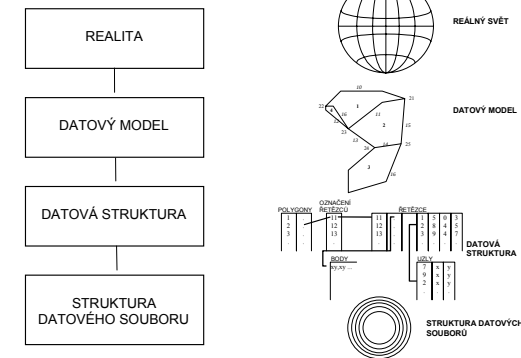


Data GIS

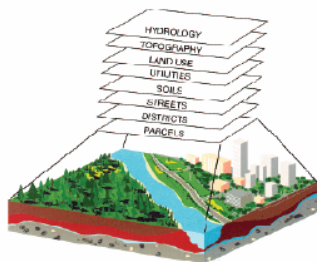
- Datové modely v GIS
- Tvorba dat pro GIS

Přednáška 2.

Úrovně abstrakce reality



Model reálného světa



- Reálný svět
- Vrstvy
- Objekty

Poloha v prostoru

Geometrie

- Pojem geometrie má mnoho významů („měření na Zemi“, odvětví matematiky...)
- Geometrie v GIS: geometrická část popisu objektu

Topologie

- obor matematiky
- zabývá se popisem a analýzou prostorových vztahů mezi geometrickými objekty
 - v reálném světě intuitivně chápány
 - pro GIS musí být vyjádřeny počítači srozumitelným způsobem

Poloha v prostoru

- Aby bylo možné definovat geoobjekty v prostoru (tj. definovat jejich geometrii a topologii jednoznačným způsobem), potřebujeme **souřadnicový systém** (spatial reference system)

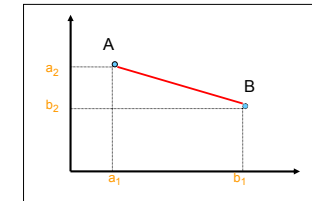
Musí splňovat tři podmínky:

1. Definice polohy musí být jednoznačná
2. Definování polohy musí být kvantifikovatelné
3. Musí být definována **metrika**

Euklidovská metrika

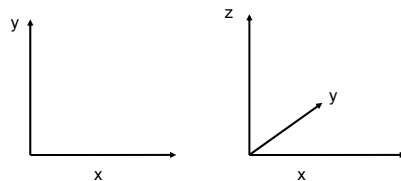
- nejběžnější způsob definování metriky v n-dimenzionálních prostorech

$$d(A, B) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad A(a_1, \dots, a_n) \text{ a } B(b_1, \dots, b_n)$$



Karteziánský souřadnicový systém

- Nejdůležitější systém pro aplikace v geovědách
- N-rozměrný karteziánský systém se vytváří párováním jednorozměrných prostorů reálných čísel R (souřadnicových os, které se protínají ve svých nulových polohách (v počátku souřadnicového systému) a mají definované nějaké měrné jednotky
- V geovědách hlavně 2 a 3rozměrné souř. Systémy – osy x, y, z



Datové modely

- Data v GIS jsou modelem reálného světa
- Objekty reálného světa jsou generalizovány a formálně reprezentovány způsobem, který je možno uložit a zpracovat v počítači

VEKTOR

Model je vytvořen z čar, vznikajících spojením bodů, přes které přecházejí. Body jsou reprezentovány svými souřadnicemi. Popisné informace mohou být uloženy ve zvláštních záznamech, podrobněji charakterizujících objekt

x

RASTR

Prostorově na sebe navazující množina buněk, které kompletně vyplňují celou zkoumanou plochu. Každá buňka má hodnotu, která vyjadřuje hodnotu popisné informace v této poloze

Vektorová reprezentace prostorových objektů

- název odvozen z matematické terminologie – přímá čára s danou velikostí a směrem = **vektor**
- jednotlivé základní **geometrické** prvky jsou vyjádřeny sestavou bodů, případně bodů spojených úsečkami
- **Bod** - vyjádřen diskrétní polohou určenou souřadnicemi x,y
- **Linie** - vyjádřena posloupností orientovaných úseček definovaných souřadnicemi jejich počátečního a koncového bodu
- **Plocha** - vyjádřena pomocí uzavřené obrysové linie

Topologie ve vektorové reprezentaci

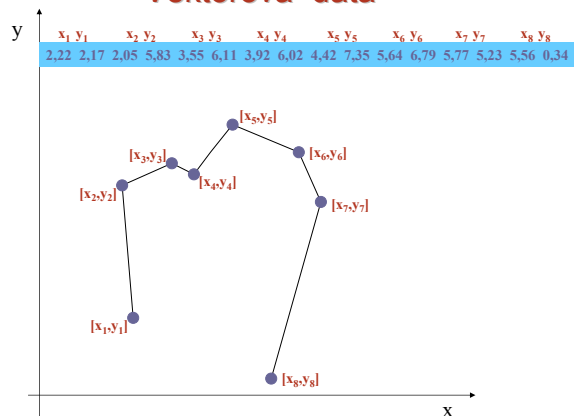
v **topologickém** smyslu vektorový objekt tvoří:

- **Uzly** – bezrozměrný objekt
- **Hrany**
 - hrana je propojením dvou uzlů
 - má-li význam **pořadí uzlů** při definování hran, má **směr**
- Základní prostorové vztahy mezi těmito elementy jsou formulovány v teorii grafů

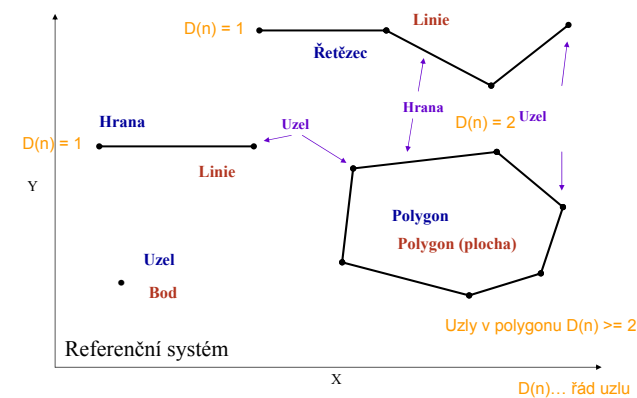
Základní topologické vztahy jsou:

- **Konektivita** – spojitost (connectivity)
- **Orientace** – směr (orientation)
- **Přilehlost** – sousednost (adjacency)
- **Obsahování** - (containment)

Vektorová data



Geometrické a topologické chápání objektů ve vektorové reprezentaci

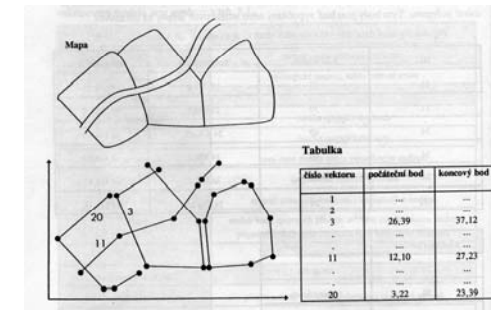


Ukládání vektorových dat

- Databáze souřadnic různé struktury
- Většina použitelných vektorových datových struktur může být zařazena do jednoho z následujících "klasických typů":

- špagetový model
- topologický model
- hierarchický model

Špagetový datový model



Špagetový datový model

- Soubor bodů a vektorů v podobě souřadnic x,y, **bez informace o vzájemných vazbách**
 - o při vykreslení vytvářejí změť protínajících se čar - podobné "špagetám"
 - o nejsou vyznačeny průsečíky čar ani údaje o logické provázanosti objektů
 - o polygony - nezávislé řady souřadnic, společná hranice mezi dvěma polygony je v databázi uložena dvakrát
 - o **není vhodný pro prostorové analýzy** - jakákoliv informace o vztazích mezi objekty, jež je zřejmá při pohledu na původní analogový dokument, se musí odvozovat výpočtem
 - o ukládání a prohlédávání dat - sekvenční - trvá dlouho
 - o používá se především v počítačové grafice, digitální kartografii

Topologický datový model

Topologický model

- Vyjadřuje spojení a vazby mezi objekty nezávisle na jejich souřadnicích
- Vyžaduje, aby všechny čáry byly spojité a polygony uzavřené
- V GIS je nejvíce používán pro ukládání prostorových dat
- **Vhodný pro analytické úlohy**

Topologický datový model - NAA

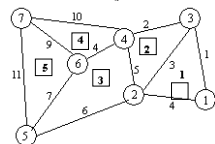
- NAA (Node-Arc-Area)** - uzel - hrana - plocha: obecně užívaný topologický model

Složky:

zorientovaný oblouk - hrana - arc
průsečík - uzel - node
plocha - polygon - area

Pravidla:

- Každá hrana má jeden počáteční a jeden koncový uzel
- Každý uzel musí být začátkem nebo koncem nejméně jedné hrany
- Každá plocha je ohraničena jednou nebo více hranami
- Hrany se mohou protínat jen ve svých uzlových bodech
- Každá hrana má přesně jednu plochu nalevo a napravo
- Každá plocha musí být pravou nebo levou plochou nejméně jedné hrany



Lze snadno popsat v relačním databázovém modelu

Rastrová reprezentace prostorových objektů

- Prekrytí studované oblasti pravidelnou sítí (mřížkou)
- Popisné hodnoty jsou vztaženy k polím této sítě
- Polohová lokalizace entity je určena souřadnicemi polí, které ji představují.

Základní jednotky

- buňky rastru (cells)** nebo **pixels** (pixel - picture element).
- 3D: **voxel**

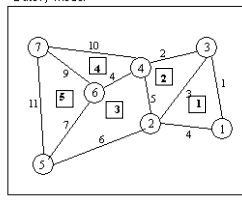
Dělení prostoru:

- pravidelná (regular)** - tvar buněk je přesně definován a vždy stejný - čtverec, šestiúhelník, trojúhelník.
 - rastry se stejnou rozlišovací úrovní - jednotlivé buňky jsou stejně velké
 - rastry s nestejnou rozlišovací úrovní nebo hierarchické - velikost buněk se definovaným způsobem mění
- nepravidelná (irregular)** vytvářejí se buňky různého tvaru i velikosti

Topologický datový model

topologický datový model a struktura

Datový model



(Zpracováno podle Peugnot a Marlin, 1990)

Soubor topologických vztahů

hrana	pravý polygon	levý polygon	počátek v bodě	konec v bodě
1	1	0	3	1
2	2	0	4	3
3	2	1	3	2
4	1	0	1	2
5	3	2	4	2
atd.				

Soubor souřadnic bodů

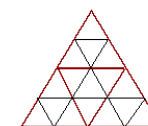
uzel	x-ová souřadnice	y-ová souřadnice
1	23.5	8.8
2	17.6	17.1
3	29.1	15.4
4	26.3	20.5
atd.		

Rastrová reprezentace prostorových objektů

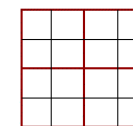
Nejpoužívanější je čtvercová mřížka – 3 hlavní důvody:

- Kompatibilní se strukturami datových posloupností používaných ve výpočetní technice (matice)
- Kompatibilní s mnoha hardwarovými zařízeními pro záznam a výstup údajů (skenery, tiskárny, plotry)
- Kompatibilní s karteziánskými souřadnicovými systémy

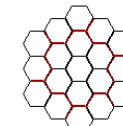
Pouze buňky tvaru čtverce lze rozdělit na menší čtverce se stejným tvarem a orientací.



trojúhelníková



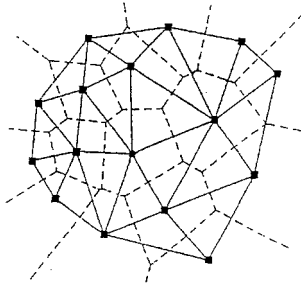
čtvercová



šestiúhelníková (hexagonální)

Triangulace

Delaunay (TIN) x Thiessen (Voronoi)

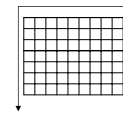


1	2	2	2	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	1	1	1	1	1	1	4	4	2	2	2	2	2	2	2	2	2	2	2
3	1	1	1	1	1	1	1	4	4	2	2	2	2	2	2	2	2	2	2
3	3	1	1	1	1	1	1	4	4	4	4	4	4	5	5	5	5	5	5
3	3	1	1	1	1	1	1	4	4	4	4	5	5	5	5	5	5	5	5
3	3	1	1	1	1	1	1	4	4	4	5	5	5	5	5	5	5	5	5
3	3	1	1	1	1	1	1	4	4	4	5	5	5	5	5	5	5	5	5
3	3	3	3	1	1	1	1	4	4	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	4	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5
3	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	5

Geometrie v rastrové reprezentaci

Rastr lze geometricky definovat:

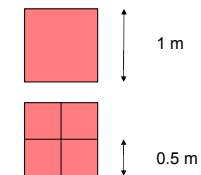
- stanovením směru os X a Y
 - sloupce buněk podle osy X
 - řádky podle osy Y.
- stanovením velikosti krokové vzdálenosti δX a δY
 - velikost strany buňky, rozlišovací schopnost (rozlišení) rastru
- velikosti (rozsáhlosti) rastru – dána rozdílem hodnot nejnižších a nejvyšších souřadnic (X_1, Y_1) a (X_n, Y_n) buněk rastru. Používají se hodnoty souřadnic spodních levých nebo horních levých rohů buněk.



Lokalizace rastru

- Definice každé buňky dvojicí souřadnic georeferenčního systému
- **Diskrétní rastrový prostor:** k definici polohy buněk se použije dvojice indexů (i, j) – definice polohy vzhledem k počátku rastrového systému
 - i = index sloupců buněk rastru
 - j = index řádků buněk rastru

Rozlišení rastru

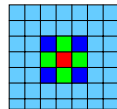


- Velikost souboru > 4x
- Vyšší rozlišení = vyšší přesnost

880 dpi

Topologie v rastrové reprezentaci

- implicitně definována v geometrii rastru
- každý element má **dva druhy sousedů**:
 - **Plné – dokonalé**
 - sousední elementy ve stejných řádcích a sloupcích
 - **Diagonální**
 - sousední buňky na rozích zkoumaného elementu
- Zvlášť – **čtyřčlenné sousedské okolí**
- Dohromady – **osmičlenné sousedské okolí**



Tyto topologické vazby jsou základem pro vykonávání různých operací a prohledávání rastru z různých hledisek.

Výhody a nevýhody vektorů

Výhody:

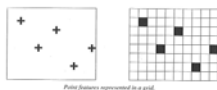
- je možné pracovat s jednotlivými objekty jako se samostatnými celky
- menší náročnost na paměť
- dobrá reprezentace jevové struktury dat
- kvalitní grafika, přesné kreslení, znázornění blízké mapám
- vysoká geometrická přesnost
- jednoduché vyhledávání, úpravy a generalizace objektů a jejich atributů

Nevýhody:

- složitější odpovědi na polohové dotazy
- obtížná tvorba překryvů vektorových vrstev
- komplikovanost datové struktury
- problémy při náročných analytických operacích, složitost výpočtů
- problémy při tvorbě modelů, komplikované využití pro simulaci jevů

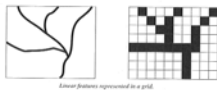
Tématika v rastrové reprezentaci:

- téměř ve všech praktických aplikacích se používá vrstevový přístup
- Jednoduchost geometrie a topologie rastrového modelu je v kontrastu s přesností reprezentace polohy a tvaru geoobjektů:

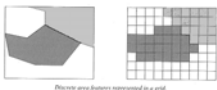


→ přesnost je limitována velikostí buněk

Bodové objekty - reprezentovány individuálními buňkami



Liniové objekty - aproximovány sekvencí sousedních buněk



Plošné objekty - aproximovány množinou souvisejících, sousedících buněk

Výhody a nevýhody rastrů

Výhody:

- rychlé dotazování
- snadné překrývání a kombinace obrazů s různým obsahem
- jednoduchost datové struktury
- jednoduchá kombinace s jinými daty rastrové povahy (DPZ)
- snadné provádění analytických operací

Nevýhody:

- značná paměťová náročnost (velký objem dat)
- omezená přesnost, daná rozlišením rastru a orientací rastru (výpočty délek, vzdáleností, ploch ...)
- kvalita výstupů závislá na rozlišení rastru
- nižší vizuální kvalita rastrových výstupů
- nevhodnost pro síťové analýzy



Specifika prostorových objektů

Dimenze

- Prostorové objekty mohou být modelovány v různých **dimenzích** (**dimenze** geoobjektu **charakterizuje jeho rozšíření do různých směrů prostoru**)
- Pro potřeby geometrického modelování můžeme uvažovat s objekty těchto dimenzí:
 - bezrozměrné - 0D – body
 - jednorozměrné - 1D - (přímé) úseky čar
 - dvourozměrné - 2D - polygony
 - trojrozměrné - 3D - tělesa, u nichž lze určovat objem
- 4. dimenze - čas
- "tématická dimenze"** - počet definovaných atributů.

Prezentace prostorových dat - shrnutí

Objekt	Vektorový formát		Rastrový formát	
	Digitální	Analogová	Digitální	Analogová
Bod	souřadnice x,y	•	Pixl	
Linie	posloupnost souř. x,y		Pixly	
Plochy	uzavřená posloupnost souř. x,y		Pixly	

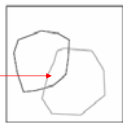
Specifika prostorových objektů

Měřítko

- Geoobjekt může být modelovaný v různých prostorových **měřítkách**, tj. s různou **prostorovou rozlišitelností**.
- Prostorová rozlišitelnost je v úzkém vztahu ke kartografickému měřítku → větší rozlišitelnost = větší měřítko

Shapefile

- Nejjednodušší
- Neukládají topologické vztahy
- Vytváření map a některé analýzy
- obsahuje prvky jen jednoho geometrického typu
 - Body a vícenásobné body (point a multipoint)
 - Linie (line a polyline)
 - Plochy (polygony)
- !!! Shapefile se skládá z několika souborů stejného jména !!!



Coverage

- Tradiční formát pro komplexní zpracování geodat (*geoprocessing*)
- Ukládá topologii = vhodné pro analýzy

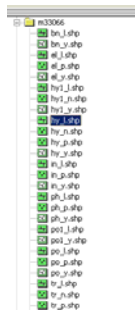
Prvky coverage

- Primární (Primary features)
- Sekundární (Secondary features)
- Složené (Composite features)

Systém souborů a ArcCatalog

	Průběh	Velikost	Datum
1. smk	shp	23 884	27.01.2004 12
2. smk	dbf	240 753	22.01.2004 12
3. smk	shx	24 468	22.01.2004 12
4. smk	shp	2 692	22.01.2004 12
5. smk	shp	1 121 368	22.01.2004 12
6. smk	shp	13 316	22.01.2004 12
7. smk	shp	8 309	13.12.2003 18
8. smk	shp	54 548	13.12.2003 15
9. smk	shp	128 169	13.12.2003 15
10. smk	shp	54 368	13.12.2003 15
11. smk	shp	5 015	13.12.2003 15
12. smk	shp	8 240	13.12.2003 15

Shapefile v systému souborů



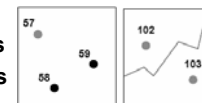
Shapefile v ArcCatalog

- Poznámé geometrický typ
- Shapefile vidíme jako samostatnou třídu prvků

Primární prvky coverage

Body

- Points
- Labels

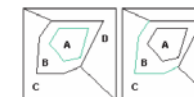


Linie

- Atribut mohou mít i nody



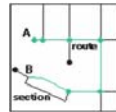
Polygony



Složené prvky coverage

Trasy (routes) a Sekce (sections)

- Liniové prvky
- Složeny z linií a jejich částí
- Trasy definují cesty liniovou sítí
- Sekce
 - části trasy
 - Hrana (arc) nebo její část → nemusí končit nodem



Regiony (regions)

- Polygonové prvky
- Region je tvořen více polygony; nemusí být souvislý

Coverage

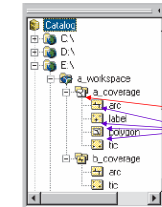
- Složka (folder) s coverage obsahuje několik souborů s informací o souřadnicích, topologii atd.

INFO

- Spravuje atributové tabulky, propojené datové tabulky aj.
- NEZASAHOVAT DO COVERAGE MIMO ARCGIS



Pracovní oblast v systému souborů



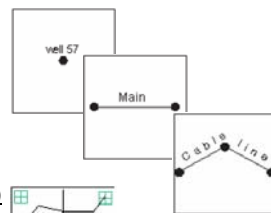
Pracovní oblast v ArcCatalog

Poznáme geometrický typ coverage i jednotlivých tříd prvků

Sekundární prvky coverage

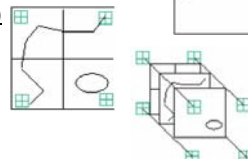
Anotace (annotation)

- Textové řetězce; popisují prvek v mapě
- Jsou umístěny k bodu, mezi dva body nebo podél řady bodů



Registrační (vlicovací) body (tics)

- Reprezentují místa o známých souřadnicích
- Slouží k transformaci souřadnic



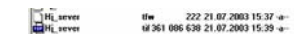
Uložení rastrů

- Obecné grafické formáty se zvláštním souborem pro georeferenci

- Tif – Tfw
- Bmp – Bpw
- Jpg – Jpw

- Některé sw mají své vlastní rastrové formáty – georeference je v nich již obsažena

- *.ras v TopoLu
- *.img Erdas Imagine
- ArcInfo grid – jedna vrstva je v jednom adresáři, nutný adresář info



Imeno	Typ	Velikost	Datum	Atribut
[...]	<DIR>	29.09.2004 14:35		
[...]	<DIR>	22.09.2004 16:39		
[...]	<DIR>	29.09.2004 14:50		
[...]	<DIR>	29.09.2004 14:12		
[...]	<DIR>	29.09.2004 14:35		
[...]	<DIR>	29.09.2004 14:04		
[...]	<DIR>	29.09.2004 14:50		
[...]	<DIR>	22.09.2004 15:30		
[...]	<DIR>	22.09.2004 15:14		

Imeno	Typ	Velikost	Datum	Atribut
[...]	<DIR>	22.09.2004 16:39		
[...]	<DIR>	32.22.09.2004 16:39		
[...]	<DIR>	308.22.09.2004 16:39		
[...]	<DIR>	516.22.09.2004 16:39		
[...]	<DIR>	32.22.09.2004 16:39		
[...]	<DIR>	2.040.22.09.2004 16:39		
[...]	<DIR>	23.006.000.22.09.2004 16:39		
[...]	<DIR>	553.500.22.09.2004 16:39		

Jak data pro GIS vytvořit

Pořizování dat

Podle způsobu vzniku:

1. **Primární data** - vznikají přímým měřením v terénu (geodetická, meteorologická měření, mapování v terénu...)
2. **Sekundární data** - získávají se z materiálů již dříve pořízených (mapy, statistické ročenky.....)

Podle formy:

1. Data v analogové formě
2. Data v digitální formě

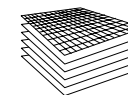
Vznik rastrových dat

- Přímou v digitální podobě
– družicové snímky
- **Skenováním** z analogové podoby:
 - Mapy
 - Letecké snímky na fotografickém materiálu



Skenování

- Postup a volba technického vybavení závisí na podkladu a požadovaném výsledku:
 - Formát (A4, A1....)
 - Požadavky na přesnost a rozlišení...
 - Tematický nebo kontinuální rastr



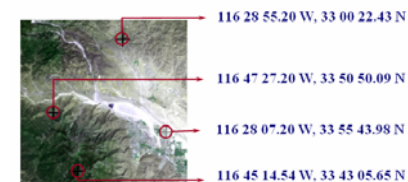
Obrázek

- Všechny tyto obrázky mají po pouhém naskenování z hlediska práce v GIS obdobný přínos...



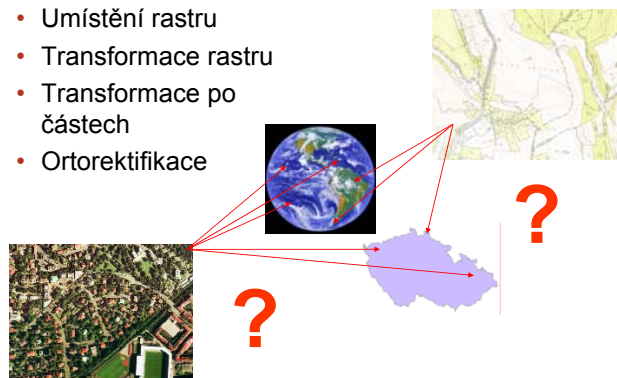
Jak?

- Musíme mít body o známých souřadnicích a vědět, který bod neumístěného podkladu představují
 - Geodetická měření
 - GPS



Kam s ním?

- Umístění rastru
- Transformace rastru
- Transformace po částech
- Ortorektifikace



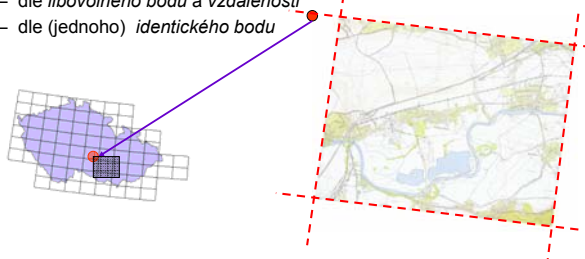
Identické body

- Známé souřadnice získáme pomocí existující vrstvy GIS, změříme pomocí GPS, „totálky“, teodolitu...



Umístění rastru

- Jednoduchá specifikace polohy (posun)
- Pouze pro případy, kdy není třeba korigovat geometrické zkreslení rastru
- Různé způsoby, např.:
 - dle *Pravého horního rohu* a *šířky* celého souboru
 - dle *libovolného bodu* a *vzdálenosti*
 - dle (jednoho) *identického bodu*



Transformace, ortorektifikace

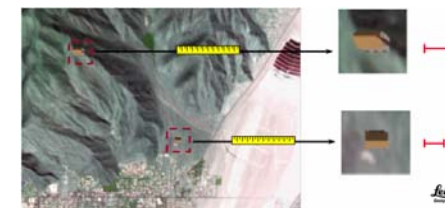
- Složitější způsob umístění rastru, kdy jsou korigována geometrická zkreslení



Zkreslení:
Do roviny je snímán
nerovinný útvar
(tvar planety, reliéf)

Zkreslení

- Různé měřítko v různých částech obrazu
- Výška a náklon senzoru
- Interní chyba senzoru



Metody transformace

PODOBNOSTNÍ

- Posunutí, otočení a změna měřítka v obou směrech
- Minimálně dva páry identických bodů

AFINNÍ

- Posunutí, otočení, zkosení a rozdílná změna měřítka v obou směrech
- Minimálně tři páry identických bodů

Metody transformace

KOLINEÁRNÍ

- Nezachovávají se úhly, měřítko se mění v obou směrech nezávisle
- Minimálně čtyři páry identických bodů

POLYNOMICKÁ DRUHÉHO STUPNĚ

- Minimálně šest párů identických bodů

POLYNOMICKÁ TŘETÍHO STUPNĚ

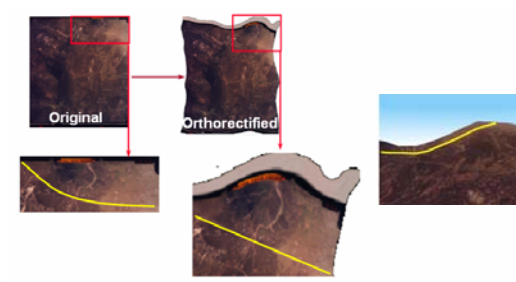
- Minimálně deset párů identických bodů

Transformace po částech

- Interpoluje po částech mezi identickými body
- Vhodnou volbou bodů lze zmírnit chybu způsobenou reliéfem

Ortorektifikace

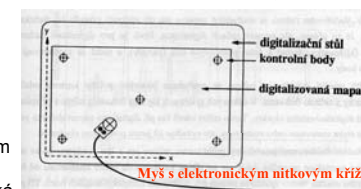
- Přesnější a náročnější
- Kromě identických bodů využívá digitální model terénu a údaje o senzoru



Vektorizace

1. Z papírového podkladu pomocí digitizeru

digitizer (tablet) = deska se zabudovanou jemnou, hustou a velice přesnou sítí vodičů → souřadnicový systém
Pojeno s počítačem – vektorizovaná kresba se zobrazuje na monitoru



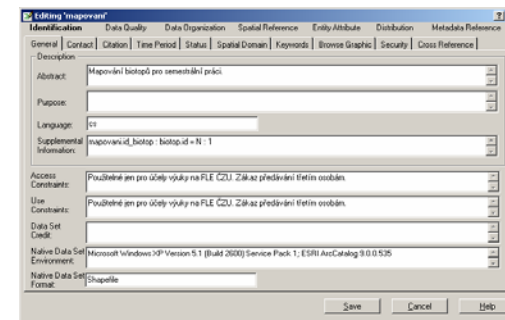
2. Vektorizace nad rastrem
 - A) Ruční
 - B) Poloautomatická
 - C) Automatická

Metadata

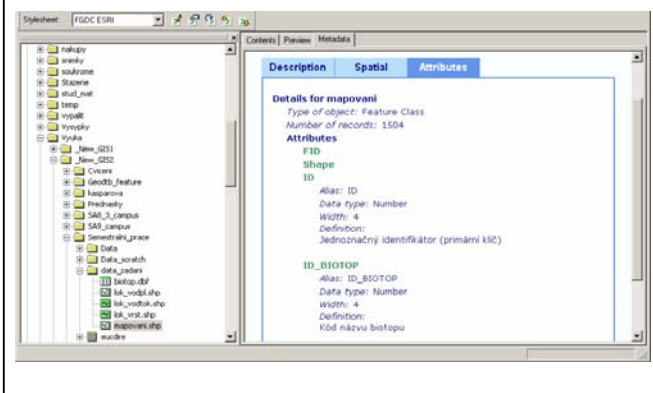
- „Data o datech“
- Popisují data tak, aby bylo vše srozumitelné jejich uživatelům
- Forma různá (text, tabulka, html...)
- Základní údaje pro vrstvy GIS
 - Co vrstva představuje, přibližné umístění a rozsah dat, souřadný systém, referenční měřítko, popis atributů – tabulky a vztahy mezi nimi, význam sloupců, význam zapisovaných kódů (popř. omezení hodnot), autor dat, zpracovatel dat, vlastník dat, za jakých okolností mohou být data poskytnuta dalším subjektům

XXX

Metadata v ArcGIS

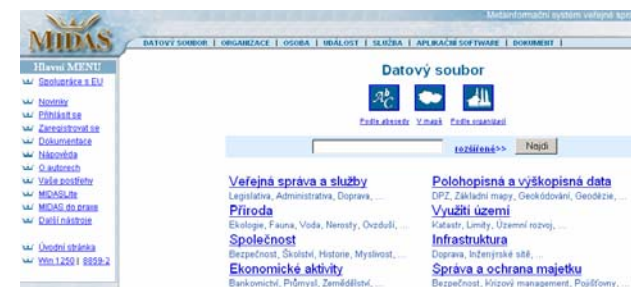


Metadata v ArcGIS



Příklad - Metainformační systém Midas

- Cenný zdroj informací o existujících geodatech
- <http://gis.vsb.cz/midas/>



Kontrolní otázka

- Provedli jste v terénu mapování (např. biotopů, krajinného rázu, lesních porostů...). Všechny výsledky máte zakresleny v papírovém mapovém podkladu. Pro další práci potřebujete topologicky správnou polygonovou vektorovou vrstvu.

Co budete dělat?