

Hva er digitale kart?

Gunnar Tenge, 27.9.05

Et digitalt kart har mange likheter med et papirkart. Et papirkart blir ofte beskrevet som ”et lite utvalg av virkeligheten som er kraftig forminsket, sett ovenfra.”

For å lage et papirkart må man

- gjøre et utvalg av objekter og fenomener som man vil ta med på kartet, både ut fra formålet med kartet (skal det bli et veikart, et o-kart, et arealressurskart eller et kart over sykdomstilfeller, osv.) og målestokken det skal forminskes til (ved liten målestokk tar man med et mindre utvalg av objekter og fenomener enn hvis man bruker en større målestokk).
- måle og kode de objektene og fenomenene man velger å ta med, f.eks. lite hus som liten sort firkant, vei som rød strek, sti som stipla sort strek, vann som blå flate, terrengoverflate som brune høydekurver, osv.
- tegne (lagre) på et papir

Et digitalt kart blir ofte definert på denne måten:

”En utvalg av romlige objekter og fenomener fra virkeligheten som er kodet og lagret i geografiske databaser”.

For å lage et digitalt kart må man

- gjøre et utvalg av objekter og fenomener, både ut fra formålet med det digitale kartet og målestokken det skal forminskes til/presenteres i (et digitalt kart lagres strengt tatt i målestokk 1:1 og det ser ikke ut som et kart før det blir presentert på skjerm eller papir. Det er den anbefalte målestokken ved presentasjonen som er interessant her)¹.
- måle og kode (se Dataklassifisering mot slutten av artikkelen) de objektene og fenomenene man velger å ta med, f.eks. lite hus som ordinalt målingsnivå og kodes som heltall 6315 (brønn). Vanlig vei som 7000, sti som 7414, vann som 3102, terrengoverflate som linjer med kode 2001 osv., hvilket målingsnivå tilhører dataene man registrerer/velger å ta med, hvilken datatype bør velges (heltall, float, char, etc.)
- digitalisere (ofte med å tegne med musa) og lagre i et valgt geodataformat i en geodatabase.

Det digitale kartet er nå fysisk bare magnetisk/ikke magnetisk, strøm/ikke strøm, 0/1 inne i datamaskinen. Ved hjelp av datastrukturer og kartvisningsprogrammer i datamaskinen og regler som implementert i disse programmene, kan man presentere disse digitale kartene som kartbilder på skjermen eller skrive det ut på en skriver.

Både papirkart og digitale kart representerer et ”snapshot” av virkeligheten – slik det så ut på det tidspunktet man gjorde kartleggingen. Man sier gjerne at kart er statiske. Det er mange som tror at når kartet er ”på data” så er det oppdatert til en hver tid – slik er det ikke. Likevel er en av fordelene med digitale kart at de er lettere å oppdatere enn papirkart. Man kaller det

¹ Presentasjonsmålestokk: Det lagres sjelden anbefalt presentasjonsmålestokk sammen med dataene. Det er vanligere å lagre en kode for geometrisk nøyaktighet til objektene som ligger i basen. F.eks. en kode som betyr +/- 2m som tilsvarer ca. 1:5000.

gjærne å ajourføre geodatabasen som kan være å slette ting eller legge til nye ting. I tillegg har digitale kart den store fordel at de kan manipuleres og analyseres med GIS-programvare. Dette er det mest spennende med GIS og er beskrevet i de to neste tekstene.

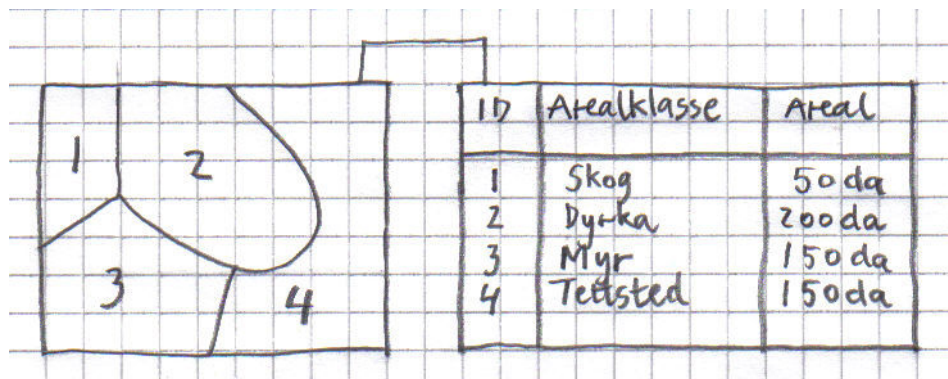
Kjært barn har mange navn og digital kart kalles like ofte geografiske data, geodata, geodatasett, stedfestet informasjon, romlige data eller spatial data på engelsk. Noen liker også å bruke ordet ”geografisk informasjon”, men for meg er dette noe mer – nemlig geodata som gjerne er koblet sammen med noen andre data og presentert på en slik måte at det fremstår som informasjon.

Geodata innholder informasjon om

- Hvor – sted, geometrisk beskrivelse av fenomenet, gjerne en god del koordinater
- Hva - egenskaper til stedene
- Når – ble observasjonen gjort og hvor lenge ”lever” den

Det særegne er altså dette med hvor/sted.

En vanlig måte å tegne et digitalt kart på er med georelasjonsmodellen som vist i dette eksempelet med noen polygoner som representerer forskjellige arealklasser.



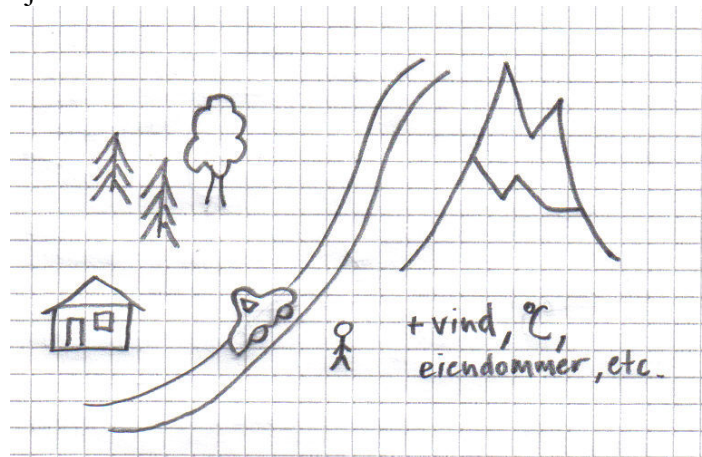
Geometri/steder

Egenskaper/beskrivelse av stedene

Fra virkeligheten til geodatabase

For å gå den lange veien fra virkeligheten til geodatabasen må det gjøres en del valg som til slutt gjør at man har geodata lagret på en strukturert måte inne i datamaskinen.

Virkeligheten består av objekter og fenomener. Her er det tegnet opp noen trær (en skog), et hus, en vei, en bil, et menneske, et fjell for å representere topografien, og det er ramset opp noen ting som vi ikke kan se men som er der likevel. Alt dette er eksempler på stedfestede fenomener og objekter.



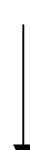
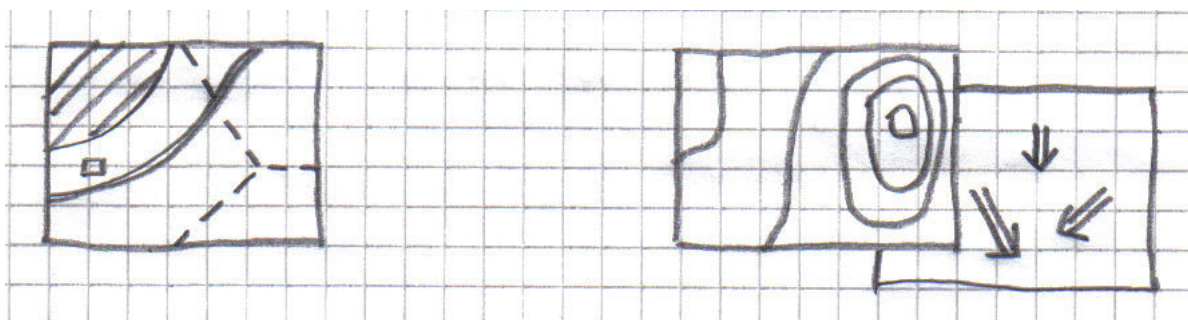
For det første så gjøres det gjerne et tøft bortvalg (det motsatte av utvalg!) av objekter og fenomener som man ikke vil ta med. Biler og mennesker tæs ofte ikke med.

Så kan virkeligheten deles i diskre objekter og kontinuerlige fenomener, object view og field view (Goodchild, 1992, Wang and Howarth, 1994).

I dette eksemplet blir veien, huset, eiendomsgrensen og skogen sett på som diskre objekter, mens topografien og vinden er tatt med for å illustrere noen kontinuerlige fenomener.

Object view

Field view



Diskré objekter

Ofte menneskeskapte objekter

f.eks. hus, veier, eiendomsgrenser, dyrkamarkgrenser, maurtuer

Finnes bare noen steder

Data blir samlet inn ved feltarbeid, landmåling, fly-/satellittbilde tolkning

Data samles inn ved å velge egenskapen (eg. hus) og så måle stedet (x,y)!

Objektene blir nesten alltid representert som punkt, linje og polygon i **vektorsmodellen – vektor GIS**.

God forankring i tradisjonell kartografi. Man er vant med å representere kontinuerlige fenomener som diskre objekter. Eg. høyder som høydekurver, bonitetsoverganger i skog som en klar avgrensning/strek, etc.

Kontinuerlige fenomener

Ofte ”naturskapte” fenomener

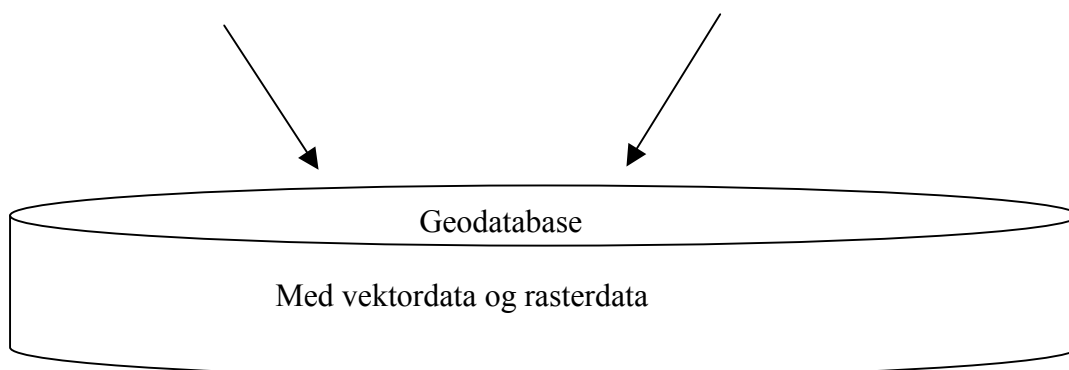
f.eks. overflater (høyder), temperatur, nedbør, vind

Finnes alle steder

Data blir samlet ved å sample i felt og fra bilder – kan ikke samle alle steder!

Data samles inn ved å velge stedet (x,y) og så måle egenskapen (eg. 150 moh)!

Disse stedene (helst alle steder) med verdier representeres ofte som pixler/celler med en pixelverdi/celleverdi i **rastermodellen – raster GIS**. Samplene kan også lagres som de er, og så kan man senere gjette på verdien i alle steder mellom samplingsstedene – interpolere, eller som TIN (triangulated irregular networks), etc



Representasjon av spatial relationships/romlige sammenhenger

Objekter og fenomener påvirker hverandre. Det er formulert i First law of Geography – Tobler 1970:

“Everything is related to everything else, but near things are more related than distant things”

Romlige sammenhenger håndteres på to måter i GIS.

- Enten ved å lagre dem i sammen med dataene i geodatabasen – i datastrukturen. Dette kalles **topologi** og betyr sammenhenger mellom objekter. Uttrykkes brukes mest i vektormodellen. Eg. for hver linje lagres det hvilke linjebiter denne består av med alle start- og endepunkter. Da har vi linje-topologi som vi er helt avhengig av for å få svar

på spørsmål av typen – finn korteste vei. For hver linje kan det også lagres hvilket polygon som ligger på hver side av linjen, og polygonene kan lagres som henvisninger/pekere til hvilke linjer som omslutter dem. Dette kaller jeg ordentlig polygontopologi (planar enforced datastrukture). Dagens mest kjente vektorformat. Shape-formatet og de vektorformatene som lagres i de store geodatabasene, lagres som såkalte Simple Features og har ikke har planar enforced datstruktur. Det er et klart savn – nok om det!

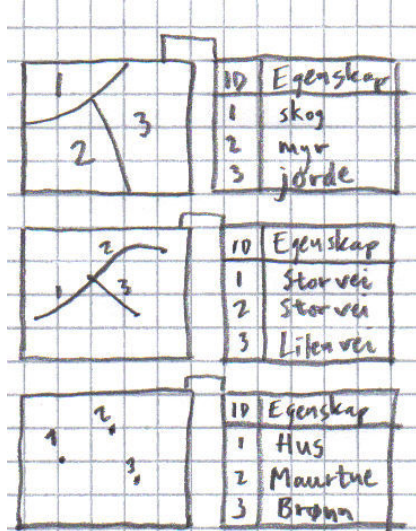
- b. Eller ved å finne dem når de trengs ved hjelp av kommandoer i GIS-programvaren. Eg. finn alle hus i nærheten av flyplassen. Definer nær til å være f.eks. 1000m og lag buffer rundt flyplass og select innenfor buffer...

Romlige sammenhenger er lettere/mulig å håndtere i GIS enn i manuelle kartsystemer. Manipulering og analyse av romlige sammenhenger er i grunnen noe av det mest spennende med GIS. Det er lett å håndtere f.eks. ”innenfor, utenfor og krysser” mens det er noe verre å håndtere ”nær, langt fra og nabo”.

Sammenlikne vektor og raster modellen

Her følger en sammenlikning av vektor og rastermodellen:

Vektormodellen



Polygon-, linje- og punkteksempel.

kompleks datamodell – bla. topologi

tung dataprosessering

tar liten plass

muligheter for topologisk prosessering eg. nettverksanalyser

vanskelig overlay

god kartografiske output

vanskeligere med romlige analyser

egenskaper i databasertabeller

Rastermodellen

100	90	80	50
100	70	50	30
90	50	20	0

Eksempel med høydeverdier lagret i hver celle. Rasteret består av rader og kolonner.

enkel datamodell – kun et todimensjonalt array

enkel dataprosessering

tar stor plass

enkel topologi, man kjenner bare nabocellen, men det er nok til å modellere f.eks. spredning

enkel overlay

dårlig kartografiske output

bra til romlige analyser

enkel egenskapshåndtering

Hvis man tenker seg en celle/et pixel som et lite kvadratisk polygon, så er i grunnen rastermodellen det samme som vektormodellen med små kvadratiske polygoner.

Representasjon av tid (temporal relationships)

Representasjon av tid er også lettere/mulig å håndtere i GIS enn i manuelle kartsystemer.

Inntil nå har dette stort sett vært løst ved å ta være på det gamle geodatadatasettet når det har kommet et nytt. Men interessen er økende og mulighetene større med moderne databasestrukturer. Det er lettere å modellere tidsdimensjonen.

Viktige egenskapene til en temporal prosess er når (generation time) den oppstår og hvor lenge (duration time) den varer. Operer også med noen som heter temporal målestokk. Eg. et

tre vokser i 100år, det skal vises på skjermen på 1 min, temporal målestokk = 1/52mill. Kan vises på skjermen som dynamiske kart, som små filmer – symboler som gror og blir så blir borte.

Dataklassifisering – måle og kode opplegg

Det er enda viktigere med en plan og et opplegg for dataklassifisering i GIS enn i manuelle kartsystemer. For det første ønsker vi å bruke geodataene til flere ting enn bare å lage et kart. Vi har flere formål med geodatene enn et kart. For det andre har er en datamaskin meget dårlig evne til å gjette på hva som egentlig var ment.

Det er altså viktig med en klassifiseringsplan som passer til det du ønsker å bruke geodataene til senere. Hvor nøyaktig skal inndelingen være? Hva slags kodesystem skal vi lage/bruke? Hvilket målingsnivå trenger vi? Hvilken datatype skal informasjonen lagres med?

Koder

I Norge har vi allerede et meget omfattende kodeverk for stedfestete objekter. SOSI (Samordnet Opplegg for Stedfestet Informasjon) inneholder en kodeliste med fire- og to sifrede koder for det nesten alt. Tipset er å se om det du ønsker å lagre i geodatabasen allerede har en offisiell kode.

Målingsnivåer

Det er viktig å gjøre seg opp en mening om hva man ønsker, og hva man kan gjøre, med informasjonen som blir samlet inn. Målingsnivåer/measurement scales som ble definert av Stevens i 1946 gir oss hjelp her.

Å kjenne målingsnivået er viktig fordi det

- styrer hva slags operasjoner og analyser som er tillatt å utføre på målingene
- gir føringer for hvordan målingene bør visualiseres (nesten 1:1 med visuelle variable i kartografien. Se teksten om kartografi)
- gjør det lettere å forstå betydningen av målingene

Nominal – Den enkleste. Bare klasse, navn eller id. Kan bare tilhøre en klasse.

eg. Arealklasseskog -dyrka, vann

Idnr – 213, 5, 13

Veinavn – Karl Johan, Drøbakveien

Operasjoner – eg. velge en, omkode, telle dem opp hver for seg eller alle sammen

Visuell variabel – de kvalitative variablene - retning, ulikt symbol, ulike farger

Ordinal – viser orden og rangeringer. Kan bare tilhøre en klasse.

Eg. Bystørrelse – liten, middels, stor

Egnethet for utbygging – lav, middels, høy

Operasjoner – som over + sortere dem

Visuell variabel – noen av de kvantitative med varsomhet fordi man ikke kjenner sprangene - størrelse, korning, gråtone/fargetone (lys til mørk), evt. trafikklys

Nominale og ordinale sier man er kategoriske

Intervall – kjente intervaller som kan ordnes og regnes på uten naturlig 0-punkt. Relativ fra et tilfeldig 0-punkt.

Eg. temperatur i C – 10 grader C, 12 grader C, 24 grader C

Operasjoner – som over + kan rangeres, kan subtraheres, men ikke multipliseres (40C er ikke dobbelt så varmt som 20!)

Visuell variabel - de kvantitative med varsomhet fordi man ikke kjenner

0-punktet - størrelse, korning, gråtone/fargetone (lys til mørk), evt. trafikklys

Ratio - kjente intervaller som kan ordnes og regnes på med et kjent 0-punkt.

Eg. temperatur i K- 100 kelvin, 200 kelvin

Høyde – 1m, 10m, 11m

Operasjoner – som over + kan rangeres, +, -, x, /. 20m er dobbelt så høyt om 10m!

Visuell variabel - de kvantitative. størrelsesproporsjonale symboler, korning, gråtone/fargetone (lys til mørk), evt. trafikklys

Her ga Stevens seg i 1946, men folk har kommet etter ham og ønsket noen presiseringer.

Cyclic ratio – som ratio men de gjentar seg

Eg grader i en sirkel – 0, 180, 360(som er 0), 45 etc.

Absolutt/Absolute – som ratio men hele skalaen er kjent, også sluttpunktet

Eg. sannsynlighet – [0 – 1]

Operasjoner – kan ikke ganges eller legges sammen

Telling/Count – et nøyaktig antall av et eller annet innenfor et område

Eg antall mennesker pr. kommune – 5213, 3457, 45356

Personene er diskrete objekter og kan ikke deles i to

Kan ikke regnes på så fritt som ratioer

Mens en befolkningstetthet $5213/20\text{km}^2 = 260$ pr, km^2 er en ratio

Under er det eksempel på en på hvor galt det kan gå når man legger sammen innholdet i informasjonsvariable med forskjellig målingsnivå.

<i>NEW CUYAMA</i>	
<i>Population</i>	<i>562</i>
<i>Ft above sea level</i>	<i>2150</i>
<i>Established</i>	<i>1951</i>

<i>Total</i>	<i>4663</i>

Dat typer

Når data skal lagres i databaser må man også ta stilling til hva slags datatype de skal lagres i. Som heltall/integer, flyttall/float, char, blob (binary large object), desimaltall med x antall desimaler, etc.

Geodatabaser

Geodata bør lagres i geodatabaser. Databaserteori og er et stort felt som jeg kun vil berøre her.

Det viktigste nå er å vite at geodata før ble lagret som filer og at det jobbes det veldig hardt med å legge dette i inn geodatabaser – strukturerte lagringsplasser for geodata. Geodatabaser er den viktigste byggestenen i en moderne geografiske infrastrukturer. Det er først de senere år Software Engineering, en veldefinert metode for bla. bygging av robuste databasesystemer, har blitt tatt i bruk for bygging av geodatabaser. Det er også vanlig å si at det først nå, ved fokus på geodatabaser, at GIS-miljøet har blitt en del av mainstream IT.

Noen viktige ord og uttrykk:

En database kan beskrives på

- konseptuelt nivå – en lettlest ”skisse” med de objekter som skal inngå i databasen
- logisk nivå – en mer strukturert beskrivelse av de objektene som skal inngå i databasen
- fysisk nivå – en detaljert beskrivelse av hvordan objektene skal lagres, fysisk, i databasen

Relasjonsdatabaser – en godt definert databasemodell som baserer seg på tabeller som kan kobles sammen vha. koblingsnøkler.

Georelasjonsmodellen – en beskrivelse av geodata på et logisk nivå. Viser geometri og egenskapstabeller koblet sammen. I denne modellen samlet objekter som passer sammen, i lag. Eg. et lag for veier, et lag for markslag, et lag for høydekurver, etc.

Objektorienterte databaser – mer et abstrakt konsept enn en formalisert datamodell med definerte regler og prosedyrer. I stedet for å fokusere på tabeller som i relasjonsdatabasemodellen eller som lag som i georelasjonsmodellen, fokuseres det på objekter – eg. hus, skog, menneske, vei, hendelse. Så ser man på ”hva er et hus?”, ”hva ønsker man å gjøre med et hus?”. Objektorientert modellering starter på det konseptuelle nivået. Så beskrives det hvilken datastruktur man skal lagre ”hus” i og hvordan man skal implementere operasjonene som skal virke på ”hus”. På denne måten beveger man seg nedover via logisk nivå til fysisk nivå i databasebyggingen.

Til slutt et forsøk på å klassifisere digitale kart fra enkle til veldig gode

Dette er noe jeg brukte i forelesninger for 10-15 år siden. Det taes med for å fullføre artikkelen om digitale kart.

- a. et digitalt bilde av et kart på f.eks. jpg eller tif format, uten georeferering
- b. et digitalt bilde av et kart med georeferering
- c. et digitalt ortofoto eller et satellittbilde
- d. vektordata uten topologi uten georeferering, ”veldig dårlig spagetti”
- e. vektordata uten topologi med georeferering, ”spagetti”
- f. et utvalg av målinger av et kontinuerlig fenomen, samples, med georeferering

- g. et egenskapskodet rasterkart med georeferering
- h. vektordata med enkel topologi, simpel features, med georef
- i. vektordata med linje og polygontopologi, planar enforced datastructure, med georeferering