



Geoinformatika

II – GIS jako zpracování dat

jaro 2014

Petr Kubíček

kubicek@geogr.muni.cz

**Laboratory on Geoinformatics and Cartography (LGC)
Institute of Geography
Masaryk University
Czech Republic**



Komplexní GIS schéma

Transformace dat

- modelu
- polohy
- formátu

Sběr dat

- editace
- import

Uložení dat

- | | |
|-------------|------------------------------|
| Návrh | - struktura
- datové typy |
| Manipulace | - dotazování
- indexování |
| Dokumentace | - metadata |

Analýza dat

- průzkum
- modelování

Prezentace dat

Vizuální

- kartografická
 - statické mapy
 - dynamické mapy
 - uživatelské rozhraní
- nekartografická
 - grafická
 - textová

Nevizuální

- export
- řídicí povely

Historie GIS

- **V 50. letech 20. století začaly pokusy s automatizovaným mapováním za využití výpočetní techniky.**
- **V roce 1963 zavedl pojem GIS Kanadán R. F. Tomlinson a označil tak nové technologie pracující s daty a podávajícími informací o terénu pomocí výpočetní techniky.**
- **Jak dostat mapu do počítače?**

Historie GIS

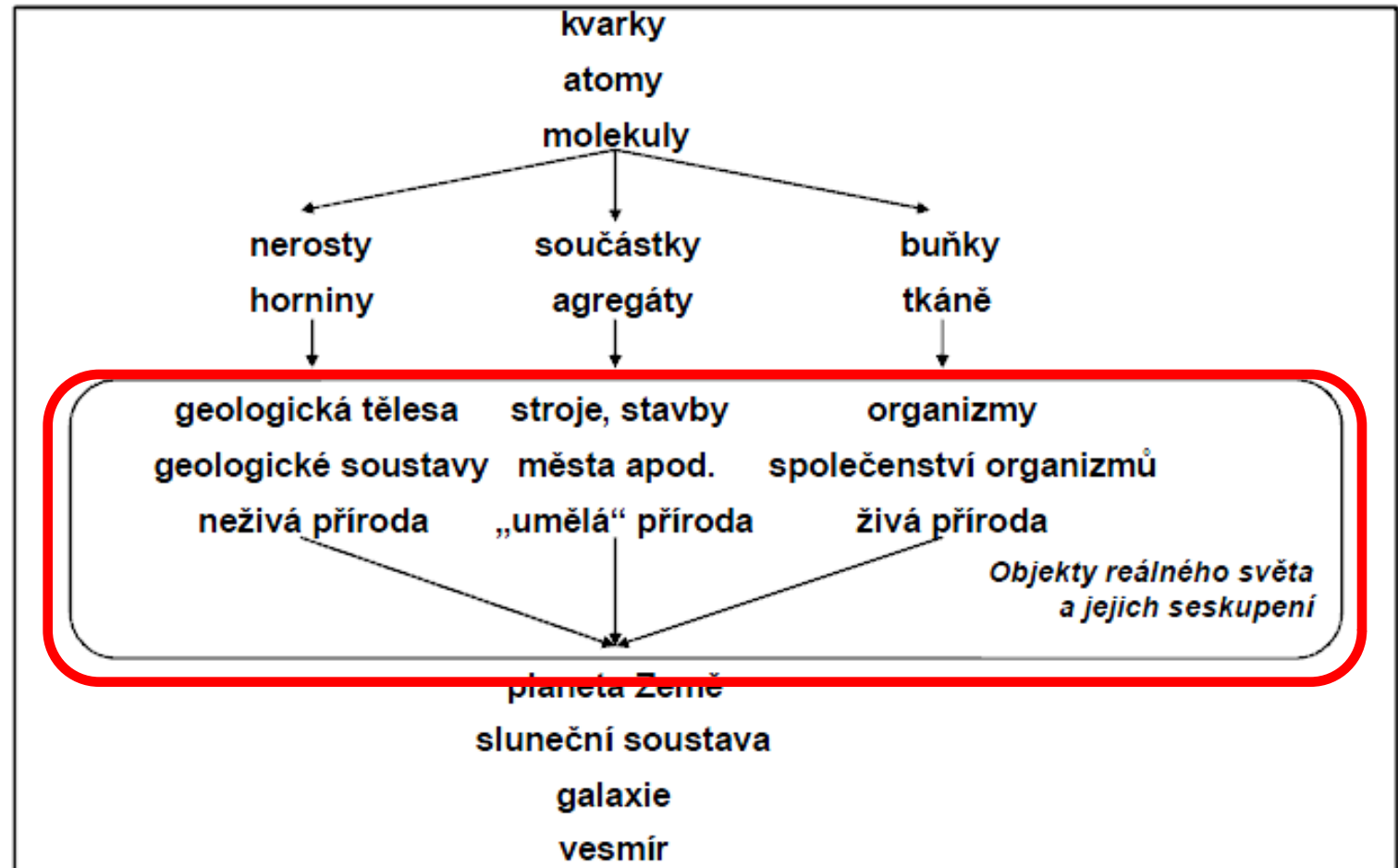
- **Pionýrské období (konec 60. let až 1975) - hlavně průkopnické práce, univerzity - důraz na digitální kartografii.**
- **1975- začátek 80. let - ujednocení pokusů s institucemi na lokální úrovni - první LIS.**
- **1982 - konec 80. let - komercializace problematiky – běžně dostupné softwarové systémy pro GIS (ESRI, Intergraph, ...), první systémy založené na CAD (systémy před tím měly minimální grafické možnosti). Autocarto, EuroCarto (1987 Brno).**
- **1986 – P. Burrough - učebnice GIS**

Historie GIS

- **NCGIA** – National Centre for Geographic Information and Analysis ;GI science (USA). Změna pojetí GIS jako nástroje (research with GIS) na přijetí GIS jako výzkumného směru (research about GIS – spatial information theory).
- **90. léta** - počátky standardizace, uživatelské GIS, Desktop GIS, otevřené systémy (Open GIS), Internet.
- **AGILE**; Conference on Spatial Information Theory (**COSIT**).
- **Současnost** - vývoj objektově orientovaných systémů, masivní propojení s databázemi, vzdálený přístup přes Internet/Intranet, webové služby
- **Mobilní GIS** ...

Pohledy na reálný svět

Tři základní pohledy - Objektový , Jevový, Procesní





Jevový a procesní pohled

- **Dělení jevů** – kvalitativní x kvantitativní; statické x dynamické; kontinuální x diskrétní.
- **Procesní pohled** – vnáší do reálného světa **dynamiku** a ovlivňuje jak jevy, tak objekty.
- Pro modelování jevů a objektů využívá geoinformatika odlišné datové modely.

The slide features a decorative header with a grid of vertical bars in various colors (blue, green, yellow, orange, pink, purple) on the left side. In the top-left corner, there is a logo for 'IGC' (Institute of Geomatics) which includes a stylized globe and the letters 'IGC' in blue. The main title 'Modelování geografických objektů' is written in a large, bold, dark blue font on the right side of the header.

Modelování geografických objektů

- v geoinformatice se nezaobíráme reálnými objekty, ale z důvodu zjednodušené reprezentace - **modelem reality**.
- **Modelování** - abstrahování týkající se všech součástí geografické informace:
 - geometrické,
 - topologické,
 - tematické
 - dynamické.

Proces modelování

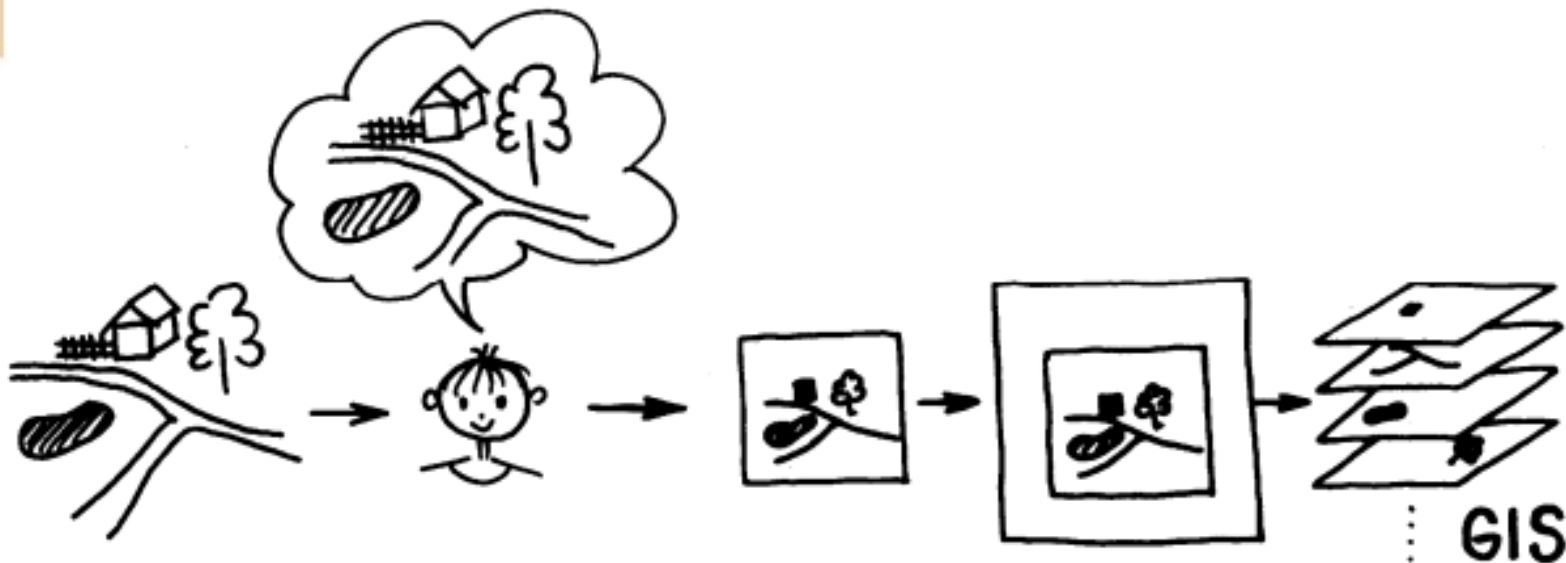
- **modelování** – je proces abstrakce, při kterém jsou podstatné elementy reálného světa zdůrazněny a nepodstatné eliminovány (s ohledem na cíl, který má toto modelování splnit):
 - úmyslně – zobrazují se jen ty elementy, které jsou předmětem zkoumání, ostatní se potlačují
 - neúmyslně – v dané fázi poznání jsou nedostupné či nepoznatelné
- Principem modelování je snaha o poznání vlastností studované části reality.



GIS jako obraz reálného světa

- **Reálný svět** je pozorován pozorovatelem. Ten vytváří na základě svých vjemů vnitřní (**mentální model**) tohoto světa.
- **Mentální model** je velmi blízký reálnému světu, ale vždy je tu jisté zjednodušení.
- Pro poskytnutí tohoto modelu dalším uživatelům je nutné ho převést do podoby **papírové mapy**.
- Pro převedení mapy do **GISu** je potřeba ji **zdigitalizovat**. Jednotlivé geoprvky jsou nahrazeny základními geometrickými prvky a ty jsou dále tříděny podle tématu do jednotlivých vrstev.

GIS jako obraz reálného světa



Výsledný obraz reálného světa v GISu je:

- složen z bodů, linií a polygonů
- rozříděných do jednotlivých vrstev
- dvourozměrný
- statický
- zjednodušený (chybí mnoho informací)
- obsahuje mnoho chyb a nepřesností vzniklých z převodem reality do podoby GIS.

Složky geografických dat

- **Neprostorová složka (tzv. Atributy)**
 - Čísla – kvantitativní hodnoty
 - Řetězce znaků – kvalitativní hodnoty
 - Datum – časové určení
 - Komplementární atributy – linky, videa, dokumenty ...
- **Časový aspekt** odráží změnu jevu v čase (od jednoho záznamu ke druhému)
- **Prostorová složka (tzv. Geometrie)**
 - tvar
 - poloha
 - topologie
- Všechny měřitelné nebo popsatelné vlastnosti reálných entit spadají do jednoho z aspektů: **prostoru**, **tématu** nebo **času**.



Neprostorová složka - atributy

Popisující geografické objekty a jejich vlastnosti

- **Typy atributových dat**

- *poměr* - např. procenta
- *interval* - např. celá čísla z intervalu (0,10), desetinná čísla z intervalu (0.5-14.0)
- *pořadí* (ordinální) - řadová číslovka
- *výčet* (nominální) - např. pro typ silnice to může být (dálnice, rychlostní silnice, silnice 1.třídy, silnice 2.třídy, ostatní silnice)

Příklad:

- **Objekt** = lesní porost
- **Atribut** = dřevinná skladba, průměrná výška porostu, věková struktura, apod.
- Atributy jsou neprostorové (nereprezentují informaci o lokalizaci či o prostorových vztazích), mají vytvořenou vazbu na prostorové prvky atributové hodnoty, reprezentující kvalitu geografického objektu, nelze vždy měřit nebo udávat v jednotném měřítku.
- Př. Borovicový porost není nikdy 100% složen pouze z borovice. Při analýzách to nevadí, ale je nutno s touto skutečností počítat (míry kvality).

Časová složka

- **Čas – dynamický popis**
 - dynamika charakterizuje časovou variabilitu geografických objektů.
 - tyto změny se mohou týkat geometrie, topologie i tematického popisu.
- **Modelování** dynamických prostorových procesů v rozměrném prostoru vyžaduje složité modely a metody.

V praxi se ale používá zjednodušení:

- 1. Analýza časové série na jednom měřícím bodě - **časová změna.**
- 2. Prostorová změna atributové hodnoty mezi dvěma body v tom samém čase - **prostorová změna.**

Prostorová složka - tvar

- **Dvě chápání prostoru – vektorový a rastrový**
 - Jev -> jeho vymezení -> hranice -> **objekty**
 - Kartografický model reality
 - Bod, Linie, Plocha
 - Uspořádaný soubor souřadnic (vektor)
 - Prostor -> jeho rozdělení -> jaký **jev** je přítomen v dílu
 - Vzorkování
 - Různé druhy členění podle charakteru buněk
 - Pravidelnost, komplexita
 - Matice hodnot – u pravidelného a jednoduchého členění.

Datové modelování – vektorová reprezentace

Základní geometrické objekty

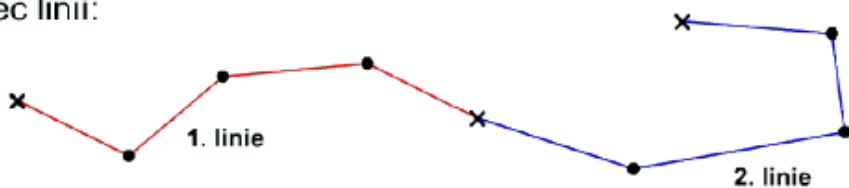
Bod:

x
 $[x,y]$

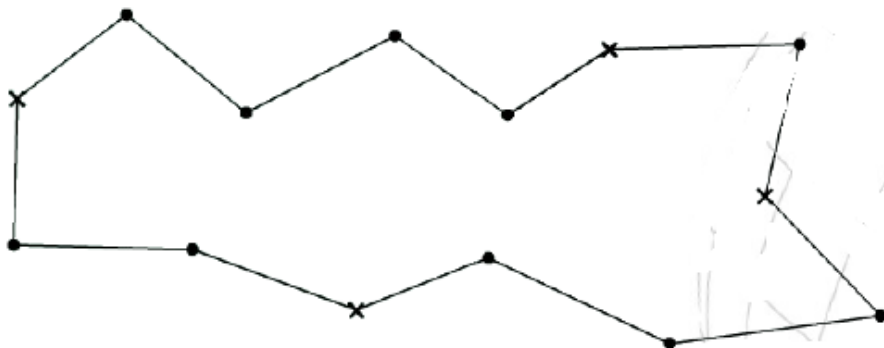
Linie:



Řetězec linií:



Plocha - uzavřený řetězec linií:





Typy geometrických objektů

- **Bod**

- nemá délku, hloubku ani šířku - bezrozměrný (0D) prvek
- je jednotlivý pár souřadnic X, Y, reprezentující geografický prvek
- je příliš malý na to, aby byl zobrazen jako linie či plocha.

- **Linie**

- má délku, ale nemá šířku ani hloubku - jednorozměrný (1D) geografický prvek;
- je sled orientovaných úseček (hran) definovaných souřadnicemi vrcholů (vertex) mezi dvěma uzly (nodes);
- tvar reprezentovaného geografického prvku je příliš úzký na to, aby mohl být zobrazen jako plocha.

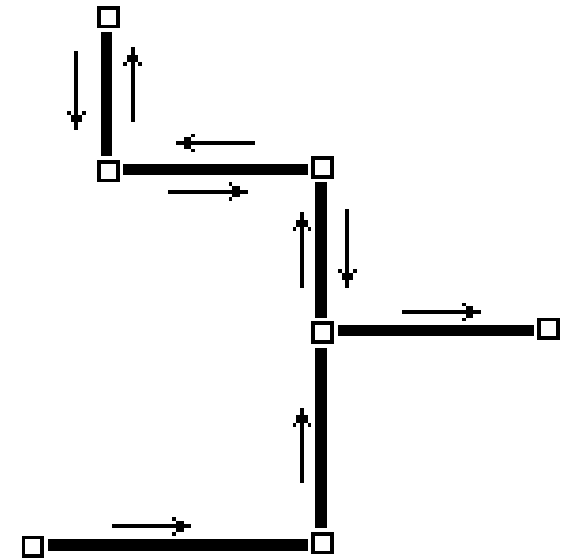
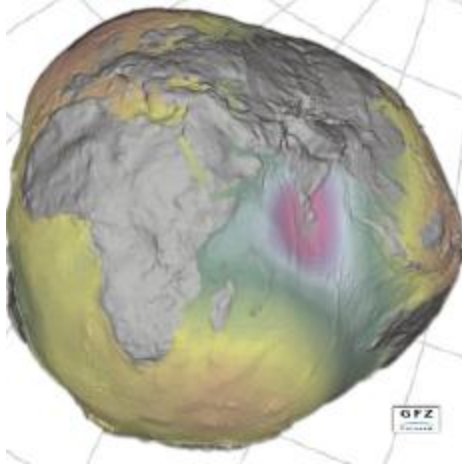
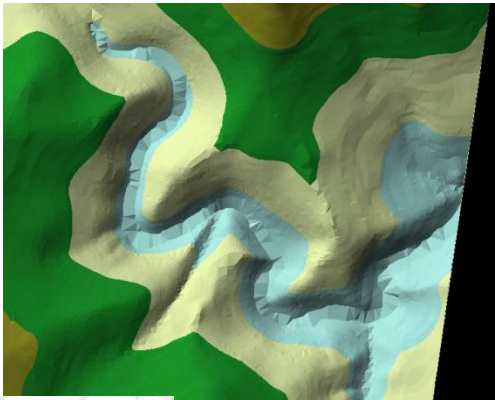
- **Polygon (plocha)**

- mají délku a šířku, ale nemají hloubku - dvojrozměrný (2D) geografický prvek;
- je uzavřený obrazec, jehož hranicí je uzavřená linie.



Sítě

- systém linií s topologickou strukturou;
- je řada vzájemně propojených linií, podél níž probíhá tok informací.



Povrchy

- je to souvislá entita, pro kterou v každém bodě existuje nějaká hodnota (2,5 D) x, y, z .

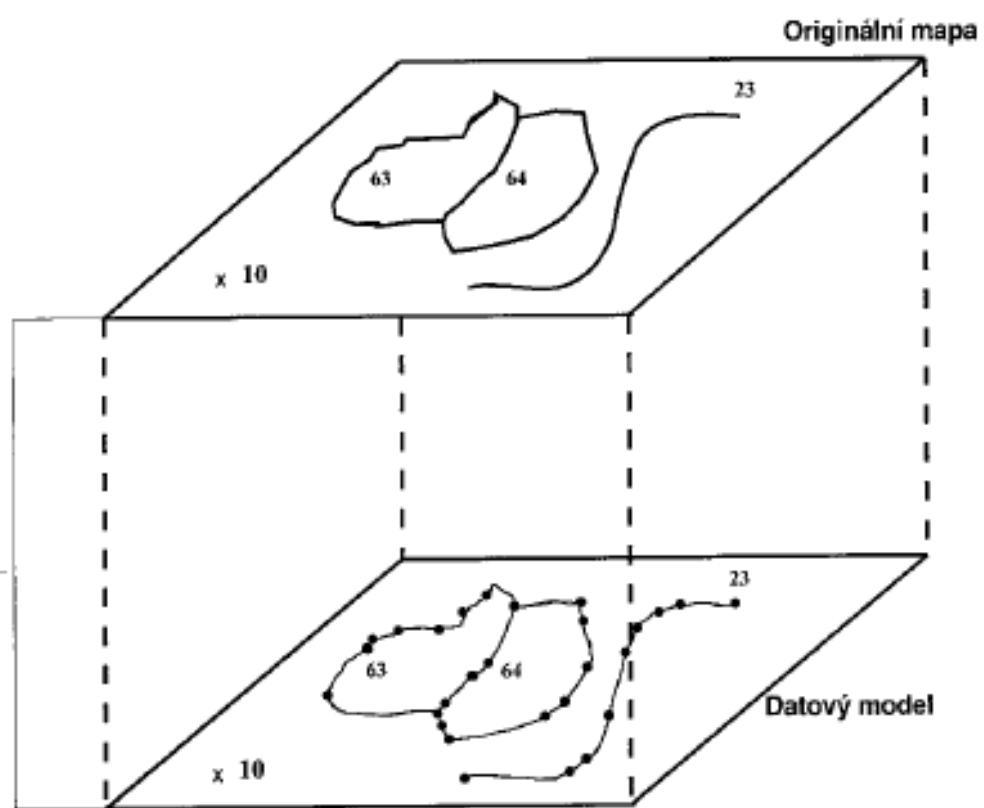
Objemy

- mají všechny rozměry (délku, šířku, hloubku) - trojrozměrné (3D) geografické prvky.



Špagetový datový model

- Nejjednodušší
- Objekt na mapě se reprezentuje **jedním logickým záznamem** v souboru a je definovaný jako **řetězec x,y souřadnic**.
- Nevýhody - ačkoli jsou všechny objekty v prostoru definovány, struktura **neposkytuje informace o vztazích mezi objekty**.
- **Společná linie je pro každý polygon ukládána dvakrát**.
- Pro většinu prostorových analýz je tento model nevhodný, protože veškeré potřebné prostorové vztahy musí být spočítány před každou analýzou



Datová struktura

Objekt	Číslo	Poloha
Bod	10	X, Y Jednotlivý bod
Čára	23	$X_1 Y_1, X_2 Y_2, \dots, X_n Y_n$ Řetězec
Polygon	63	$X_1 Y_1, X_2 Y_2, \dots, X_1 Y_1$ Uzavřená smyčka
	64	$X_1 Y_1, X_2 Y_2, \dots, X_1 Y_1$



Topologický datový model

- V tomto modelu každá linie začíná a končí v bodě nazývaném **uzel - node**.
- **Dvě linie se mohou** protínat opět jenom v uzlu. Každá část linie je uložena s odkazem na uzly a ty jsou uloženy jako soubor souřadnic x,y. Ve struktuře jsou ještě **uloženy identifikátory označující pravý a levý polygon vzhledem k linii**. Tímto způsobem jsou zachovány základní prostorové vztahy
- **Použitelné pro analýzy**. Navíc tato topologická informace umožňuje body, linie a polygony uložit v neredundantní podobě (bez opakovaného zápisu).

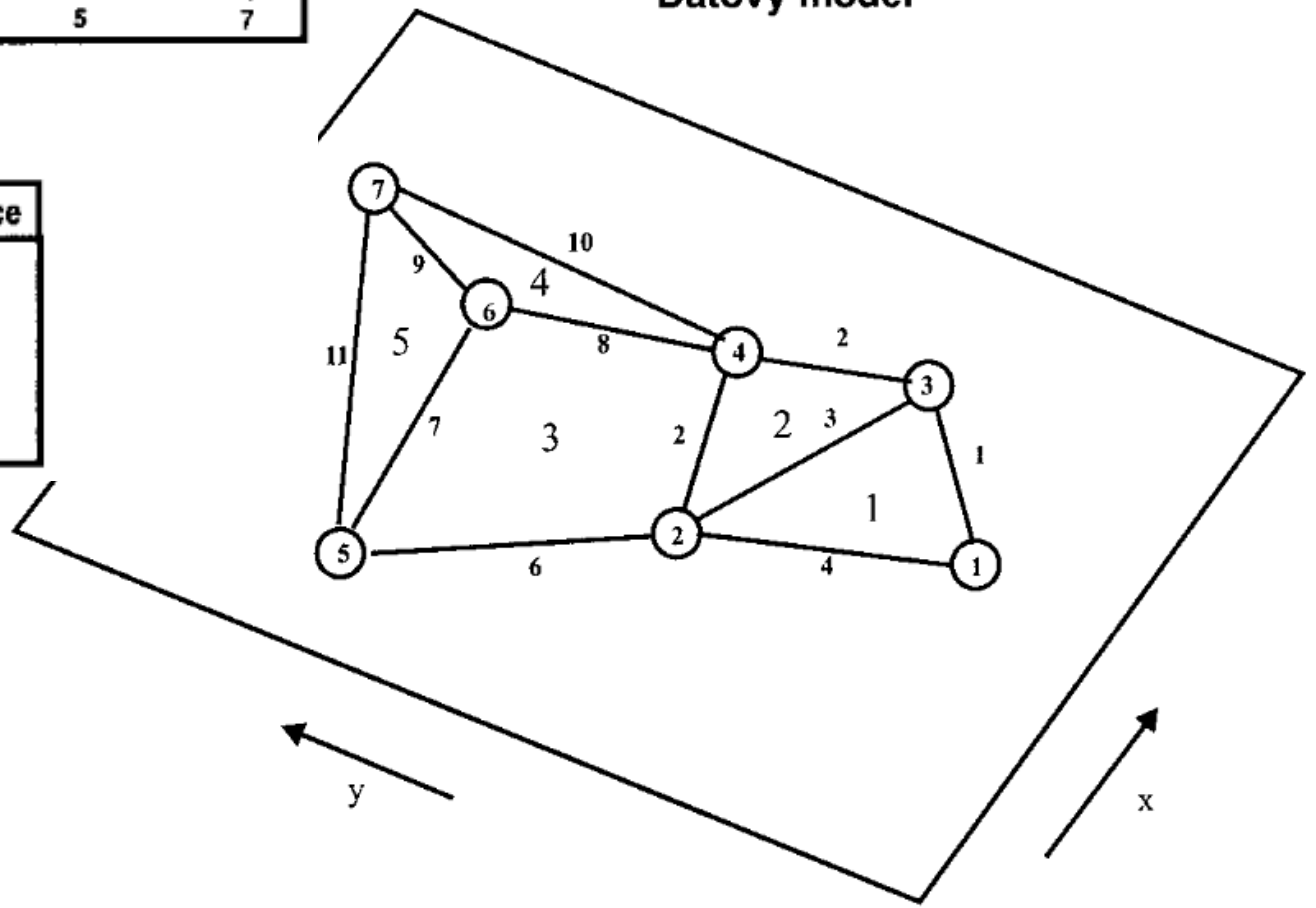
Soubor topologických vztahů

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
6	3	0	5	3
7	3	5	5	6
8	3	1	6	4
9	4	5	7	6
10	0	4	7	4
11	5	0	5	7

Soubor souřadnic bodů

Uzel	X souřadnice	Y souřadnice
1	23	8
2	17	17
3	29	15
4	26	21
5	8	26
6	22	30
7	24	36

Datový model





Vektorová reprezentace - topologie

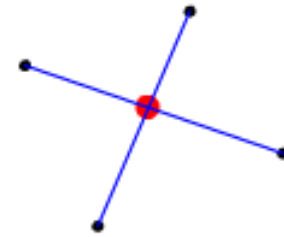
- **Topologie je matematický způsob, jak explicitně vyjádřit prostorové vztahy mezi jednotlivými geometrickými objekty.**
- **Proč vůbec topologie? Má jisté výhody, například:**
 - Umožní ukládat data efektivněji.
 - Mnoho analýz v GIS využívá pouze topologické a nikoli geometrické vztahy.
- **Důvod pro využívání topologie (ESRI 1995):**
- "Topology is useful in GIS because many **spatial modeling operations don't require coordinates, only topological information**. For example, to find an **optimal path between two points** requires a **list of the arcs** that connect to each other and the **cost to traverse each arc in each direction**. Coordinates are only needed for drawing the path after it is calculated."



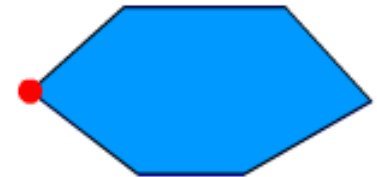
Tři základní topologické koncepty

- **Konektivita** – dvě linie se na sebe napojují v uzlech.
- **Definice plochy** – linie, které uzavírají nějakou plochu, definují polygon.
- **Sousednost** - linie mají směr a nesou informaci o objektech nalevo a napravo od nich.

Konektivita



Definice plochy



Princip okřídlené hrany:



Vektorová data

Výhody

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

Nevýhody

- výpočtová náročnost (problémy při náročných analytických operacích);
- komplikovanost datové struktury;
- složitější odpovědi na polohové dotazy;
- obtížná tvorba překryvů vektorových vrstev (overlay)
- problémy při modelování a simulaci jevů.