp världen i rektangulära **celler = pixlar**
- Registrerar grid-hörnen till en geografisk punkt

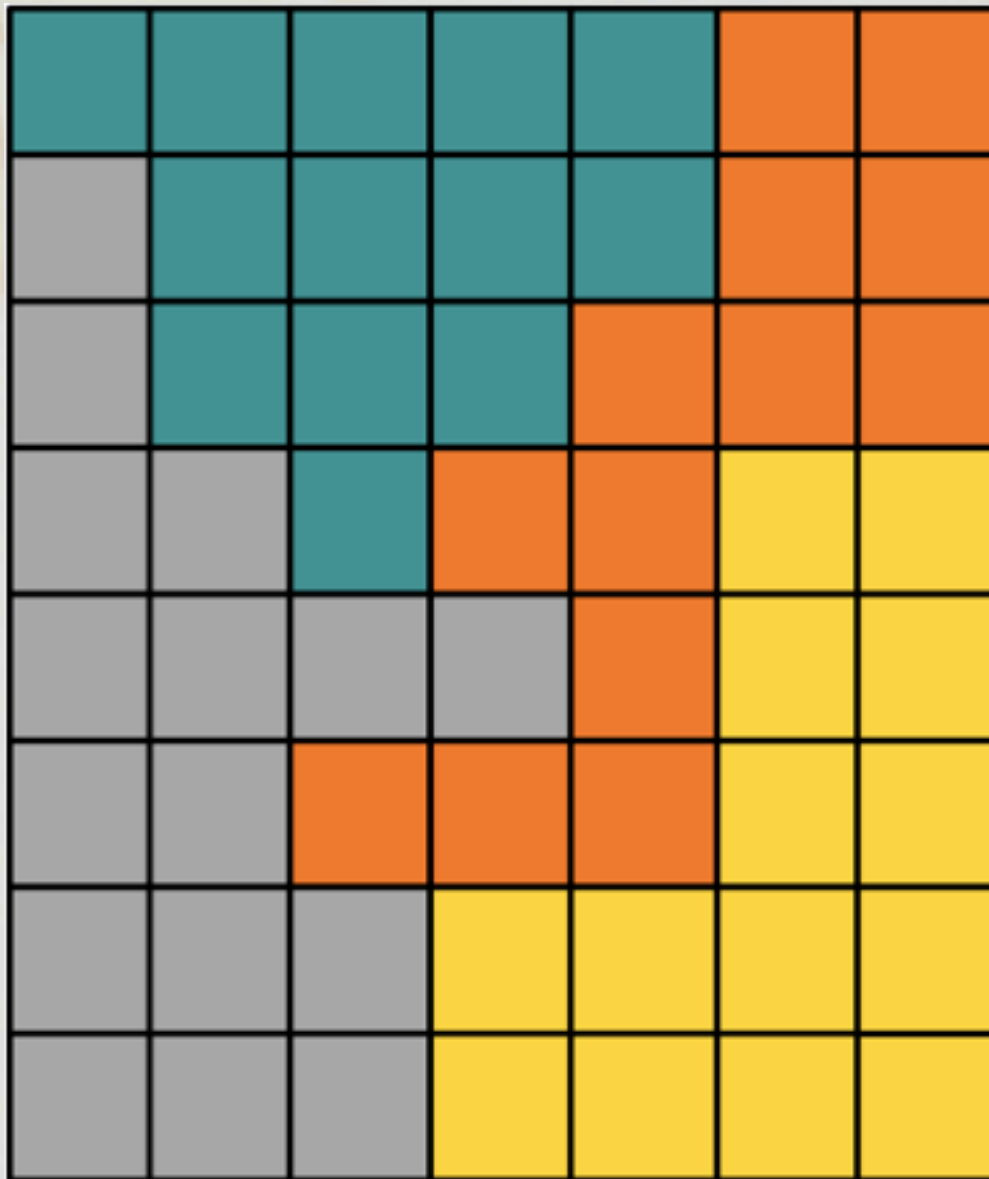


diskreta objekt
grupper av celler



Kontinuerliga fält
cellvärdet = fältvärdet

Raster data struktur

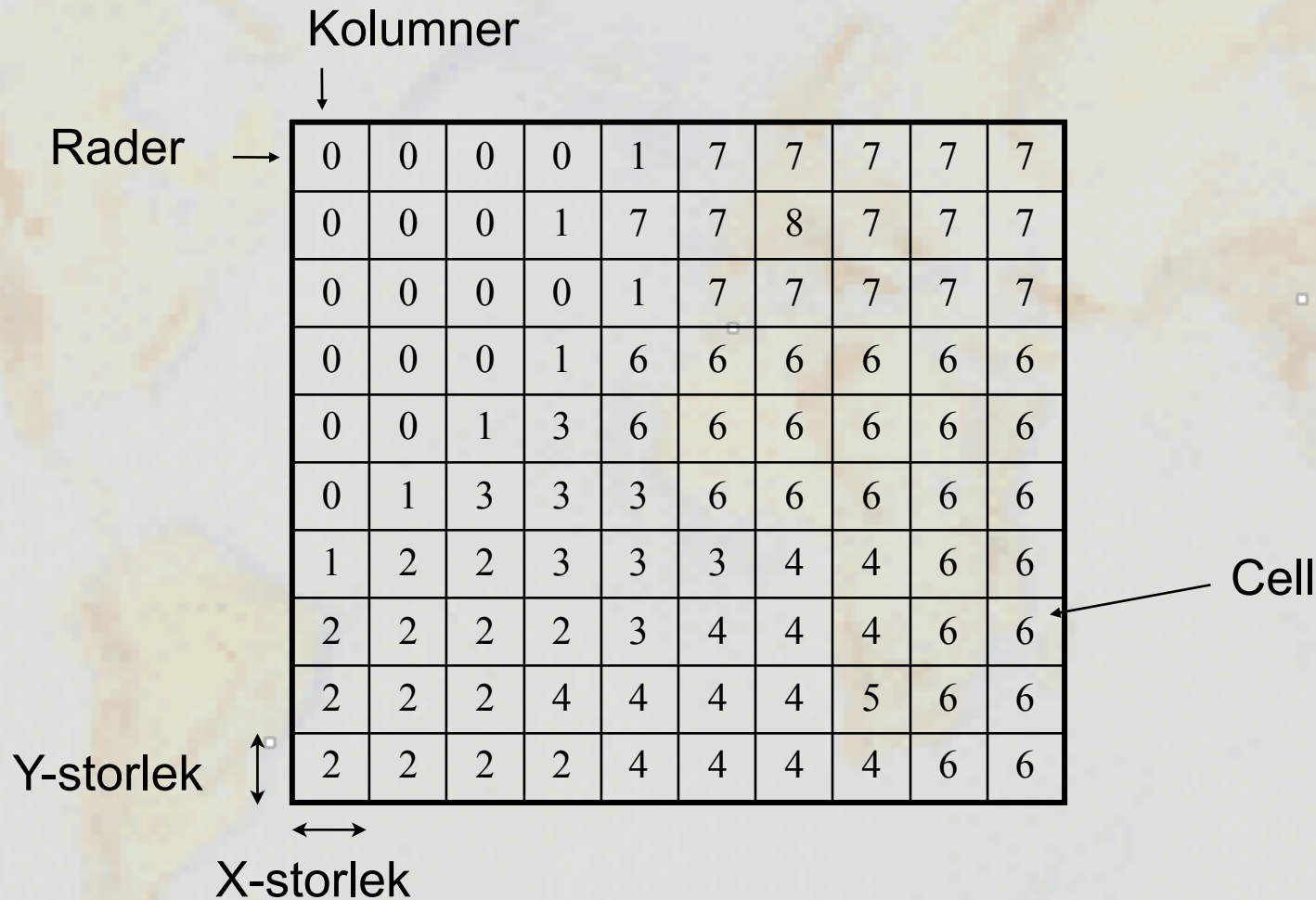


Raster data struktur
för diskreta objekt -
markanvändning.

Raster

- Pixel storlek
 - Storleken på cellen eller bildelementen som definierar den rumsliga detaljeringsgraden/ upplösningen
 - kan vara olika i x och y
- Tilldelning av cellvärden - värdet på en cell kan representera
 - medelvärdet för cellens yta
 - typvärdet för cellens yta
 - mittvärdet för cellets yta

Raster data struktur



Cell/pixel storlek = **rumslig upplösning**

- definierar **detaljnivå för rumsliga objekt**
- variationer **inuti pixeln** går förlorad

Raster

Mixade pixlar - ett problem med raster



Vatten dominant

V	V	L
V	V	L
V	V	L

Segraren tar allt

V	L	L
V	L	L
V	L	L

ekotoner som egen klass

V	E	L
V	E	L
V	E	L

Raster

Lagring av rasterdata

- Sekventiell lagring
 - Byte Interleaved by Pixel - BIP
 - Byte Interleaved by Line - BIL
 - Band Sequential - BSQ
- Blockkodning
- Kedjekodning
- Radlängdskodning
- Kvadratträd

Raster

Filstorlek

$\text{rader} * \text{kolumner} * \text{Byte per pixel} = \text{filstorlek}$

Storlek på fil med byte-värden (1 byte per pixel)

$\text{rader} * \text{kolumner}$

Raster

metadata och huvudfil

Exempel I: Byte data
ERmapper

```
DatasetHeader Begin
  Version          = "5.5"
  Description      = "NOAA-AVHRR NDVI annual average "
  DataSetType     = ERStorage
  DataType        = Raster
  ByteOrder       = LSBFirst
  CoordinateSpace Begin
    Datum          = "CLARKE 1866"
    Projection     = "ALBERSEA"
    CoordinateType = EN
    Rotation       = 0:0:0.0
  CoordinateSpace End
  RasterInfo Begin
    CellType       = Unsigned8BitInteger
    NullCellValue = 0
    CellInfo Begin
      Xdimension   = 8000
      Ydimension   = 8000
    CellInfo End
    NrOfLines     = 360
    NrOfCellsPerLine = 450
    RegistrationCoord Begin
      Eastings     = -3920000
      Northings    = 3250000
    RegistrationCoord End
    NrOfBands     = 1
    BandId Begin
      Value        = "Pseudo"
    BandId End
  RasterInfo End
DatasetHeader End
```


Raster

metadata och huvudfil

Exempel 1: Byte data
ArcView

```
;
;ArcView Image Information
; NOAA-AVHRR NDVI annual average
; Projection: ALBERS (Albers Equal Area Conic)
; Units: METERS
; Spheroid: CLARKE1866
; 1st standard parallel (dms): -19 00 0.000
; 2nd standard parallel (dms): 21 00 0.000
; central meridian (dms): 20 00 0.000
; latitude of projection origin: 1 00 0.000
; false easting (meters): 0.00000
; false northing (meters): 0.00000
;
NCOLS 450
NROWS 360
NBANDS 1
NBITS 8
LAYOUT BIL
BYTEORDER 1
SKIPBYTES 0
MAPUNITS METERS
ULXMAP -3916000
ULYMAP 3246000
XDIM 8000.00000
YDIM 8000.00000
```

Raster

metadata och huvudfil

Exempel I: Byte data
IDRISI

file format : IDRISI Raster A. I
file title : NOAA-AVHRR NDVI annual average
data type : byte
file type : binary
columns : 450
rows : 360
ref. system : albersaf
ref. units : m
unit dist. : 1.0000000
min. X : -3920000.0000000
max. X : -320000.0000000
min. Y : 370000.0000000
max. Y : 3250000.0000000
pos'n error : unknown
resolution : 8000.0000000
min. value : 0
max. value : 255
display min : 0
display max : 255
value units : unspecified
value error : unknown
flag value : none
flag def'n : none
legend cats : 0

Raster

metadata och huvudfil

Exempel 1: Byte data ENVI

ENVI

description = {
 NOAA-AVHRR NDVI annual average }

samples = 450

lines = 360

bands = 1

header offset = 0

file type = ENVI Standard

data type = 1

interleave = bsq

sensor type = AVHRR

byte order = 0

map info = {Albers_NDVI_ADDS, 1.0000, 1.0000, -3916000, 3246000, 8.0000000000e+003, 8.0000000000e+003, , units=Meters}

projection info = {9, 6378206.4, 6356583.8, 1.000000, 20.000000, 0.0, 0.0, -19.000000, 21.000000, Albers_NDVI_ADDS, units=Meters}

wavelength units = Unknown

band names = {

NDVI}

Raster

metadata och huvudfil

Exempel 1: Byte data
DIVA

```
Version=4.1
Title=NDVlg Annual mean 2004
Created=20050306
[GeoReference]
Projection=ALBERS
Datum=CLARKE1866
Mapunits=m
Columns=450
Rows=360
MinX=-3920000
MaxX=-32000
MinY=37000
MaxY=3250000
ResolutionX=8000
ResolutionY=8000
[Data]
DataType=BYTE
MinValue=0
MaxValue=255
NoDataValue=-9999
Transparent=1
Units=NDVI
[Application]
```

Raster

metadata och huvudfil

Exempel 1: Byte data

JPG (*.jpw, *.jpgw)

TIF (*.tfw)

BMP (*.bmpw)

8000

0

0

-8000

-3916000

3246000



Raster

metadata och huvudfil

Exempel 2: Integer data
ERmapper

```
DatasetHeader Begin
  Version          = "5.5"
  Description      = "NOAA-AVHRR NDVI annual npp "
  DataSetType     = ERStorage
  DataType        = Raster
  ByteOrder       = LSBFirst
  CoordinateSpace Begin
    Datum          = "CLARKE 1866"
    Projection     = "ALBERSEA"
    CoordinateType = EN
    Rotation       = 0:0:0.0
  CoordinateSpace End
  RasterInfo Begin
    CellType       = Unsigned 16BitInteger
    NullCellValue = 0
    CellInfo Begin
      Xdimension   = 8000
      Ydimension   = 8000
    CellInfo End
    NrOfLines     = 360
    NrOfCellsPerLine = 450
    RegistrationCoord Begin
      Eastings     = -3920000
      Northings    = 3250000
    RegistrationCoord End
    NrOfBands     = 1
    BandId Begin
      Value        = "Pseudo"
    BandId End
  RasterInfo End
DatasetHeader End
```

Raster

metadata och huvudfil

Exempel 3: real data
ERmapper

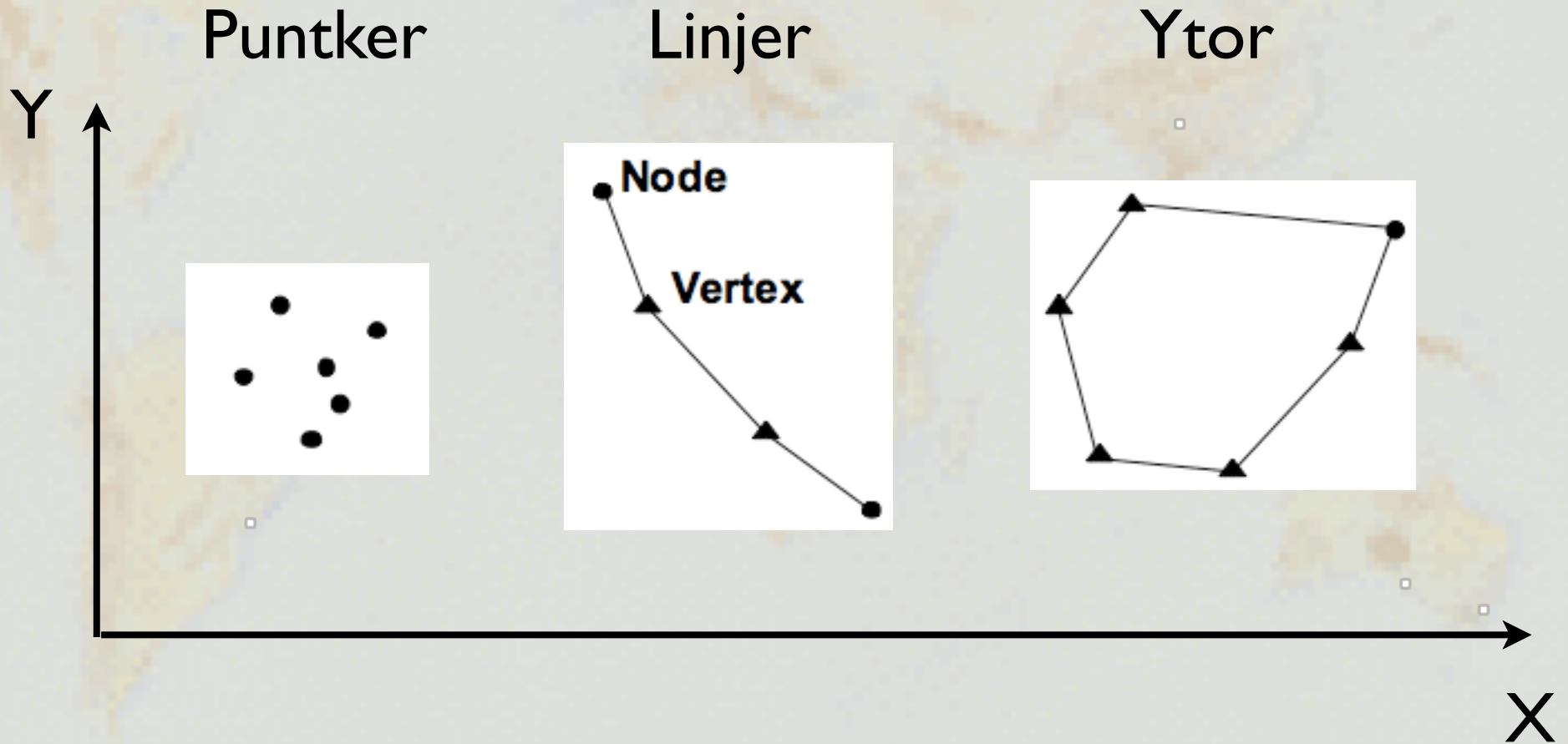
```
DatasetHeader Begin
  Version          = "5.5"
  Description      = "NDVI annual max trend 1982-2004"
  DataSetType     = ERStorage
  DataType        = Raster
  ByteOrder       = LSBFirst
  CoordinateSpace Begin
    Datum          = "CLARKE 1866"
    Projection     = "ALBERSEA"
    CoordinateType = EN
    Rotation       = 0:0:0.0
  CoordinateSpace End
  RasterInfo Begin
    CellType       = IEEE32REAL
    NullCellValue = 0
    CellInfo Begin
      Xdimension   = 8000
      Ydimension   = 8000
    CellInfo End
    NrOfLines     = 360
    NrOfCellsPerLine = 450
    RegistrationCoord Begin
      Eastings     = -3920000
      Northings    = 3250000
    RegistrationCoord End
    NrOfBands     = 1
    BandId Begin
      Value        = "Pseudo"
    BandId End
  RasterInfo End
DatasetHeader End
```

Vektor data model

Verkliga eller fiktiva objekt representerade som punkter, linjer och ytor

- punkter representerar objekt utan utbredning, eller med för skalan irrelevant utbredning
- linjer knyter samman punkter till start-, bryt-, och stoppunkter
- ytor (polygoner) byggs upp av slutna linjer

Vektor data model



Vektor data model

Precision och noggrannhet

- Objekt definieras av x,y koordinater relaterade till ett koordinatsystem (long/lat eller x,y).
- Precision (upplösning) i koordinater beror på binär lagringsform (6-15 decimaler), men är ofta hög
- Noggrannheten i data oftast mer begränsande än upplösning

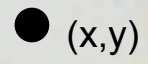
Vektor data model

Precision och noggrannhet

- Precision är det minsta avstånd mellan två intilliggande objekt som uppmätts och lagrats.
- Noggrannhet är frånvaro av fel
- Osäkerhet är ett mer generellt begrepp, och inkluderar både precision och noggrannhet.

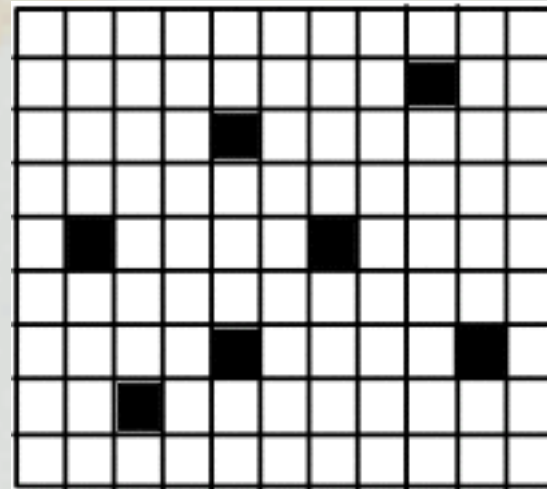
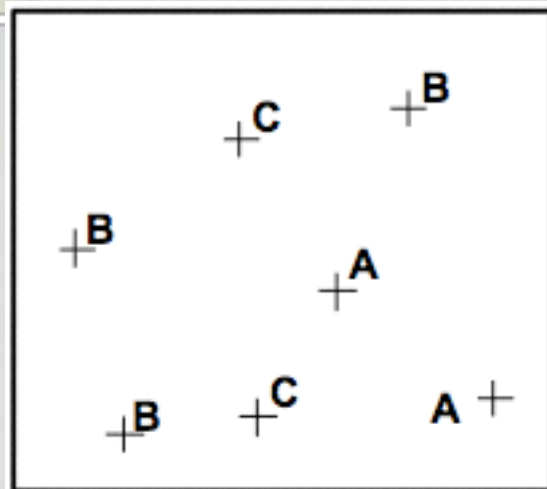
Vektor data model

Punktdata



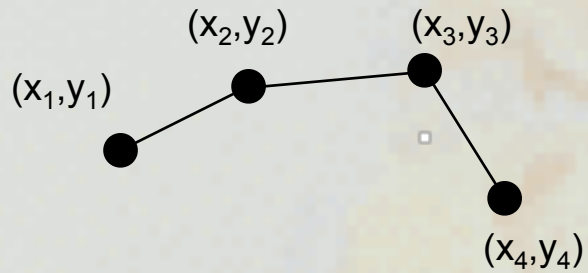
(x,y)

Flaggstång
Byggnad
Stad

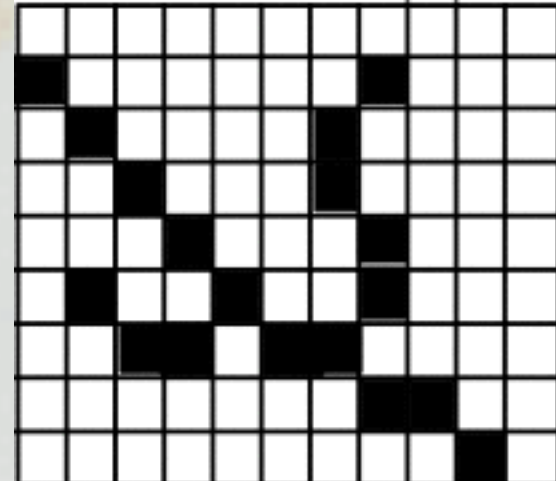
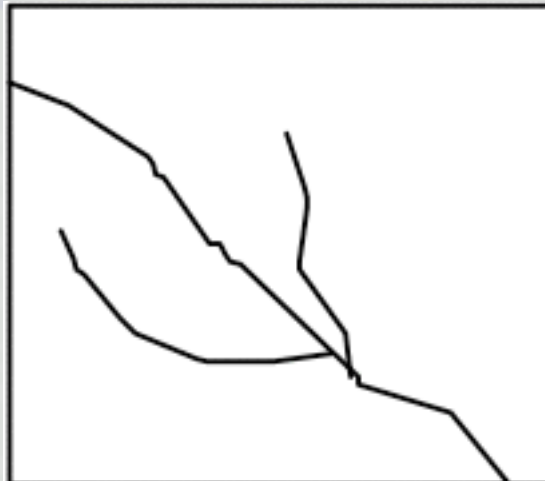


Vektor data model

Linjedata

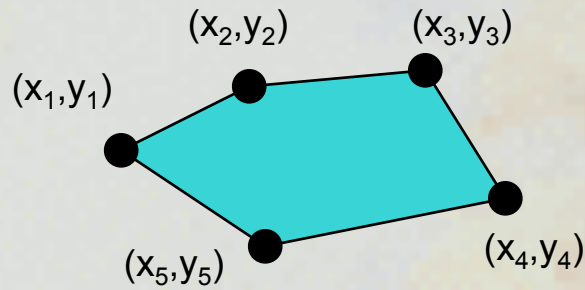


- Vattendrag
- Väg
- Järnväg
- Staket

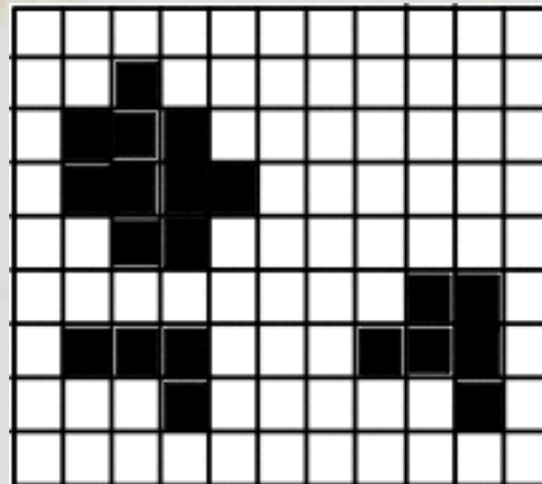
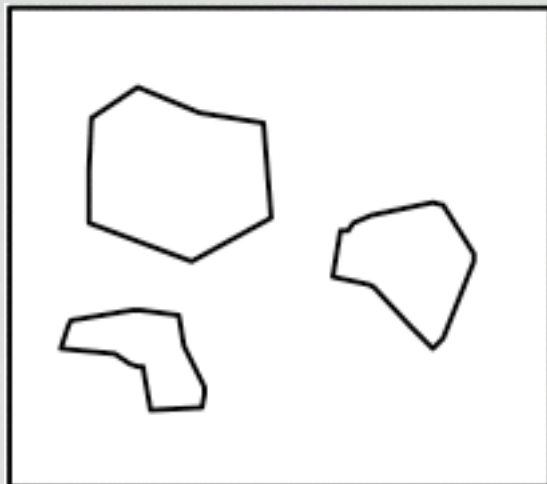


Vektor data model

Areadata



Sjö
Skog
Stad
Fastighet



Vektor data model

Tre huvudsakliga modeller för att lagra vektorer

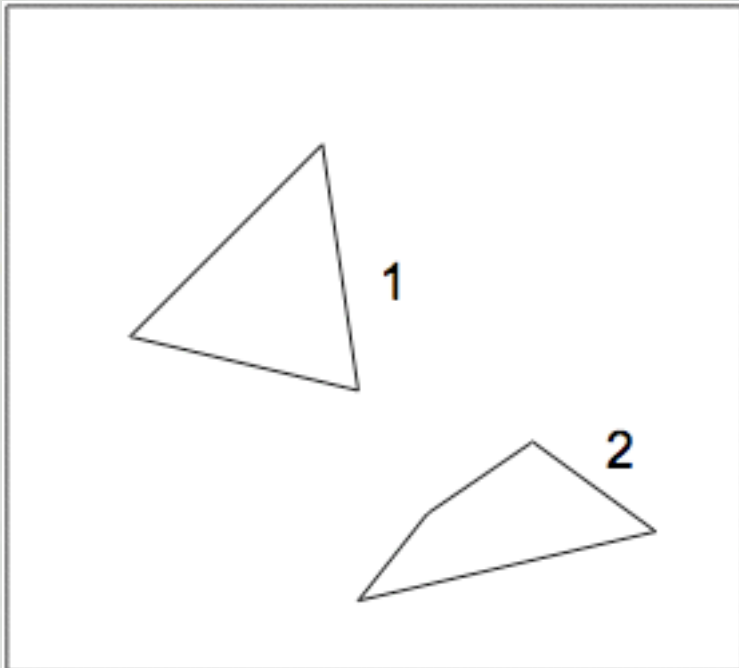
- Enkel (eller spaghetti) data struktur
 - Ingen logik, dubblering av data (inom ett lager)
- Punkt listor
 - Ingen logik, ingen dubblering
- topologisk struktur
 - Logik, ingen dubblering

Vektor data model

Spaghetti vektor data model

Varje punkt, linje eller polygon lagras i en post (“record”) som innehåller Id och koordinater som definierar geometri (de första GIS-programmen hade spaghetti data struktur)

Polygoner



ID	Coordinates
1	(2,4), (4,3), (3,6) , (2,4),
2	(3,1), (5,2), (4,3), (3,2), (3,1)

Vektor data model

Spaghetti vektor data model

■ Fördelar

- enkelt

- effektiv för display och utskrift

■ Nackdelar

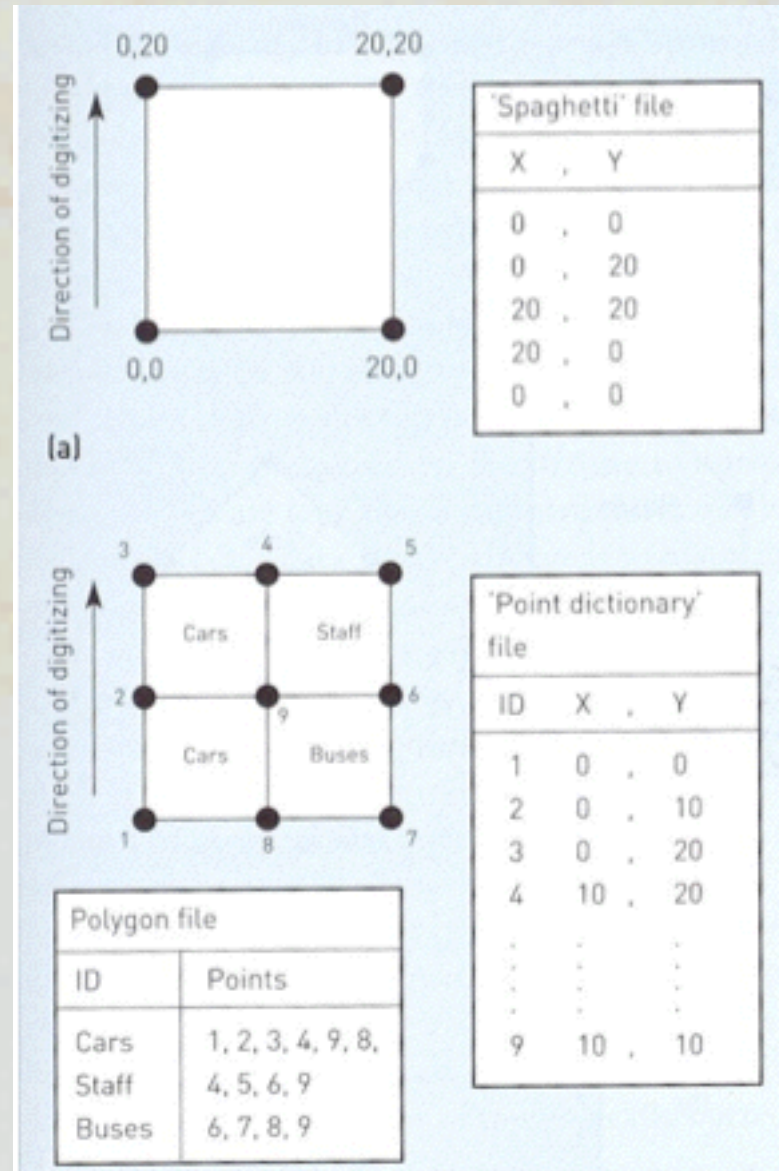
- Ineffektivt för rumsliga analyser

- och generaliseringar

Vektor data model

Punkt data struktur

Ingen data redundans
Ingen topologi



Vektor data model

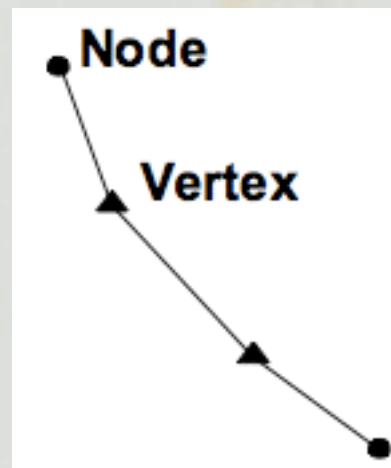
Topologisk data struktur

Nätverkstopologi

kallas även “ark-nod” modellen

ark = linje

nod = slutpunkt på en linje, eller en punkt där en linje splittras eller linjer går samman



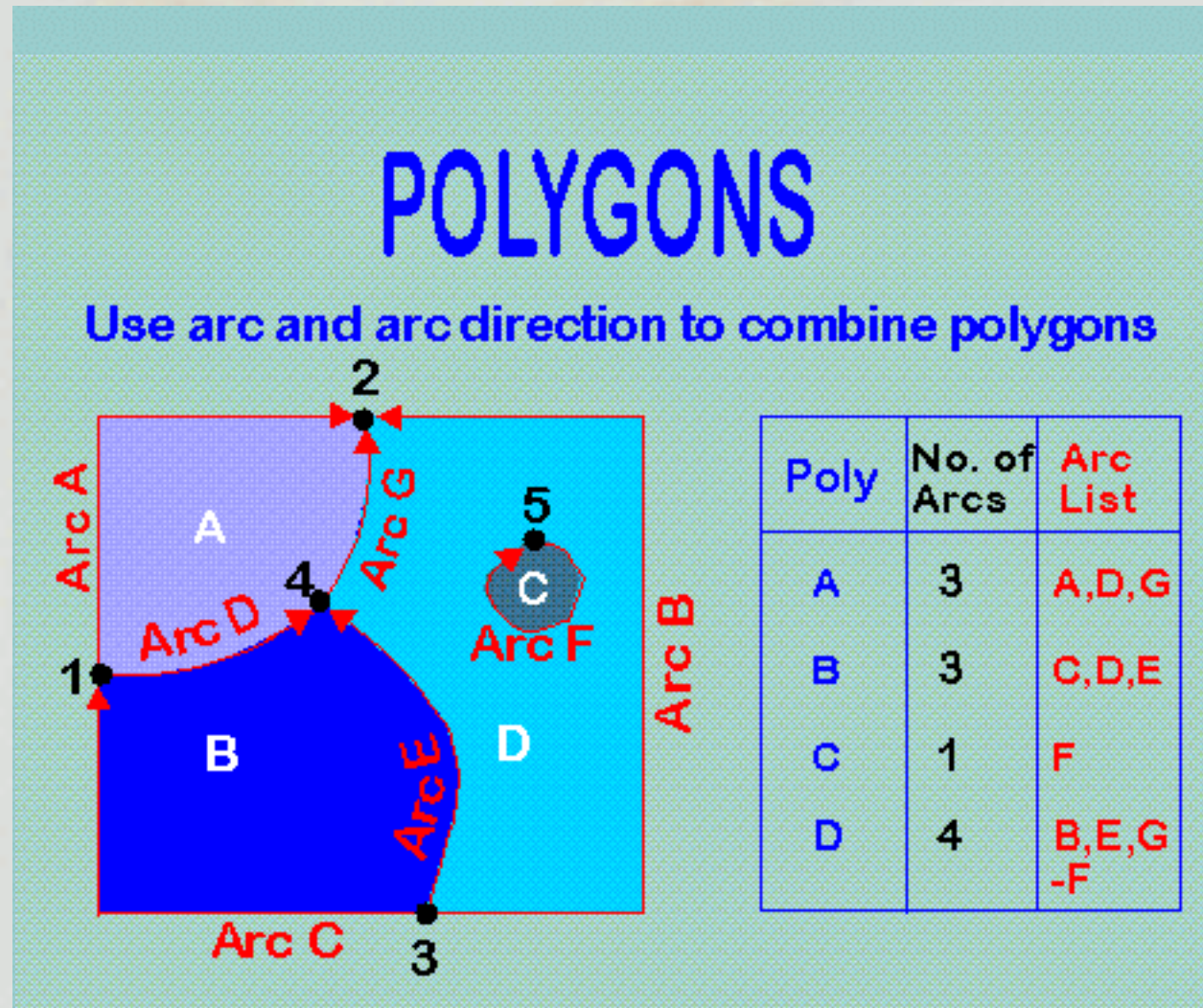
Vektor data model

Topologisk data struktur

- registrerar x/y koordinater av rumsliga objekt

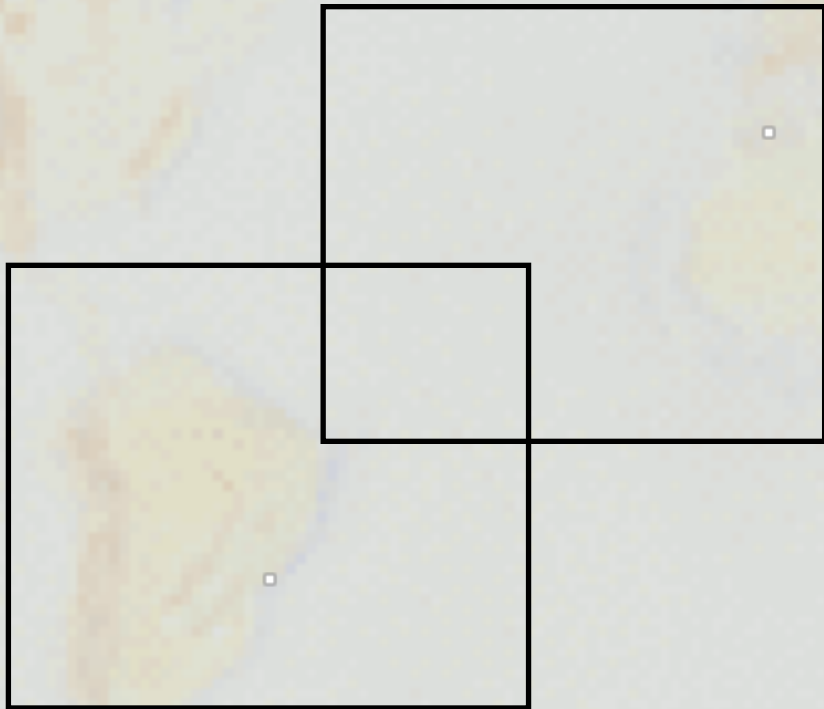
Kodar rumsliga relationer:

- vilka arkar kopplar till vilken nod
- vilka ytor ligger på sidorna av en ark
- vilka arkar bygger en polygon



Vektor data model

Spaghetti modell och topologisk modell



Spaghetti: registrering
som 2 eller 3 ytor

Topologiskt: registrering
som 3 ytor

Vektor data model

topologisk vektor data model

- **Fördelar**
 - Rumsliga relationer är explicita
 - Rumslig analys utan koordinater möjlig
- **Nackdelar**
 - komplex data struktur
 - topologi måste omregistreras efter varje uppdatering

Fördelaktigaste systemet för flertalet användare

Jämförelse mellan raster och vektor

Fördelar

- Enkel och läsbar lagring.
- Enkelt att analysera (algoritmer från fjärranalys och bildbehandling)
Enkla att kombinera (överläggning).

Raster

Nackdelar

- Kvalitet beror på pixel-storlek.
- Kräver mycket fysisk lagringskapacitet: grid formatet växer kvadratisk när cellstroleken minskar.

Fördelar

- Enkelt att skala om, kvalitet behålls vid transformationer.
- Enkelt med topologiska och nätverksberäkningar.
- Effektivt utnyttjande av fysisk lagringskapacitet.

Vektor

Nackdelar

- Beräkningsmässigt mer krävande för flera standardberäkningar (Filtrering, överläggning).

Skala och upplösning

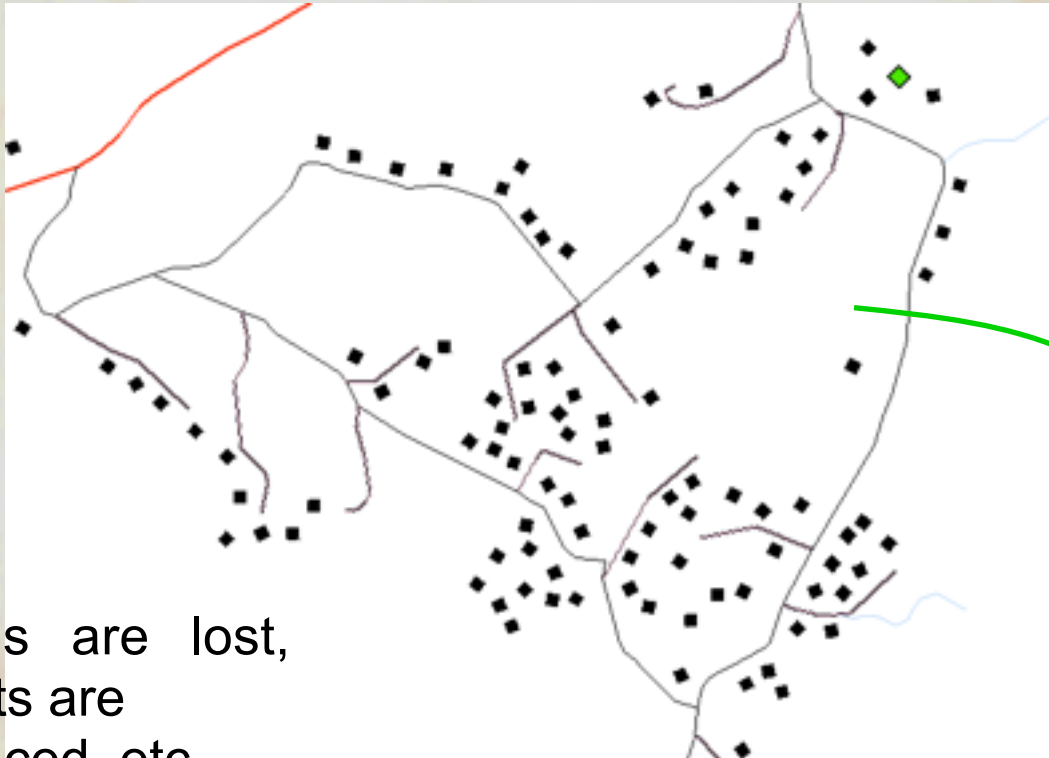
► Scale

- The ratio of distance on the map and distance on the ground
- E.g. 1:10 000, 1:50 000, etc.
- constant for paper maps / changeable in a GIS

► Resolution

- describes the level of detail at which the data was collected
- **spatial resolution** = pixel size in a raster surface
- **temporal resolution** = interval between 2 samplings
(ex. remote sensing: how often does a satellite fly over Stockholm and can take a picture of this area?)

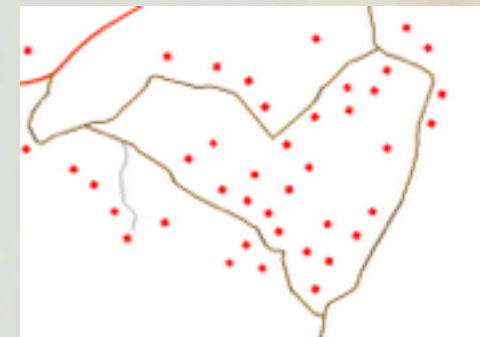
Geografisk generalisering



Details are lost,
objects are
displaced, etc.

How to find the optimal
selection of real objects
to visualise them on the
map of a certain scale?

Changing
between
scales =
**geographic
generalisation**



Reducing the level of detail in geographic data

Vilka faktorer styr generalisering

Scale – The scale determines what can be fit into the map

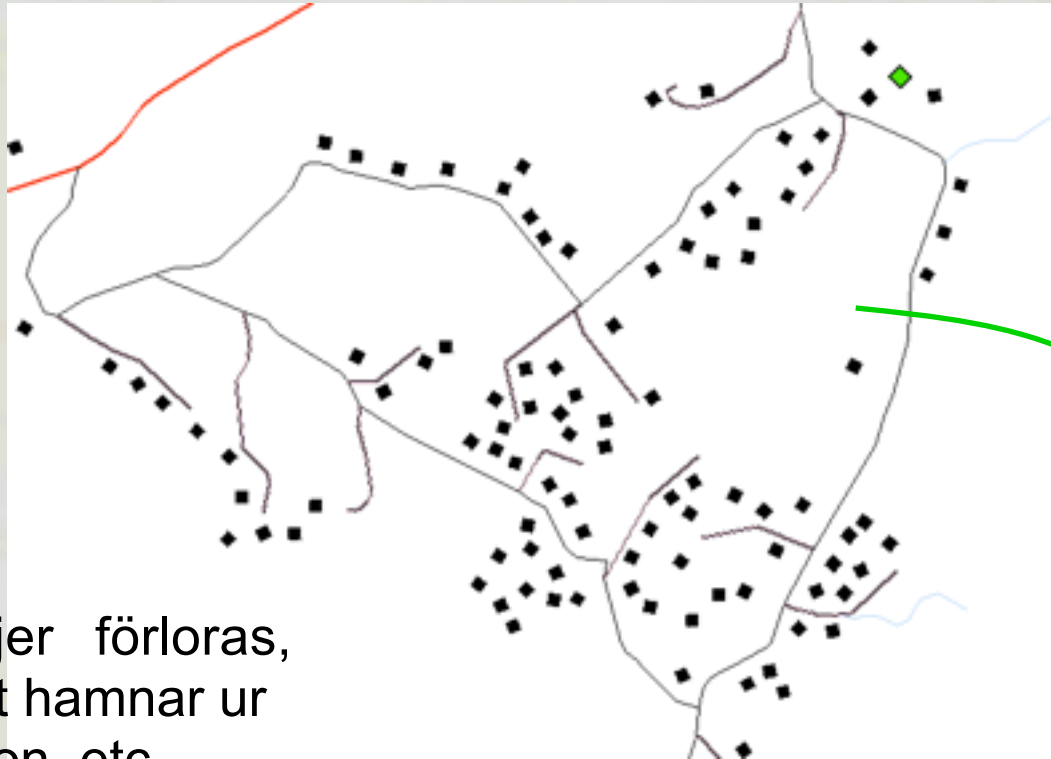
Map purpose – The purpose determines what is important to show.

Quality and quantity of available data

Graphical limits:

- choice of symbol specification
- technical reproduction capabilities

Generalisering i vektordata



Detaljer förloras,
objekt hamnar ur
position, etc.

Ökning av
skala =
**geografisk
generalisering**



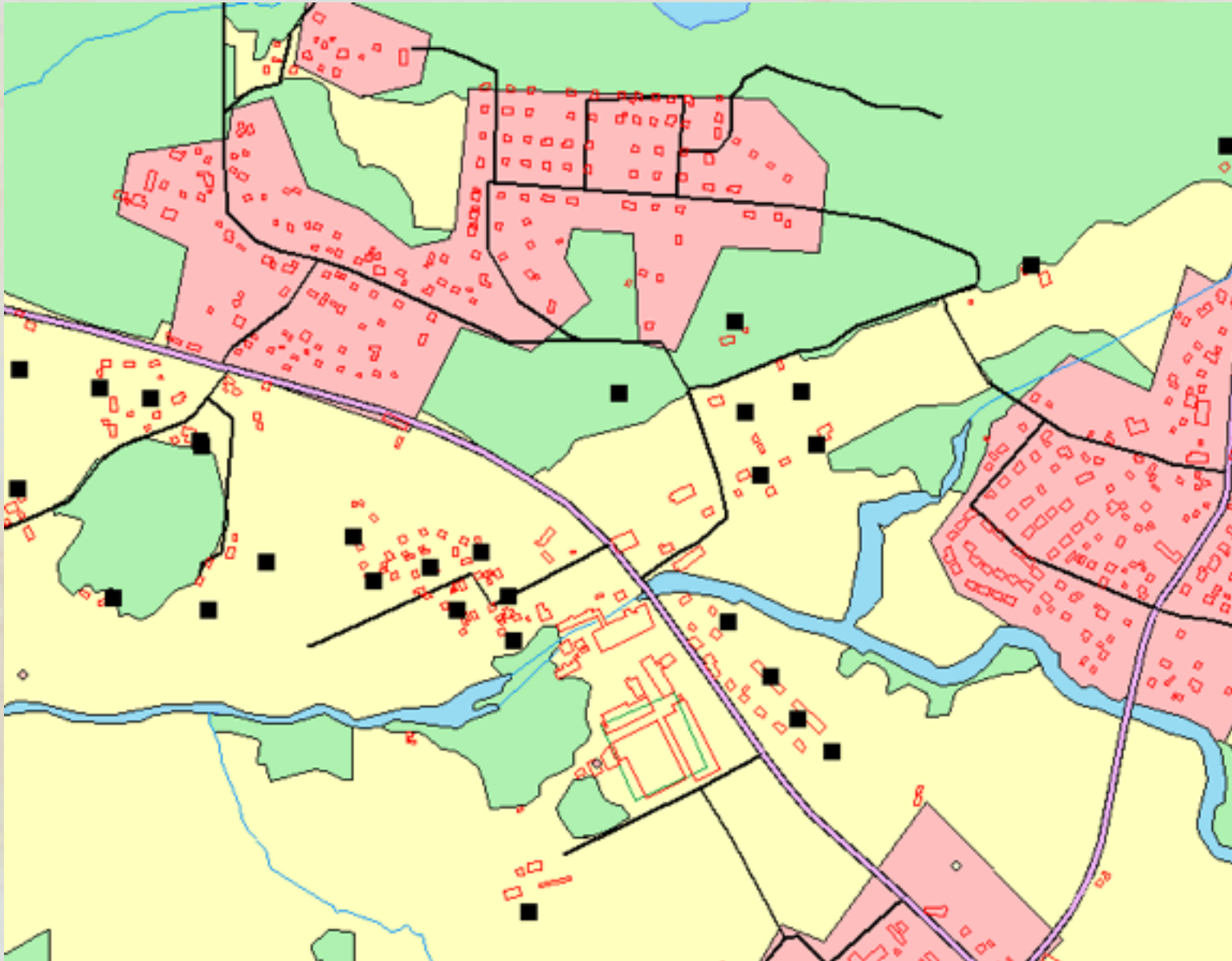
Reducerig av detaljnivå

Vektor data struktur - skala och generalisering

a city → an area on a large scale map
→ a point on a small scale map



Generalisering i vektordata



Datafångst för GIS

One of the most expensive GIS activities

Many diverse sources:

- terrestrial surveys,
- remote sensing data (satellite imagery),
- GPS measurements,
- digitising or scanning existing maps,
- socio-economic and statistical data,
- physical data,
- environmental data,
- etc.



Two capture methods

Primary (direct measurement)

Capture specifically for GIS use

Secondary (indirect derivation)

Recycling and reusing existing data

Datafangst tekniker

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

Raster Primary Data Capture



Main source – remote sensing:

- e.g. SPOT and IKONOS satellites, aerial photography, etc.
- passive and active sensors

Resolution is key consideration:

- spatial – pixel size
- spectral – wavelength
- temporal – revisit frequency



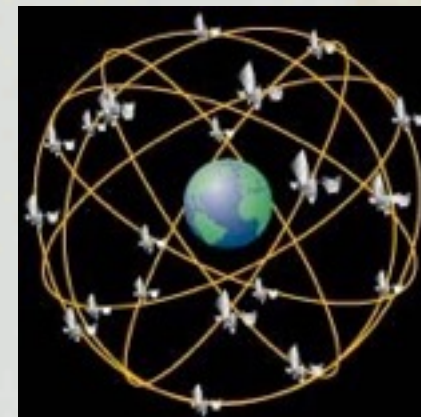
Vector Primary Data Capture

Survey measurements:

- locations of objects determined by angle and distance measurements from known locations
- uses expensive field equipment and crews
- most accurate method for large scale, small areas

Measurements by using GNSS (Global Navigation Satellite Systems):

- collections of satellites used to determine locations on Earth's surface:
 - GPS (USA) – mostly used
 - GLONASS (Russia)
 - Galileo (EU) – under development
- differential GPS used to improve accuracy



Raster Secondary 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 – no. of represented colours) resolution

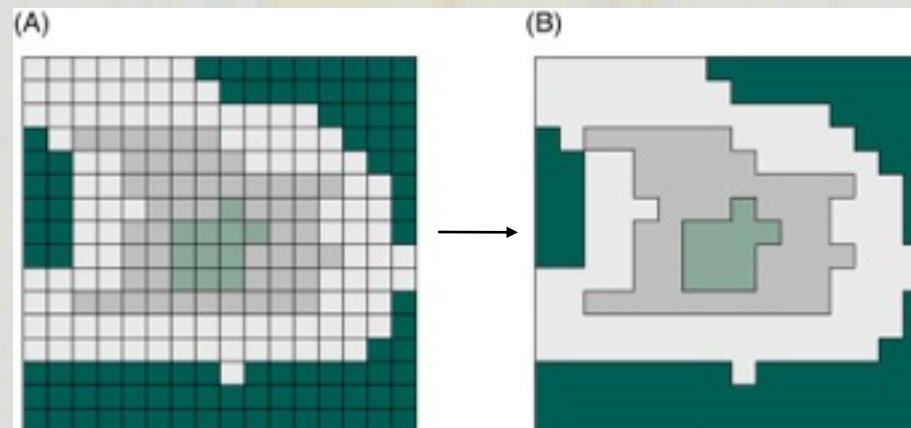


Vector Secondary Data Capture

Collecting **vector objects** from maps, photographs, plans, etc.

Digitising:

- Manual (digitising table)
- Vectorisation – converting rasters to vector data



Photogrammetry – the science and technology of making measurements from photographs.

Manual digitising:



Analog och digital kartografi

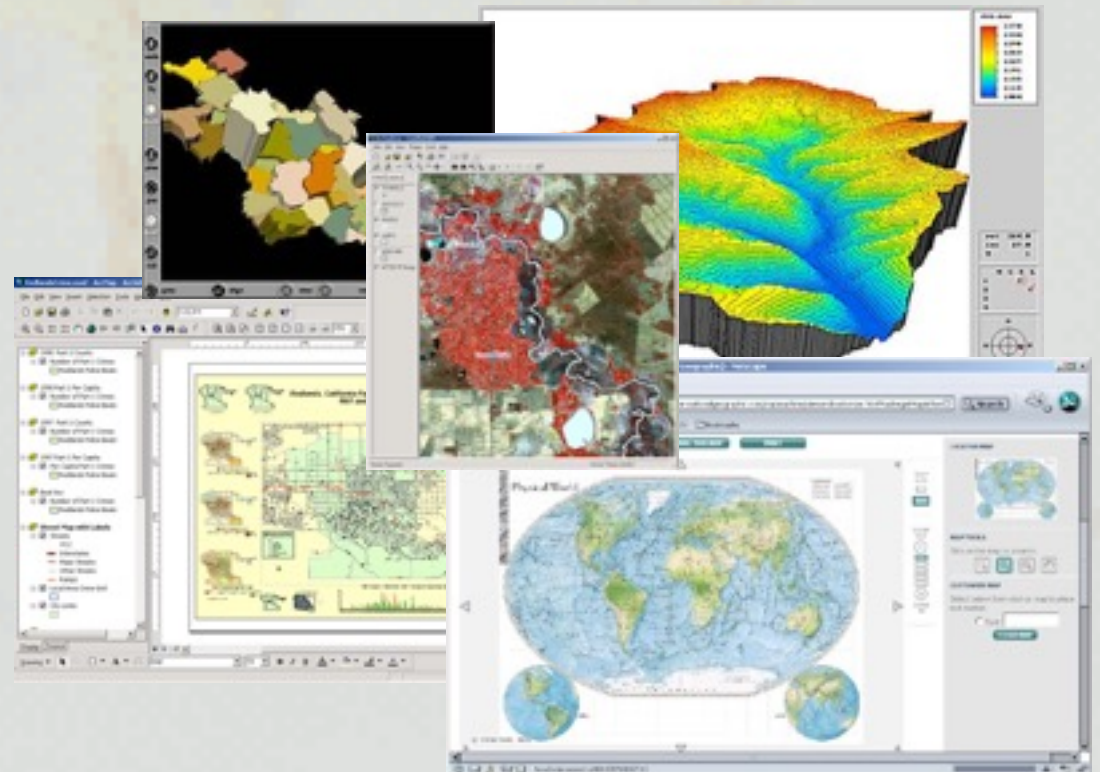
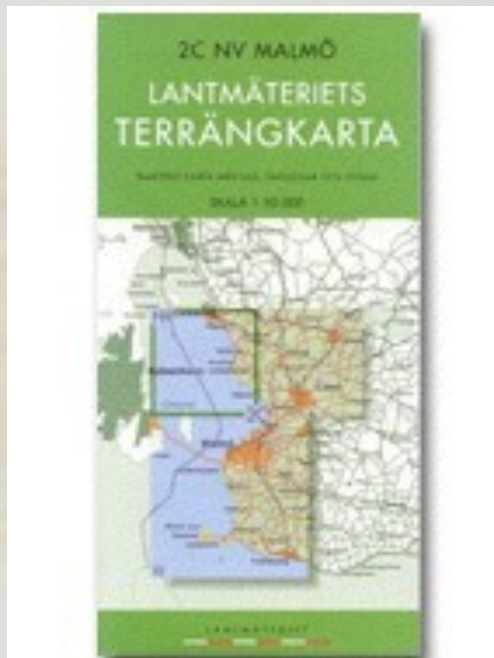
Traditionell papperskarta jämfört med digital geovisualisering

Analogt

En karta för alla ändamål, i ett format.

Digitalt

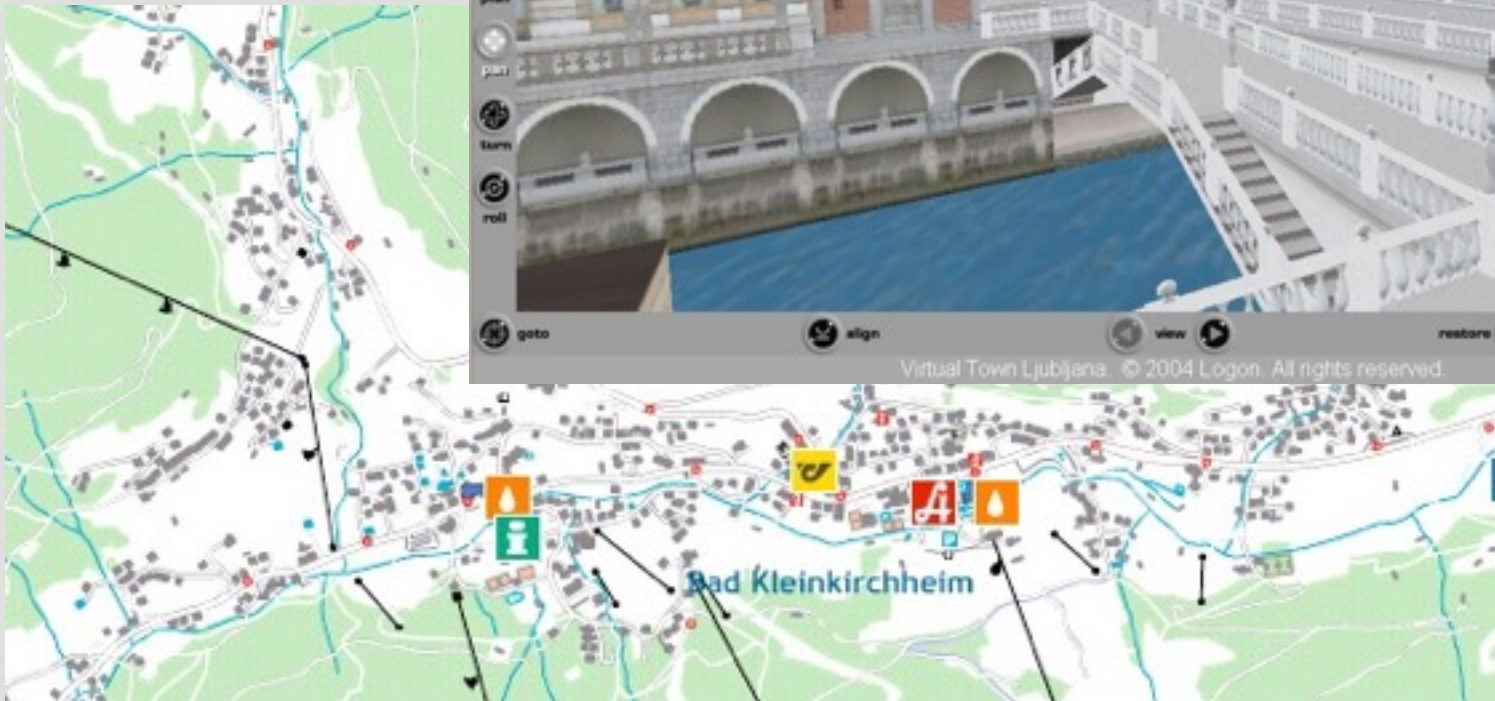
Användaren kan själv skapa sin egen 'karta'.



Analog och digital kartografi

Analogt – 2D

Digitalt – 3D



Analog och digital kartografi

Analogt - statiskt



Digitalt - mobilt



Analog och digital kartografi

Analogt – abstrakt

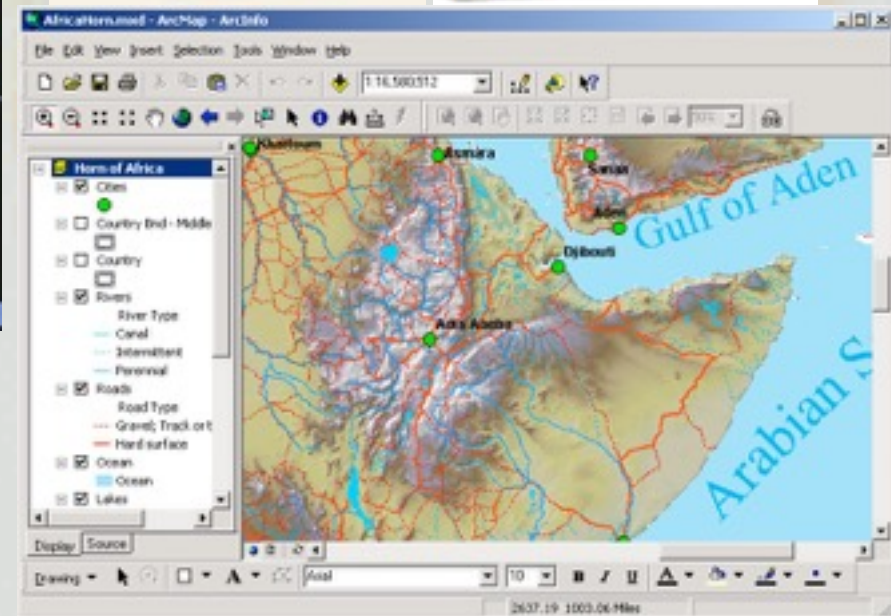
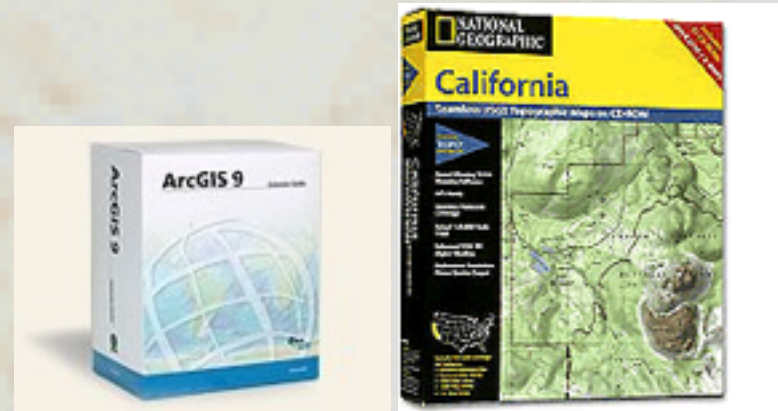
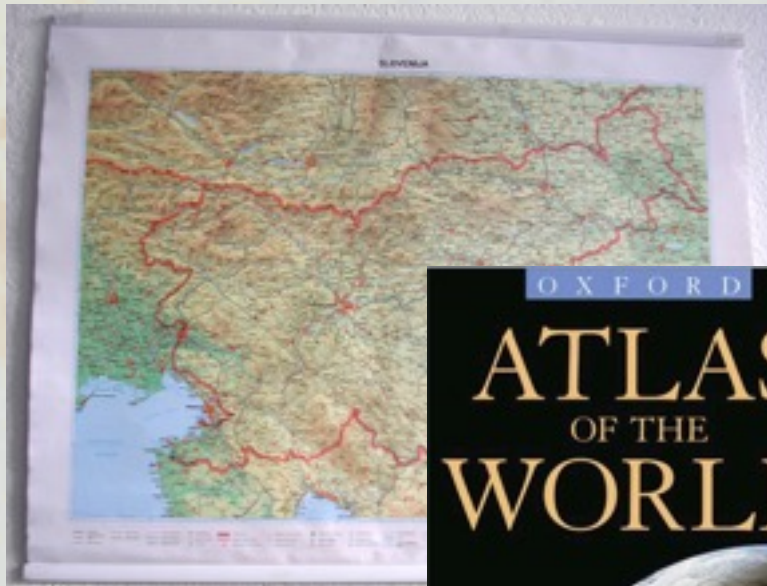
Digitalt – realistiskt



Analog och digital kartografi

Analogt – färdiggjort

Digitalt – egenproducerat



Analog och digital kartografi

Traditionell papperskarta jämfört med digital geovisualisering

Traditionell karta:

- förutbestämd skala
- fast område, närliggande områden på intilliggande blad
- statisk vy
- platt perspektiv
- data kan läggas till på enskilt blad med penna
- en enahanda vy av tema eller topografi

Digital geovisualisering:

- fritt val av skala (zoom)
- fritt val av område, panorering över stora områden (globalt)
- dynamisk visualisering (animering)
- 3D visualisering
- enkel uppdatering av attributdata
- många olika vyer skapade av olika användare