

## Nebojme se databází

Existují věci, před kterými utíkáme. Mezi nimi jsou i nové poznatky, nové teorie a nové přístroje. A přitom jsou pro další odborný růst v řadě profesí nepostradatelné. Co si představí většina geografů pod pojmem **databáze na počítači**? Nejčastěji právě jakéhosi strašáka, před kterým je nejlepší rychle utéct. Obzvláště když mají jména jako dBase, Oracle nebo Informix. Přitom si stačí představit například kartotéku či evidenci, která má svůj původ v klasické lístkové kartotéce známé z knihoven nebo lékařské ordinace. V geografii se databáze stávají základním prostředím správy většiny geografických dat. Proto se setkáváme s demografickými, klimatologickými i geologickými databázemi.

### Databáze je, když ...

Databáze na počítači je ve své podstatě také kartotékou, avšak umožňuje provádět mnohem více úloh. Už proto, že u lístkové kartotéky je přístup omezen jedním, většinou abecedním uspořádáním evidenčních lístků. Charakteristickou vlastností počítačových databází je možnost **zpracování velkého množství dat**. Dalšími vlastnostmi je **rychlý přístup k datům**, **snadná aktualizace**, **rychlé zpracování**. Data uložená v databázích mívají **trvalý charakter**. Oproti hardwaru a softwaru vykazují z časového hlediska delší trvalost. Jsou platná dokud existuje realita, kterou popisují. Jakmile realita přestává existovat, slouží pro archivní účely. Proto mají například data o zemském povrchu delší platnost než data o počasí, obojí jsou ovšem velmi důležitá. Z hlediska ekonomického je pořízení dat nejnákladnější součástí informačního systému, hovoří se až o 80 % z celkových nákladů jeho budování. Ovšem náklady se velmi rychle vracejí snadným a přehledným přístupem k datům a jejich rychlým zpracováním.

### Relační databáze

Nejběžnějším typem počítačových databází jsou v současnosti relační databáze. Jejich základním rysem je ukládání dat ve formě **vzájemně propojených tabulek**. Databáze proto nejsou jen jednou tabulkou, ale souborem různého počtu tabulek. Objevuje se otázka, proč se nepoužívá namísto relačních databází jednoduchý tabulkový procesor, který umí s tabulkami pracovat, například běžně dostupný Microsoft Excel? Odpovědí je potřeba definování vztahů mezi tabulkami.

V záhlaví tabulky se definují vlastnosti objektů - **atributy**. V každém řádku tabulky je uložen jeden **záznam** popisující právě jeden objekt například město. Atributy mohou být různého typu - textový řetězec, číslo nebo příznak ano/ne. V některých databázových programech, např. Microsoft Access, lze uložit ve formě atributu i fotografie či textové dokumenty. V databázi měst jsou typickými atributy název města, rozloha, počet obyvatel, nadmořská výška, počet ZŠ a dále třeba letecké ortofoto či textový dokument s historií města.

Jaký je **vztah mezi tabulkami**? Do jedné tabulky uložíme například údaje o krajích a do druhé údaje o městech.

Kraj		Město			
Kraj	Rozloha [km <sup>2</sup> ]	Název města	Počet obyvatel	Nadm. výška [m n.m.]	Ortofoto
Jihočeský	10056	České Budějovice	99708	379 - 561	D:\cb.jpg
Jihomoravský	7062	Třeboň	9083	435	D:\tr.jpg
Olomoucký	5139	Brno	383 569	190 - 425	D:\br.jpg
		Olomouc	103 293	208 - 420	D:\ol.jpg
		Přerov	47 988	210	D:\pr.jpg
		Prostějov	49477	225	D:\pv.jpg

Z tabulek je patrné, že do jednoho kraje spadá více měst. Vazba mezi tabulkami se provádí pomocí **primárního klíče** a **cizího klíče**. Do tabulky Kraj se přidá identifikační číslo, které nemá charakter popisného atributu, ale jen funkci primárního klíče, který jednoznačně identifikuje kraj. Do tabulky měst se přidá identifikační číslo kraje, do kterého dané město patří. Zde hovoříme o tzv. **cizím**

**klíči.** Tímto způsobem se realizuje vazba (relace) mezi tabulkami. Při výpisu libovolného přehledu o městech a krajích, se identifikační číslo nezobrazuje. Je skryto uvnitř relační databáze a pouze pracuje v náš prospěch.

Číslo kraje	Kraj	Rozloha [km2]
36	Jihočeský	10056
116	Jihomoravský	7062
124	Olomoucký	5139

Město

Číslo kraje	Název města	Počet obyvatel	Nadmořská výška [m.n.m.]	Ortofoto
36	České Budějovice	99708	379 - 561	D:\cb.jpg
36	Třeboň	9083	435	D:\tr.jpg
116	Brno	383 569	190 - 425	D:\br.jpg
124	Olomouc	103 293	208 - 420	D:\ol.jpg
124	Přerov	47 988	210	D:\pr.jpg
124	Prostějov	49477	225	D:\pv.jpg

### Vztahy

Databázové vztahy mohou být různé, hovoří se o stupních vztahů. Výše uvedený příklad je typický vztah **1:N** (jedna ku N) – do jednoho kraje patří více měst. Dalším příkladem může být: do jednoho úmoří náleží více řek. Jednodušším stupněm vztahu je vztah **1:1** (jedna ku jedné). Příkladem takové vazby je vztah kraj - krajské město, dále například jeden občan ČR - jedno rodné číslo. V geografické praxi je vždy potřeba pečlivě zvážit, zda je rentabilní zakládat pro takovéto objekty samostatné tabulky, protože je mnohem úspornější sloučit je do jediné tabulky. Ovšem mohou existovat i důvody, proč je ponechat v samostatných tabulkách, např. velký počet atributů jednotlivých objektů. Nejsložitějším databázovým vztahem je vztah **M:N**. To je totéž jako m-krát opakovaný vztah 1:N. Příkladem může být vztah, kdy jedna parcela má více vlastníků, nebo naopak, kdy jeden vlastník má více parcel. Vztah M:N je v datovém modelu nutné převést (dekomponovat) na dva vztahy 1:N. A to z důvodu, že databázové systémy neumí tento vztah realizovat přímo.

Ukážeme si to na příkladu úředních jazyků. Vazba stát – úřední jazyk je typu M:N. Dekompozicí vznikne třetí tabulka Vazba, která je **spojující tabulkou**.

Stát

ID_stát	Stát
101	Belgie
102	Lucembursko
103	Francie

Vazba

ID_stát	ID_jazyk
101	1
101	2
101	3
102	1
102	3
102	4
103	1

Jazyk

ID_jazyk	ÚředníJazyk
1	francouzština
2	nizozemština
3	němčina
4	lucemburština

Vraťme se zpět ke klasické papírové kartotéce. Předmět evidence (mapy, vrty, sčítací obvody, města aj.) je uspořádán podle určitého pravidla, například v abecedním pořadí podle názvu. Vyhledání je díky používání papírových záložek s písmeny abecedy na horním okraji poměrně rychlé. Vkládání nové papírové karty je ale složitější a pomalejší, protože musíme novou kartu umístit přesně na pozici podle abecedy. V počítačových databázích se uloží nový záznam vždy nakonec. Velký rozdíl nastane, jestliže požadujeme vyhledání podle jiného kritéria než je název, například města podle počtu obyvatel. V papírové kartotéce je to mnohem obtížnější, protože musíme projít všechny lístky několikrát a postupně vyhledávat města s maximální počet obyvatel a ty odkládat. Díky manuálnímu třídění kartotéky (a po skončení práce zpětnému sestavení) je práce s počítačovou databází mnohonásobně efektivnější.

Proto je nespornou výhodou počítačových databází snadné **řazení záznamů** podle potřeby - abecedně podle názvu, abecedně podle názvu kontinentů a v rámci kontinentů abecedně podle názvu států, dále třeba podle rozlohy (od nejmenšího k největšímu a naopak.), atd.

## Databázové úlohy

K čemu se naplněné databáze nejčastěji používají? K **vyhledávání**, to je nejčastější úloha v databázích. K vyhledávání potřebných informací se v databázích sestavují tzv. **vyhledávací dotazy**. Při práci s již fungující databázovou aplikací jsou dotazy předchystány. Geograf pouze zadává parametry dotazu čili specifikuje podmínku vyhledávání. Například při vyhledávání informací o státu zadaného jména obdržíme jako odpověď všechny údaje o jednom konkrétním státu - samozřejmě ty, které jsou v databázi uloženy. Na dotaz "vyber všechny sčítací obvody s volební účastí přes 60 %", získáme přehledný výpis obvodů (předem nevíme kolik jich bude) a okamžitě můžeme nad touto sestavou provádět nejrůznější statistické výpočty. Jiným druhem dotazů jsou **agregační dotazy** s formulací například "kolik vrtů v lokalitě je hlubších než 100 metrů". Odpovědí je počet vrtů, nikoliv jejich seznam. V databázích je možné také vypočítávat údaje, například hustotu zalidnění (jsou-li v databázi údaje o rozloze státu a počtu obyvatel), a ukládat je do nových atributových položek. Kromě dotazů, které jsou nejmocnějším prostředkem k získávání informací z tabulek, je možné sestavit i **akční dotazy** pro změny v tabulkách. Změny se mohou týkat jednoho objektu (např. aktualizace údaje o počtu obyvatel) nebo hromadně všech (např. všem vybraným zemím změň měnu na Euro).

## Kam za databázemi?

Většina geografů bude pracovat již s hotovými počítačovými relačními databázemi. Tedy s takovými, které již někdo sestavil (navrhl strukturu tabulek a vazeb) a naplnil je daty. Na webových stránkách Agentury ochrany přírody a krajiny ČR (<http://www.nature.cz>) je ke stažení Informační systém ochrany přírody ISOP. Na serveru Ministerstva práce a sociálních věcí (<http://www.mpsv.cz>) je veřejně přístupný Územně identifikační registr adresních míst v ČR UIR-ADR, který je pravidelně 1x týdně aktualizován.

Registr adres České republiky, verze dat 327 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address [http://uiradr.mpsv.cz/uir/aplikace\\_v3/default2.asp?najit=ulici](http://uiradr.mpsv.cz/uir/aplikace_v3/default2.asp?najit=ulici)

**Hledat** ↑

okres

obec

pr. obvod

měst. část

část obce

**ulici**

objekt

adresu

poštu

[www.mpsv.cz](http://www.mpsv.cz)

**Omezit území - vyberte obec**

okres **Tábor** 3308 (77)

obec libovollná (16)

**Jméno ulice:** obsahuje Skálova vyhledat

PSČ:  Zobrazit:  kódy  datumy  zkratky  odkazy

#	ulice	obec
1	Skálova	Tábor

Stránka 1/1; počet záznamů = 1

Jiné geografické databáze prodává Český statistický úřad (<http://www.czso.cz>), Geofond ČR (<http://www.geofond.cz/>), ČHMÚ a tisíce dalších organizací v Česku i po celém světě. Vyhledávat ve statických datech zejména o populaci světa a ekonomických ukazatelích lze na adrese [www.geohive.com](http://www.geohive.com).

Velice často se na webových stránkách o klimatických údajích linkují údaje z databází, neboť se jedná o údaje, které se mění každý den. Tudíž nemohou být tyto údaje součástí statických webových stránek. Příkladem je [www.povodi.cz](http://www.povodi.cz) - Stavy a průtoky povodí Moravy, kde naleznete aktuálně průtoky a stavy na jednotlivých měřicích stanicích 4x za den.

Nejužitečnější by samozřejmě byla databáze o databázích. Taková se v Česku vytváří v rámci aktivit České asociace pro geoinformace ČAGI. Jedná se o systém MIDAS, který eviduje informační zdroje veřejné správy a obsahuje důležité informace o existenci a vlastnostech jiných dat. MIDAS (<http://gis.vsb.cz/midas/>) je rozsáhlou evidencí geografických dat zejména z okresních úřadů ale i soukromých subjektů.

Je zbytečné se databází bát. Jsou užitečné a prospěšné. Obzvláště geografům ...

Ing. Zdena Dobešová  
Katedra geoinformatiky  
Univerzita Palackého v Olomouci