

Vedecká rada Stavebnej fakulty  
Slovenskej technickej univerzity v Bratislave

Ing. Ivana Ivánová

**Autoreferát dizertačnej práce**

**Data quality in spatial datasets**  
**Kvalita údajov v priestorových databázach**

na získanie

vedecko-akademickej hodnosti philosophiae doctor

v odbore doktorandského štúdia:

39-31-9 Geodézia a geodetická kartografia

Bratislava 2006

Dizertačná práca bola vypracovaná v internej forme doktorandského štúdia na Katedre geodetických základov Stavebnej fakulty STU v Bratislave.

**Predkladateľ:** Ing. Ivana Ivánová  
Katedra geodetických základov  
Stavebná fakulta STU Bratislava  
Radlinského 11, 813 68 Bratislava

**Školiteľ:** doc. Ing. Ernest Bučko, PhD.  
Katedra geodetických základov  
Stavebná fakulta STU Bratislava  
Radlinského 11, 813 68 Bratislava

**Oponenti:** prof. Ir. Henri J.G.L. Aalders  
OTB Research Institute fo Housing, Urban and Mobility  
Studies, Jaffalaan 9, 2628 BX Delft

doc. Ing. Petr Rapant, PhD.  
Inštitút geoinformatiky  
VŠB-TU Ostrava  
17. listopadu 15, 708 33 Ostrava – Poruba

doc. RNDr. Dagmar Kusendová, PhD.  
Katedra humánnej geografie a demogeografie,  
Prírodovedecká fakulta Univerzity Komenského v Bratislave  
842 15 Bratislava, Mlynská dolina, pavilón CH1-B1

Autoreferát bol rozoslaný dňa:.....

Obhajoba dizertačnej práce sa koná dňa:..... o ..... h.  
pred komisiou pre obhajobu dizertačnej práce v odbore doktorandského štúdia  
vymenovanou predsedom spoločnej odborovej komisie .....

39-31-9 Geodézia a geodetická kartografia

na Katedre geodetických základov, SvF STU v Bratislave, Radlinského 11, 813  
68 Bratislava

Predseda spoločnej odborovej komisie:  
Prof. Ing. Alojz Kopáčík, PhD.  
Katedra geodézie SvF STU  
Radlinského 11, 813 68 Bratislava

# 1. ÚVOD

Informačná spoločnosť kladie najväčší dôraz na výmenu a zdieľanie znalostí. Výmena znalostí je výmenou základných charakteristík prvkov alebo výmenou skúseností o používaní týchto prvkov v príslušnej problémovej doméne. Priestorové dáta sú súčasťou základnej informačnej infraštruktúry, ktorej súčasťou sú rezorty z takmer všetkých oblastí ľudskej činnosti. Počas vývoja priestorovej informačnej infraštruktúry na rôznych úrovniach vzrástla potreba definície a zabezpečenia kvality údajov, ktoré sú jej základným stavebným prvkom. Pred tým ako sa začne diskusia o parametroch, hodnotení a zabezpečení kvality priestorových dát je nevyhnutné začať s diskusiou o tom čo rozumieme pod pojmom „kvalita“ z pohľadu priestorovej informačnej infraštruktúry. Je kvalita priestorových dát charakterizovaná polohovou presnosťou, ktorá je kľúčová z pohľadu geodetov? Alebo je charakterizovaná atribútovou presnosťou, kľúčovou z pohľadu tematickej kartografie? Je azda kvalita priestorových údajov z konceptuálneho hľadiska niečo celkom iné?

Dizertačná práca je zameraná na problematiku definovania kvality v priestorových databázach v rovine abstrakcie reálneho sveta do priestorových databáz za účelom poskytnutia informácie o kvalite prostredníctvom metadát. V práci je rozpracovaná komplexná teória hodnotenia kvality údajov z pohľadu producenta priestorových dát. Teória hodnotenia kvality priestorových údajov je aplikovaná v modeli domény katastra nehnuteľností v Slovenskej republike. Výsledky dizertačnej práce vyúsťujú do návrhu metodiky hodnotenia kvality počas celého procesu tvorby priestorovej databázy a jej analýzy v Základnej báze pre geografický informačný systém v Slovenskej republike.

## 2. KVALITA PRIESTOROVÝCH ÚDAJOV – SÚČASNÝ STAV

Termín „kvalita“ väčšinou vyjadruje výnimočnosť tovaru alebo atest vysokého stupňa prevedenia alebo umeleckej hodnoty. Vo výrobnom procese určuje kvalita tovaru očakávaný výsledok celého procesu riadenia výroby. Kontrola kvality vo výrobnom procese má svoju dlhú históriu.

Je veľmi obtiažne definovať kvalitu údajov, pretože tieto nemajú také fyzické vlastnosti, ktoré by napomáhali ľahšiemu odhadu kvalitatívnych parametrov. Kvalita údajov je výsledkom takých čiastkových charakteristík ako sú napr. úplnosť a konzistentnosť údajov. Ak je definovaná kvalita údajov, môžu používatelia posúdiť vhodnosť použitia týchto údajov vo svojich projektoch podobným spôsobom ako si vyberajú tovar v obchode. Pri posudzovaní vhodnosti a použiteľnosti priestorových údajov treba vziať do úvahy kvalitu celého procesu modelovania reality, t.j. kvalitu konceptuálneho analytického modelu, kvalitu operačného modelu a jeho reprezentáciu bázou údajov. V procese modelovania kvalita údajov sprostredkovane svojimi spätnými väzbami ovplyvňuje realitu a jej modelovanie. Norma STN EN ISO 19101:2005 – Geografická informácia – Referenčný model definuje kvalitu geografických informácií nasledovne:

„Kvalita je súhrn charakteristík produktu, ktoré vplyvajú na jeho schopnosť splniť stanovené alebo naznačené požiadavky.“

Je možné tvrdiť, že minimálne z pohľadu priestorových informácií kvalita nie je vyjadrením v zmysle „dobré verzus zlé“, ale vyjadrením „vhodnosti na použitie“.

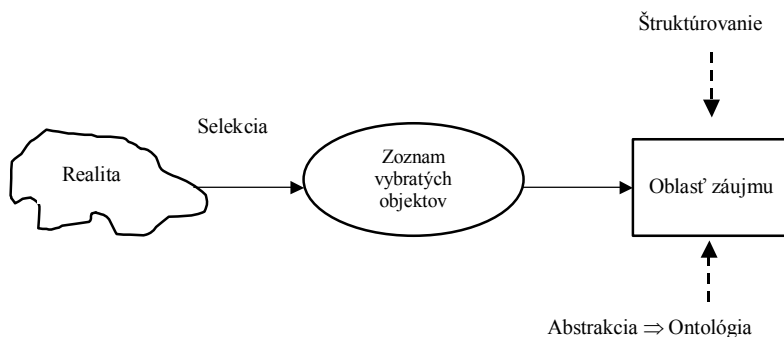
Z pohľadu tvorcu priestorovej databázy výsledky hodnotenia kvality priestorových údajov hovoria o dodržaní špecifikácie. Používateľ priestorovej databázy sa na základe informácie o kvalite a porovnaním so svojimi požiadavkami môže rozhodnúť použiť priestorové dáta pre svoj účel. Informácia o kvalite je súčasťou metadát priestorovej databázy.

## 2.1 PROCES TVORBY PRIESTOROVEJ DATABÁZY Z POHĽADU HODNOTENIA KVALITY PRIESTOROVÝCH ÚDAJOV

Oblasť záujmu (OZ z ang. Universe of Discourse) je podľa normy STN EN ISO 19101 definovaná ako pohľad na reálny alebo hypotetický svet, ktorý obsahuje všetko, čo skúmajúceho zaujíma. OZ je abstraktným modelom databázy obsahujúci výlučne objekty záujmu. Abstrakcia reálneho sveta do OZ pozostáva z nasledujúcich krokov:

- **Selekcia** objektov oblasti záujmu,
- **Abstrakcia** vybraných objektov ich popisom, definovaním atribútov a vzťahov medzi objektmi,
- **Štruktúrovanie** objektov a vzťahov medzi nimi.

Proces abstrakcie reálneho sveta do OZ je schematicky znázornený na obr.1:



**Obr. 1** Oblasť záujmu ako výsledok abstrakcie reálneho sveta

Špecifikácia OZ je aplikačne závislá a rozumie sa pri nej tvorba konceptuálnej schémy a popis obsahu bázy priestorových údajov. Predchádzajúcu definíciu je možné priblížiť pomocou jazykovednej analógie: tvorba konceptuálnej schémy je v tom prípade identická s tvorbou gramatiky jazyka a popis obsahu bázy je porovnateľný so slovníkom pre konkrétny jazyk. Proces špecifikácie môžeme rozdeliť na dva základné kroky:

- **geometrická** špecifikácia – definuje tvar a absolútnu, resp. relatívnu polohu objektov;
- **sémantická** špecifikácia – definuje popis objektov a ich atribútov, vzťahov medzi nimi a ich operácie, resp. funkcie. Sémantická špecifikácia objektov bázy priestorových údajov obsahuje tiež aj popis ich kvantitatívnych a kvalitatívnych charakteristík.

V procese špecifikácie OZ je možné uvažovať napr. aj nad špecifikáciou vstupu, ktorá definuje proces zberu objektov reálneho sveta do bázy priestorových údajov. Súčasťou špecifikácie OZ je aj množina parametrov kvality, ktorú producent očakáva po splnení podmienok pre vytvorenie predmetnej priestorových databázy.

Po abstrakcii reálneho sveta, výbere objektov do OZ a jej špecifikácii je možné prísť k samotnej tvorbe priestorovej databázy. Proces tvorby priestorovej databázy pozostáva z nasledujúcich krokov (Aalders, 2000):

- definícia objektov, ich atribútov a vzťahov medzi nimi;
- geometrická definícia referenčného systému, používaných jednotiek, rozlíšiteľnosti;
- definícia logických vzťahov v priestorovej databáze – topológie, atribútových obmedzení a pod.
- definícia kvality údajov a sémantickej kvality priestorovej databázy;
- definícia času – používanej časovej sústavy;
- aktualizácia priestorovej databázy – frekvencia, metódy a pod.
- popis pôvodu priestorovej databázy;
- možnosti aplikácie – prístup, prípady použitia a pod.

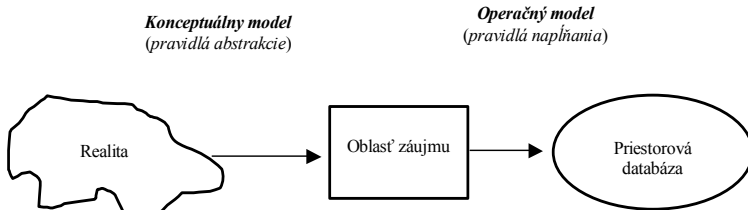
## 2.2 KVALITA PROCESU TVORBY PRIESTOROVEJ DATABÁZY

Koncepcia hodnotenia kvality vytvára priestor pre producentov aj používateľov bázy údajov pre jej ďalšie použitie. Producent hodnotí kvalitu z pohľadu dodržania vopred stanovených kvalitatívnych parametrov v špecifikácii priestorovej databázy. Používateľ sa na základe popisu kvality môže rozhodnúť pre použitie databázy na svoj konkrétny účel.

Pri modelovaní kvality priestorových údajov je nevyhnutné definovanie procesu tvorby priestorovej bázy údajov od abstrakcie objektov reálneho sveta, cez špecifikáciu oblasti záujmu až po naplnenie priestorovej databázy. Celý tento proces je v základoch ovplyvnený nasledujúcimi dvoma pravidlami (Uitermark, 2001):

- pravidlami **abstrakcie** – ktoré definujú proces abstrakcie reálneho sveta do priestorovej databázy (t.j. určujú objekty nachádzajúce sa v oblasti záujmu a žiadne iné);
- pravidlami **modelovania** – na základe týchto pravidiel prebieha naplnenie priestorovej databázy. Tieto pravidlá obsahujú pravidlá kategorizácie, zjednodušenia, agregácie a reprezentácie objektov reálneho sveta. Taktiež zahŕňajú pravidlá popisovania vzťahov medzi objektmi. Pre každý účel tvorby priestorovej databázy sú definované iné pravidlá naplnenia.

Pre jednoduchosť je možné vyššie uvedené rozdelenie zhrnúť: pravidlá abstrakcie popisujú proces selekcie a abstrakcie objektov reálneho sveta do OZ a pravidlá modelovania definujú proces transformácie objektov reálneho sveta na objekty priestorovej databázy (obr.2).



**Obr. 2 Pravidlá abstrakcie reálneho sveta do priestorovej databázy**

Pri posudzovaní kvality procesu tvorby priestorovej databázy je možné rozlíšiť dva základne druhy kvality:

- kvalitu konceptuálneho modelu – je definovaná sémantickou kvalitou priestorovej databázy;
- kvalitu operačného modelu – je definovaná súborom parametrov kvality priestorových dát ako sú polohová presnosť, tematická presnosť, časová presnosť, úplnosť, logická konzistentnosť a pod.

Posudzovanie kvality údajov v priestorovej databáze môže byť ovplyvnené aj koncepciou tvorby databázy. Medzinárodné štandardy pre geografické informácie (napr. série ISO 19100 a DIGEST) upravujú tvorbu priestorovej databázy – od definície oblasti záujmu, cez návrh špecifikácie a z nej vyplývajúcej množiny parametrov kvality, ich hodnotenie až po dokumentáciu výsledkov prostredníctvom metadát. Pri tvorbe priestorovej databázy, ako aj modelovaní jej kvality je užitočné využiť možnosti jazyka UML (Unified Modeling Language), ktorý je použitý aj vo väčšine štandardov a špecifikácií.

### 3. CIEĽ DIZERTAČNEJ PRÁCE

Dizertačná práca vychádza z požiadavky na efektívnu dokumentáciu priestorových databáz, ktorej kľúčovou časťou sú aj informácie o kvalite. Potreba hodnotenia a dokumentácie kvality priestorových databáz je ovplyvnená aj nasledujúcimi faktormi:

- vzrastajúca výmena a zdieľanie priestorových dát používateľmi aj z iných rezortov ako sú tvorcovia dát
- vzrastajúca dôvera spoločnosti v priestorové dáta
- zabezpečenie prístupu k dátam prostredníctvom ich dokumentácie
- používanie štandardov pri tvorbe priestorových databáz a ich dokumentácii
- snaha predchádzať stratám a zneužitiu dát v dôsledku ich nevhodného použitia.

Cieľom dizertačnej práce je návrh definovania, hodnotenia a spôsobu dokumentácie kvality v priestorových databázach s prihliadnutím na medzinárodné štandardy súvisiace s kvalitou priestorových informácií. Na splnenie cieľa dizertačnej práce bolo navrhnuté riešenie nasledujúcich čiastkových úloh:

- *definovanie parametrov kvality priestorových údajov*

- *definovanie metód hodnotenia kvality priestorových údajov a možných mier výsledkov hodnotenia kvality*
- *popísanie štandardizovaných spôsobov dokumentácie informácií o kvalite priestorových údajov*
- *popísanie princípov kontroly kvality prostredníctvom systému manažérstva kvality podľa série ISO 9000:2000 v organizácii produkujúcej priestorové dáta*
- *demonštrovanie definície, princípy hodnotenia kvality jej dokumentácie vo zvolenej doméne*

## 4. KVALITA ÚDAJOV V PRIESTOROVÝCH DATABÁZACH

Pri hodnotení kvality (aj údajov v priestorovej databáze) hodnotíme rozdiel výsledku a očakávania. Koncepcia hodnotenia kvality vytvára priestor pre producentov aj používateľov priestorovej databázy pre jej ďalšie použitie. Producent hodnotí kvalitu z pohľadu dodržania vopred stanovených kvalitatívnych parametrov v špecifikácii produktu (v tomto prípade produkt = priestorová databáza). Používateľ sa na základe popisu kvality údajov môže rozhodnúť pre použitie údajov priestorovej databázy na svoj konkrétny účel. Súčasná úroveň hodnotenia kvality v organizáciách, ktoré produkujú a poskytujú priestorové dáta je mierne povedané nedostačujúca.

Odborná skupina pre kvalitu (Expert Group on Quality) pri EuroGeographics v roku 1998 urobila prieskum o hodnotení kvality v Národných Mapových Organizáciách (NMO), ktoré sú členmi EuroGeographics (Jakobsson, 2001b). Výsledky ukazujú, že až 71% NMO neposkytuje žiadne informácie o kvalite. Aj keď bol prieskum uskutočnený v roku 1998, z náhodných prieskumov ExG-Q je zjavné, že situácia sa mení iba veľmi pomaly. Medzi faktory, ktoré ovplyvňujú tento stav patria nepochybne aj nedostatočné znalosti o kvalite priestorových údajov – princípov, parametrov kvality a spôsobov ich hodnotenia, dokumentácii výsledkov a kontrole kvality.

### 4.1 PARAMETRE KVALITY PRIESTOROVÝCH ÚDAJOV

Priestorovú databázu môžeme z pohľadu jej štruktúry chápať ako bázu menších súborov dát, ktoré majú zhodné triedy, atribúty alebo vzťahy. Podľa spoločných ukazovateľov môžeme posúdiť jednotlivo kvalitu týchto súborov dát. Výslednú kvalitu produktu potom určíme ako sumu kvalít čiastkových súborov tvoriacich databázu.. Rôzne zdroje (SDTS, DIGEST, ISO, CEN, ICA) definujú množinu parametrov kvality. Tieto je možné rozdeliť na dve základné skupiny – kvantitatívne a nekvantitatívne (známe aj ako opisné – napr. ISO) parametre kvality priestorových údajov.

#### **Kvantitatívne parametre kvality:**

- **polohová presnosť** – možno podľa (Fischer, 1999) a (Veregin, 1999) definovať ako očakávaný rozdiel medzi polohou objektu podľa geoúdajov v údajovom súbore a jeho skutočnou pozíciou, tvarom a veľkosťou v realite. Za skutočné hodnoty týchto parametrov považujeme také, ktoré sú určené presnejšou metódou (napr. geodetickou – merania DGPS apod.). Presnosť polohy čiarových a plošných objektov nie je doteraz jednoznačne určiteľná. Preto sa poloha týchto objektov nahradzuje oblasťou neurčitosti, tzv.  $\epsilon$ -pásom – „fuzzy pásom“, ktorý je definovaný v okolí objektu (Mitášová a kol., 2001),

- **časová presnosť** - hovorí o zhode medzi zakódovanou a súčasnou časovou súradnicou,
  - **tematická presnosť** – vyjadruje množstvo správne alebo nesprávne určených popisných údajov a závisí od formy vyjadrenia atribútov diskrétno alebo spojito. Ohodnotenie presnosti atribútov vyjadruje percentuálny obsah správne alebo nesprávne určených popisných údajov a môžeme ho realizovať na základe kontrolných meraní v teréne, resp. s využitím ortofotosnímkov,
  - **sémantická presnosť** – vyjadruje kvalitu definície objektov vzhľadom na špecifikáciu oblasti záujmu. Ja vyjadrená súhrnom parametrov ako sú sémantická úplnosť, sémantická konzistencia, atribútová správnosť, aktuálnosť,
  - **úplnosť** - porovnáva objekty v báze údajov a jej špecifikáciou, pričom je potrebné si podľa (*Veregin, 1999*) uvedomiť existenciu dvoch typov úplnosti:
    - o *údajovej úplnosti*, ktorá je merateľnou odchýlkou zanedbania zistenou medzi bázou a jej špecifikáciou a
    - o *úplnosti modelu*, ktorá hovorí o zhode medzi špecifikáciou bázy a realitou.
- Prakticky je nemožné vyjadriť, či sa v báze údajov nachádzajú všetky informácie o všetkých geobjektoch v záujmovom území, preto má zmysel hovoriť aj o treťom type úplnosti – o úplnosti kontroly, ktorá predstavuje množstvo a rozloženie kontrolných meraní,
- **logická konzistencia** – určuje stupeň dodržania logických vzťahov prvkov v báze. Pri posudzovaní kvality logických vzťahov v databáze hodnotíme kvalitu geometrických vzťahov (topológie) a kvalitu vzťahov medzi geometrickými a negeometrickými (atribútovými) dátami,
  - **správnosť** – tento parameter vyjadruje správnosť reprezentácie reality priestorovou databázou. Reprezentácia reality je správna vtedy, ak výsledky operácií vykonaných v realite korešpondujú s výsledkami zodpovedajúcich operácií vykonaných v databáze.

#### Nekvantitatívne parametre kvality:

- **účel** – zámer používania bázy údajov,
- **použitelnosť** – poskytuje informácie o použití priestorových dát v rôznych aplikáciách. Jeho definícia je špecifikovaná okrem ISO 19113 aj v norme ISO 9242-11, ktorá odporúča definovať (systémovú použitelnosť) rozsah, v ktorom je možné používať priestorovú databázu pri zachovaní stanovenej kvality v jej špecifikácii. Tento parameter kvality je predmetom najnovších diskusií v komunite GI – jeho definícia nie je doteraz vymedzená.
- **pôvod** – popisuje zdroje a metódy zberu a tvorby priestorovej databázy.
- **rozlišiteľnosť** – udáva hustotu údajov v záujmovej lokalite. Môže byť geometrická, tematická (obsahová) alebo časová.
- **homogenita** – je to textový kvalitatívny popis očakávanej alebo overenej jednotnosti informácií o kvalite. Homogenita hovorí o meta-kvalite priestorovej databázy.
- **záväznosť** – informuje používateľa o tom, kedy bola databáza vytvorená. Tento parameter má význam v databázach, v ktorých sa hodnota údajov sa často mení.
- **dostupnosť** – informuje používateľa o autorských právach a o dostupnosti dát.
- **bezpečnosť** – informuje používateľa o bezpečnosti a ochrane dát (nechránené dáta sú pre používateľa nedôveryhodné a nespoľahlivé)



## 4.2 METÓDY HODNOTENIA KVALITY ÚDAJOV V PRIESTOROVÝCH DATABÁZACH

Metódy hodnotenia kvality údajov v priestorových databázach môžeme podľa STN EN ISO 19114:2005 Geografická informácia – Metódy hodnotenia kvality rozdeliť na:

- **priame** (automatické alebo manuálne) – založené na hodnotení prvkov v databáze. Priame metódy môžu byť:
  - o *interné* – napr. kontrola topologických vzťahov geometrických údajov,
  - o *externé* – napr. kontrola úplnosti geografických názvov, ktorá vyžaduje porovnanie tabuľky názvov v geografickej databáze s katalógom geografických názvov.
- **nepriame** – založené na hodnotení prvkov v databáze, ktoré je založené na externých znalostiach. Nepriamymi metódami hodnotíme väčšinou opisné parametre kvality priestorových dát.

Proces hodnotenia kvality priestorových dát môže byť statický alebo dynamický (aplikovateľný napr. v on-line katastrálnych databázach). Pri statickom procese hodnotenia kvality používame vyššie vymenované metódy hodnotenia kvality. Pri dynamickom procese používame nasledovné metódy hodnotenia:

- „**benchmark**“ – táto metóda je založená na zvolení frekvencie hodnotenia kvality, následne sa vyhotoví kópia databázy k zvolenej epoche a táto sa hodnotí ako pri statickom procese hodnotenia kvality.
- **priebežná** metóda – založená na hodnotení aktualizácií a ich dopadu na databázu (princíp tejto metódy vychádza z procesne orientovaného prístupu v ISO 9000)

Pri hodnotení kvality je kľúčovou voľba štatisticky významnej vzorky dát. Tento proces upravujú medzinárodné štandardy (ISO 2859 a ISO 3951), ktorých použitie odporúča aj štandard ISO 19114. Miery výsledkov hodnotenia kvality priestorových údajov bude upravovať špecifikácia ISO 19138.

## 4.3 DOKUMENTÁCIA VÝSLEDKOV HODNOTENIA KVALITY PRIESTOROVÝCH ÚDAJOV

Výsledky hodnotenia kvality je možné publikovať dvojakým spôsobom:

**súbor metadát** – tento spôsob upravuje (ISO 19113) a (ISO 19115),

**hodnotiaci správa** – obsahuje podrobnejší popis procesu hodnotenia kvality ako súbor metadát (ISO 19114).

Pre umožnenie ľahkej orientácie v záplave geoúdajov, geoinformácií a z nich vznikajúcich „geoznalostí“ v zmysle (Ivánová, 2002) je nutné vytvoriť akýsi návod na ich čítanie a chápanie. Túto funkciu zabezpečujú metadátové katalógy. Medzinárodná norma ISO 19115 definuje prvky, poskytuje schémy a zavádza terminológiu metadát. Úlohou metadát je popísať priestorovú databázu, vrátane úplnosti dátových zdrojov, vykonanej kontroly meraní a pod. Pre popis štruktúry metadátového katalógu využíva štandard ISO 19115 notáciu jazyka UML s použitím definovaného rozhrania a konceptuálnej schémy jazyka UML (ISO 19103).

Súčasnú aktivitu vo svete, ale aj na Slovensku ukazujú trend využívať pre metadáta prostredie Internetu a týmto umožniť používateľom prístup k metadátam. Čo sa týka metadát o kvalite, ukazuje sa, že z pohľadu používateľov sú zrozumiteľnejšie a prístupnejšie vhodne zvolené metadáta v číselnom formáte ako záplava textu.

## 4.4 ŠTANDARDIZÁCIA KVALITY PRIESTOROVÝCH ÚDAJOV

Vhodne navrhnutý súbor používaných štandardov pomáha zabezpečiť interoperabilitu, prenosnosť, udržovateľnosť a kvalitu priestorových dát. V organizácii produkujúcej priestorové dáta ide o množinu štandardov o:

- hardvéri a sieťach
- administrácii systému
- softvérových aplikáciách
- dátových formátoch
- o zbere, spracovaní, správe, údržbe a aktualizácii dát
- reprezentácii a vizualizácii priestorových dát
- prístupe k dátam a ich distribúci

Tvorbe štandardov pre priestorové informácie sa v súčasnosti vo svete venujú viaceré organizácie. Pre organizácie produkujúce priestorové dáta na Slovensku sú relevantné najmä:

- ISO/TC 211 – Geographic information/Geomatics – technická komisia, ktorá je tvorcom štandardov série ISO 19100 pre geografické informácie
- CEN/TC 287 – Geographic information – technická komisia, ktorá tvorí (momentálne najviac spôsobom prevzatia štandardov série ISO 19100 do sústavy noriem EN) normy pre potreby Európskej únie.
- TK89 „Geodézia a kartografia“ pri SÚTN – technická komisia, ktorá koordinuje aktivity v oblasti priestorových informácií na Slovensku.

Zoznam existujúcich a pripravovaných štandardov, ktoré priamo súvisia s kvalitou priestorových a sú relevantné pre použitie na Slovensku sú vymenované v tabuľke 1.

**Tab. 1 Zoznam štandardov série ISO 19100 priamo súvisiacich s kvalitou priestorových dát**

International Standards and Technical Reports	
ISO 19113:2002	Geographic information — Quality principles
ISO 19114:2003	Geographic information — Quality evaluation procedures
ISO 19115:2003	Geographic information — Metadata
FDTS 19138	Geographic information — Data quality measures
FDTS 19139	Geographic information — Metadata – XML schema implementation

## 4.5 KONTROLA KVALITY ÚDAJOV V PRIESTOROVEJ DATABÁZE

Pre zabezpečenie kvality priestorových dát je nevyhnutné zabezpečiť kontrolu kvality počas celého procesu tvorby priestorovej databázy. Systémy manažérstva kvality, ktoré sa v súčasnosti stávajú pre producentov v akejkoľvek oblasti ľudskej činnosti nevyhnutnosťou, pomáhajú udržať deklarovanú úroveň kvality počas celého životného cyklu organizácie. Jedným z najrozšírenejších systémov manažérstva kvality je systém ISO 9000, ktorý upravuje súbor štandardov ISO 9000:2000. Proces implementácie tohto systému prebieha v nasledujúcich etapách:

- identifikácia očakávaní organizácie od zákazníkov, zamestnancov, dodávateľov, akcionárov a spoločnosti
- stanovenie procesov nevyhnutných pre dodávku produktu zákazníkom

- plánovanie a realizácia činností
- kontrola, posúdenie zhody, certifikácia

Vzťah medzi sériou štandardov ISO 19100 a ISO 9000:2000 explicitne nie je definovaný. Ja však možné tvrdiť, že dokumentácia kvality v zmysle štandardov série ISO 19100 je súčasťou dokumentácie procesov v organizácii podľa systému manažérstva kvality ISO 9000:2000.

## 5. APLIKÁCIA

Vzhľadom k robustnosti témy dizertačnej práce je aplikácia teórie princípov hodnotenia kvality rozdelená na dve časti:

- analýza kvality priestorových údajov vo zvolenej priestorovej doméne a vytvorenie modelu kvality priestorových údajov
- návrh metodiky hodnotenia kvality priestorovej databázy počas celého procesu jej tvorby.

Zvolený postup vyplynul z nasledujúcich faktov:

- modelovanie priestorovej domény je dlhodobý proces, ktorý presahuje tematický rámec dizertačnej práce a preto bol vybudovaný iba model kvality, ktorý dopĺňa model priestorovej domény – pre potreby dizertačnej práce bola zvolená doména katastra nehnuteľností Slovenskej republiky,
- návrh metodiky hodnotenia kvality pre vznikajúcu priestorovú databázu (konkrétne ZB GIS) bol výsledkom viacerých faktorov:
  - o ZB GIS je stále vo vývoji, čo znamená, že je relevantná diskusia o princípoch analýzy, modelovania a hodnotenia kvality tejto priestorovej databázy na základe požiadaviek tvorcov databázy
  - o ZB GIS tvorí štátna organizácia, čím je zabezpečený prístup k dokumentom a používaným postupom v ZB GIS relevantným pre tému dizertačnej práce.
  - o navrhnutá metodika je výsledkom úzkej spolupráce tvorcov ZB GIS a dizertantky

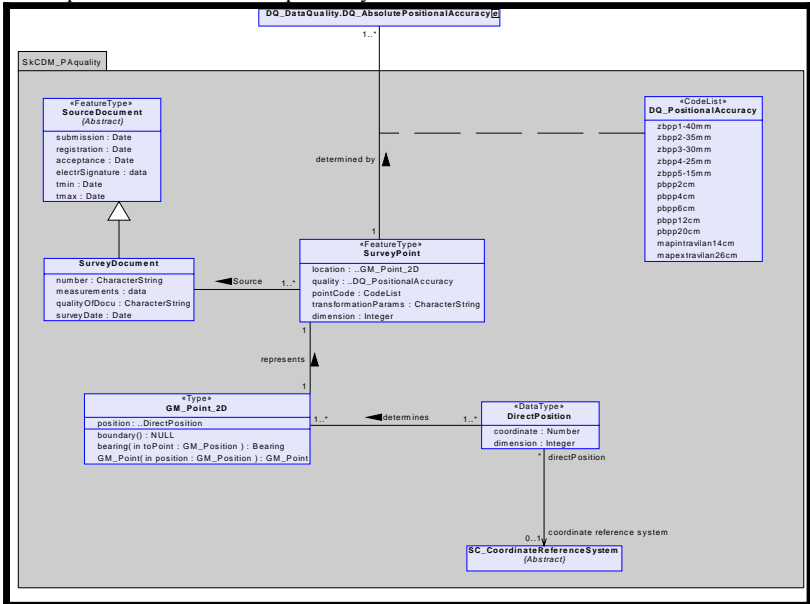
### 5.1 MODEL KVALITY V DOMÉNE KATASTRA NEHNUTEĽNOSTÍ

Model priestorovej domény katastra nehnuteľností (Stromček, 2005) je založený na princípoch Základného modelu domény katastra (Core Cadastral Domain Model – CCDM) navrhnutej výskumným tímom okolo prof. Oosteroma z TU Delft v Holandsku (Oosterom, 2004). CCDM je vytvorený v zhode s medzinárodnými štandardami série ISO 19100 a špecifikáciami OGC. Aplikácia v lokálnej (slovenskej) katastrálnej doméne taktiež rešpektuje legálne obmedzenia – zákony č.162/1995 Z.z. a č. 215/1995 Z.z. a vyhlášky č. 79/1996 Z.z. a 178/1996 Z.z.

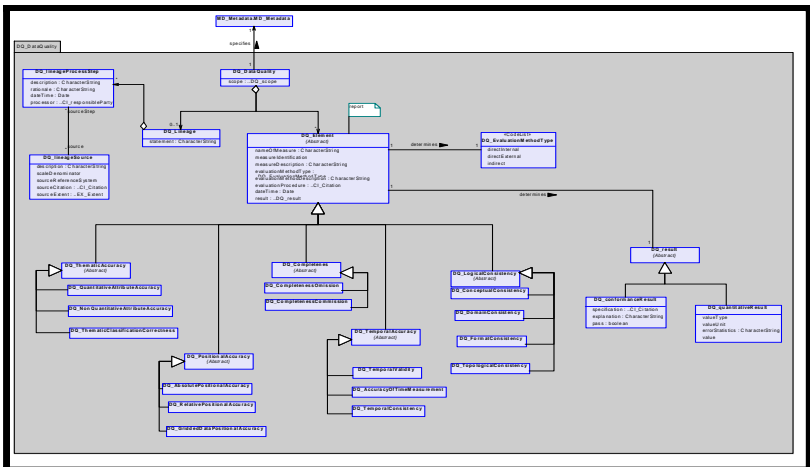
Model kvality lokálnej katastrálnej domény je vytvorený v zhode so sériou štandardov ISO 19100 a opiera sa o legálne obmedzenia, ktoré je v kontexte dizertačnej práce možné považovať za jedinú špecifikáciu pre zvolenú doménu.

Po analýze parametrov kvality v zmysle definícií uvedených v dizertačnej práci bol vytvorený model kvality s väzbou na model metadát domény katastra nehnuteľností v SR. Podľa špecifikácie domény je jediným parametrom kvality údajov, ktorý sa v doméne katastra nehnuteľností Slovenskej republiky eviduje polohová presnosť.

Tento model bol vytvorený s využitím možností jazyka UML pomocou softvéru Select Component Architecture zapožičaným firmou LBMS.



Obr. 3 Model kvality priestorových dát v doméne KN



Obr. 4 Metadáta o kvalite priestorových údajov v doméne KN

## 5.2 KVALITA ÚDAJOV V ZÁKLADNEJ BÁZE PRE GEOGRAFICKÝ INFORMAČNÝ SYSTÉM V SLOVENSKEJ REPUBLIKE

V dizertačnej práci je navrhnutá metodika tvorby priestorovej databázy v zmysle analýzy a hodnotenia kvality údajov vo zvolenej priestorovej databáze. Bola zvolená Základná báza pre geografický informačný systém v Slovenskej republike (ZB GIS). ZB GIS je databáza geografických objektov a je časťou Štátneho informačného systému, ktorý je súčasťou Národnej infraštruktúry priestorových informácií. ZB GIS sa tvorí podľa Konceptie tvorby, aktualizácie a správy Základnej bázy geografického informačného systému do roku 2005 (ďalej iba Konceptia – ÚGKK, 2002). Oblasť záujmu ZB GIS je vymedzená Katalógom objektov (KO 10/2004 – ÚGKK, 2004) vytvoreného podľa medzinárodného kódovacieho systému DIGEST/FACC (DGIWG, 2000).

Metodika pre aplikáciu princípov analýzy a hodnotenia kvality priestorových dát v ZB GIS bola navrhnutá v zmysle série medzinárodných štandardov série ISO 19100 a princípov kontroly kvality definovaných v sérii ISO 9000:

### 1. Účel tvorby priestorovej databázy založený na potrebách používateľov

Zámerom tvorcov ZB GIS je vybudovať základnú bázu referenčných objektov, ktoré sú vzťahované k povrchu Zeme na území Slovenskej republiky (SR), ktorá má slúžiť pre tvorbu tematických informačných systémov v SR.

### 2. Abstrakcia objektov reálneho sveta do oblasti záujmu (OZ)

Ontológiu ZB GIS tvorí KO 10/2004, ktorý špecifikuje obsah ZB GIS a navrhnutá metodika zberu a konverzie dát do priestorovej databázy. Je na mieste poznamenať, že, ZB GIS sa začala naplňať skôr ako bola vytvorená jej ontológia, čo sa prieči definíciou ontológie.

### 3. Vytvorenie špecifikácie produktu – ZB GIS

Špecifikácia ZB GIS je v zmysle série štandardov ISO 19100 nedostatočná – nie sú jasne definované parametre kvality a postupy ich hodnotenia a dokumentácie. Jediným parametrom kvality, ktorý definuje Konceptia je tzv. „absolútna polohová presnosť do 0,5m“. V čase dokončovania dizertačnej práce stále nebolo jasné, čo táto definícia v praxi znamená – ide o polohovú presnosť súradníc bodov objektov uložených v priestorovej databáze alebo o presnosť bodov aerotriangulácie?

### 4. Vytvorenie modelu kvality ZB GIS

Pokiaľ producent databázy nedefinuje aké parametre kvality sú relevantné pre jeho účel, nie je možné vytvoriť model kvality údajov. Pokiaľ je špecifikácia parametrov kvality v zmysle „urobme ZB GIS čo najlepšie“ je takmer nemožné kvalitu hodnotiť.

### 5. Tvorba priestorovej databázy ZB GIS – naplňanie databázy

ZB GIS je stále vo vývoji a obsah databázy sa priebežne tvorí podľa plánov producenta. Čaká sa na novú Konceptiu, ktorá by mala ujasniť, pravidlá naplňanie ZB GIS v zmysle zmenených postupov, technológií a štandardov uvádzaných v pôvodnej Konceptii.

### 6. Hodnotenie kvality ZB GIS

Ako už vyplýva z bodu č.4 hodnotenie parametrov kvality ZB GIS sa doteraz nevykonáva. Návrh metód hodnotenia parametrov kvality v zmysle série ISO 19100 je rozsiahle rozpracovaný v dizertačnej práci.

### 7. Dokumentácia výsledkov hodnotenia kvality ZB GIS

Zámerom tvorcu je dokumentovať ZB GIS prostredníctvom metadát v zmysle normy STN EN ISO 19115:2005 Geografická informácia – Metadáta, ktorá s nadväznosťou na

normy STN EN ISO 19113 a STN EN ISO 19114 upravuje aj dokumentáciu výsledkov hodnotenia kvality (spôsobom podobným ako v aplikácii pre katastrálnu doménu – pozri prílohu 3 v dizertačnej práci).

## **8. Kontrola kvality procesu tvorby ZB GIS**

Tvorca ZB GIS nepoužíva žiaden štandardizovaný systém manažérstva kvality. Pravidlá hodnotenia kvality procesov v organizácii sledujú postupy kontroly produkčného postupu tvorby kartografických diel, čo však pre potreby tvorby ZB GIS nie je dostačujúce. V ZB GIS je kontrola kvality procesov zabezpečovaná na úrovni možností softvérov použitých pre zber a spracovanie dát.

## **6. PRÍNOS PRE VEDNÚ DISCIPLÍNU A PRAX**

Pre potreby priestorových informačných infraštruktúr na všetkých úrovniach je dôležité definovať, hodnotiť a dokumentovať kvalitu priestorových informácií. Úlohou odborníka na kvalitu, je skôr než vyrábanie vzorcov a tabuliek vysvetľovať, čo kvalita znamená v príslušnej organizácii. Je takmer nemožné hovoriť o kvalite všeobecne. Téma dizertačnej práce je rozsiahla a vyžaduje viac času, strán a možno aj hláv. Na začiatku výskumu v r. 2000 bolo viac či menej známych 6, pričom koncom roka 2005 je v tejto dizertačnej práci zadaných 16 (!) parametrov kvality priestorových údajov.

Dizertačná práca sa zaoberá definíciou kvality z pohľadu celého procesu tvorby priestorovej databázy – od abstrakcie a výberu objektov reálneho sveta do oblasti záujmu, cez definovanie parametrov kvality, metód ich hodnotenia až po spôsoby ich dokumentácie. Z teórie rozpracovanej v dizertačnej práci vyplývajú nasledovné požiadavky na aplikáciu princípov hodnotenia kvality priestorových údajov:

- nevyhnutnosť nazeráť na priestorovú databázu z pohľadu princípov hodnotenia kvality počas celého procesu jej tvorby z dôvodu stanovenia relevantných parametrov kvality priestorovej databázy
- zdôraznenie potreby dôkladného spracovania špecifikácie produktu pri tvorbe priestorovej databázy, ktorá umožní jednoznačne odpovedať okrem iných aj na tieto otázky:
  - o kto sú používatelia priestorovej databázy?
  - o ktoré parametre kvality sú relevantné pre používateľov priestorovej databázy?
  - o ktoré parametre kvality sú relevantné pre tvorcov priestorovej databázy?
  - o ktoré parametre kvality je nevyhnutné dokumentovať?

Na záver možno konštatovať nasledovné výsledky dizertačnej práce:

- definované vybrané parametre kvality priestorových údajov v zmysle medzinárodných štandardov série ISO 19100 pre geografické informácie, ktoré sú analyzované v doméne katastra nehnuteľností v Slovenskej republike (kap. 8.1 a príloha 3 v dizertačnej práci)
- navrhnutá metodika tvorby priestorovej databázy v kontexte definície a princípov hodnotenia kvality priestorových údajov a v zmysle medzinárodných štandardov série ISO 19100 pre geografické informácie.

*PRE ROZVOJ VEDY V OBLASTI GEODÉZIE, GEOINFORMATIKY  
A GEODETICKEJ KARTOGRAFIE:*

- *spracovanie prvého súvislého textu na tému kvality v priestorových databázach v lokálnej (slovenskej) komunite pracujúcej s priestorovými informáciami*
- *navrhnutie definície parametrov kvality priestorových údajov vybraných v súlade s medzinárodnými štandardami série ISO 19100 pre geografické informácie*
- *vypracovanie teórie metód hodnotenia vybraných parametrov kvality priestorových údajov*
- *navrhnutie zoznamu mier, ktoré môžu vybrané parametre kvality údajov nadobúdať*

*PRE GEODETICKÚ, GEOINFORMAČNÚ A KARTOGRAFICKÚ PRAX:*

- *aplikovanie princípov definície kvality priestorových údajov v doméne katastra nehnuteľností v Slovenskej republike*
- *aplikácia je príspevkom k výskumu viacúčelového katastra*
- *návrh súčasných metód hodnotenia vybraných parametrov kvality – parametre boli vybrané v kontexte zámeru tvorcov ZB GIS*
- *navrhnutá metodika tvorby priestorovej databázy v zhode s medzinárodnými štandardmi série ISO 19100 pre geografické informácie.*

## **7. PUBLIKOVANÁ LITERATÚRA DOKTORANDKY**

- [1] Mitášová, I., Ivanová, I., Chalachanová, J. (2001): Kvality údajov v geoinformačných bázach, In: Zborník referátov sympózia „Úlohy geodézie a kartografie pri tvorbe a správe ZB GIS“, Slovenský zväz geodetov.
- [2] Ivánová, I. (2002): Štandardizácia kvality údajov v procese integrácie geoinformačných databáz, In: Sborník 4. odborné konferencie doktorského studia, Brno, ČR, 2002
- [3] Ivánová, I. (2002): Proces hodnotenia kvality z pohľadu producentov aj používateľov geografických údajov, Pedagogické listy 9/2002 – Geoinformačný model a jeho kartografická podpora, KMPU, SvF STU, Bratislava, ISBN 80-227-1828-9
- [4] Ivánová, I., Chalachanová, J. (2002): Problematika kvality v štandardizácii geografickej informácie, Geografická informácia – Terminológia v normách ISO a CEN, Geografický ústav SAV, Bratislava, ISBN 80-968365-5-2
- [5] Rapant, P., Ivánová, I. (2002): Standardizace v geografických informačních systémech v evropském kontextu, Vojenský informačný systém a prax – Zborník referátov, Banská Bystrica.

- [6] Chalachanová, J., Ivánová, I. (2002): Model hodnotenia kvality priestorovej bázy údajov – Aktivity v kartografii 2002 – zborník referátov zo seminára, Kartografická spoločnosť SR a Geografický ústav SAV, Bratislava.
- [7] Ivánová, I., Chalachanová, J. (2003): Kvalita údajov– princípy, metódy hodnotenia a dokumentácia, GaKO 7-8/2003, ČÚZK a UGKK SR, Praha, ČR
- [8] Ivánová, I., Chalachanová, J., Valachovičová, L., Fencík, R (2003): Príspevok k vybraným termínom z oblasti geoinformatiky – In: Zborník k semináru Geografická informácia – terminológia v normách ISO (2002 –2003), 12.11.2003, Geografický ústav SAV, Bratislava
- [9] Ivánová, I. (2004): 18. plenárne zasadnutie CEN/TC 287 Geografická informácia, Normalizácia 6/2004, Úrad pre normalizáciu metrologiu a skúšobníctvo SR, Bratislava
- [10] Ivánová, I. (2005): Standardization of geographic information, Kartografické listy 13/2005, Kartografická spoločnosť SR a Geografický ústav SAV, Bratislava, ISBN 80-89060-07-2, ISSN 1336-5274
- [11] Ivánová, I., Vajsáblová, M. (2005): Uncertainty in Position of Objects as an Influence Factor of the Area Value, */dexa/*, pp. 468-472, 16th International Workshop on Database and Expert Systems Applications (DEXA'05)
- [12] Ivánová, I. (2005): Sémantická presnosť – zanedbávaný parameter kvality priestorových údajov, In: Zborník z konferencie „Geodetické siete a priestorové informácie“ konanej v dňoch 24.-26.10.2005 v Podbanskom

## 8. POUŽITÁ LITERATÚRA

- [1] Aalders, H.J.G.L. (2000): The universe of discourse in GIS. In Quality of Geo-Information Course Material, GIS-section, Faculty of Geodesy, 2002/2003, TU Delft, The Netherlands
- [2] Aalders, H.J.G.L. (2002a): The Registration of Quality in a GIS, In Spatial Data Quality, edited by W. Shi, P.F. Fisher, and M.F. Goodchild, Taylor&Francis, London, UK, ISBN 0-415-25835-9, p. 186-200
- [3] Aalders, H.J.G.L. (2002b): The Standardisation of Object Definition: Internationally and The Dutch Case. In There is more than geometry, p. 1-15, VBK Editorial Management b.v. Liemeer, NL, ISBN 90-806917-1-2
- [4] Aalders, H.J.G.L. (2003): Study material from the ‘Geo-ICT Course’, Dept. Theoretical Geodesy – 13.-17.10.2003, Bratislava, SR
- [5] Aalders, H.J.G.L. (2004): 2004, the beginning of a new era for European standards, in Proceedings of a symposium GIS Ostrava 2004, [http://gis.vsb.cz/GIS\\_Ostrava/GIS\\_Ova\\_2004/Sbornik/Referaty/default.htm](http://gis.vsb.cz/GIS_Ostrava/GIS_Ova_2004/Sbornik/Referaty/default.htm) (23.1.2006)
- [6] Alesheikh et al (1999): Rigorous geospatial data uncertainty models for GISs, In Spatial accuracy assessment: land information uncertainty in natural resources, Editors: Lowell, K. and Jaton, A., Ann Arbor Press, Michigan, p. 195-202
- [7] Bossler, J., et al (2002): Manual of Geospatial Science and Technology, Taylor & Francis Ltd, London, ISBN 0-7484-0924-6
- [8] CERCO WG on Quality (1999): Good Reasons for Implementing a Quality Management System in National Mapping Agencies, <http://www.eurogeographics.org/eng/GoodReasons.html> (1.2.2006)



- [9] CERCO WG on Quality (2000): Handbook for implementing a quality in management system in National Mapping Agency, [http://www.eurogeographics.org/eng/documents/handbook\\_V1.pdf](http://www.eurogeographics.org/eng/documents/handbook_V1.pdf) (1.2.2006)
- [10] Coad, P., Yourdon, E (1990): Object oriented analysis, Yourdon Press Computing Series, New Jersey, ISBN 0-13-629122-8
- [11] Dasonville, L. et al (2002): Quality Management, Data Quality and Users, Metadata for Geographical Information. In Spatial Data Quality, edited by W. Shi, P.F. Fisher, and M.F. Goodchild, Taylor&Francis, London, UK, ISBN 0-415-25835-9, p. 202-215
- [12] Doucette, M., Paresi, Ch., (2000): Quality management in geospatial data infrastructure (GDI) – political background of quality management. In Groot, R., McLaughlin, J.: GDI – concepts, case & good practice, Oxford, Oxford University Press, 286 p., ISBN 0-19-823381-7
- [13] English, L. (2005): The High Cost of Low-Quality Data, In DM Review Magazine, January/1998, ([http://www.dmreview.com/article\\_sub.cfm?articleId=771](http://www.dmreview.com/article_sub.cfm?articleId=771)) - 24.1.2006
- [14] Frank A. (2005): Ontology for GIS (Draft – v4), <http://www.geoinfo.tuwien.ac.at/index.php> - 1.6.2005
- [15] Goodchild, M. (2003): Models of Uncertainty in Area-Class Maps, In Proceedings of The 2<sup>nd</sup> International Symposium on SDQ'03, editors: Shi, W., et al, ARC SIT, dpt. Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, p.1-9, ISBN 962-367-374-4
- [16] Goodchild, M. (2005) – A general framework for error analysis in measurement-based GIS, In Journal of Geographical Systems, Volume 6, Number 4; Springer-Verlag, ISSN 1435-5949
- [17] Gruber, T., (1993), What is an Ontology? Knowledge Systems Laboratory, Computer Systems Dept., Stanford University, Stanford, CA94305, USA
- [18] GSDI: Developing Spatial Data Infrastructures: The SDI Cookbook v.2.0, 25.January 2004, Editor: Douglas D. Nebert, Technical Working Group Chair, GSDI <http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf> (1.2.2006)
- [19] Guptill, S., Morrison, J., (1995): Elements of Spatial Data Quality, BPC Wheatons Ltd., Exeter, ISBN 008042432 p.197
- [20] Hunter, G.J., Qiu, J. (2003a): Automatic Updating of Spatial Data Quality Information, In Proceedings of The 2<sup>nd</sup> International Symposium on SDQ'03, editors: Shi, W., et al, ARC SIT, dpt. Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, p.210-214, ISBN 962-367-374-4
- [21] Hunter, G.J., Wachowicz, M. and Bregt, A.K (2003b): Understanding Spatial Data Usability”, Data Science Journal, Volume 2, p. 79-89 (e-journal located at [www.datasciencejournal.org](http://www.datasciencejournal.org)), article located at: [http://www.sli.unimelb.edu.au/hunter/2003\\_Data\\_Science\\_paper.pdf](http://www.sli.unimelb.edu.au/hunter/2003_Data_Science_paper.pdf) (3.2.2006)
- [22] Hunter, G.J. (1999): New Tools for Handling Spatial Data Quality: Moving from Academic Concepts to Practical Reality. Journal of the Urban and Regional Information Systems Association, 11, 2, p. 25-34, URISA, USA, ISSN 1045-8077 [http://www.sli.unimelb.edu.au/hunter/1999\\_J\\_URISA\\_11\(2\)\\_paper.pdf](http://www.sli.unimelb.edu.au/hunter/1999_J_URISA_11(2)_paper.pdf) (1.2.2006)
- [23] Jakobsson, A., Salo-Merta, L.: (2001a) Definition of a Basic Topographic Framework for National GI Policy – One Database for All Basic Topographic Data. In Proceeding of the 20th International Cartographic Conference, Volume 4; p.: 2197-2205

- [24] Jakobsson, A., Vaughlin, F. (2001b): Status of Data Quality in European National Mapping Agencies, [www.eurogeographics.org/eng/documents/Dataquality\\_ICC.doc](http://www.eurogeographics.org/eng/documents/Dataquality_ICC.doc) (1.2.2006)
- [25] Jakobsson, A. (2002) Data Quality and Quality Management – Examples of Quality Evaluation Procedures and Quality Management in European National Mapping Agencies, In Spatial Data Quality, edited by W. Shi, P.F. Fisher, and M.F. Goodchild, Taylor&Francis, London, UK, ISBN 0-415-25835-9, p.216-229.
- [26] Kanisová, H., Müller, P. (2004): UMLsrozumitelně, 158p. Computer Press, Brno, 2004, ISBN 80-251-0231-9
- [27] Kimlička, Š. (2004): Príklady citovania podľa noriem ISO 690 a ISO 690-2, Katedra knižničnej a informačnej vedy FiFUK, Bratislava, <http://www.fphil.uniba.sk/~kkvi/citovanie-KKIV.pdf> (3.2.2006)
- [28] Kresse, W., Fadaie, K.,(2004): ISO Standards for Geographic Information, 322 p., Springer – Verlag Berlin Heidelberg, 2004, ISBN: 3-540-20130-0
- [29] Lemmens, M.J.P.M (2003): GIS: The Data Problem, In Quality of Geo-Information Course Material, GIS-section, Faculty of Geodesy, 2002/2003, TU Delft, The Netherlands
- [30] Leung, Y., Yan, J.P. (1998): A locational error model for spatial features, In International Journal of Geographical Information Science, Volume 12, Number 6, p. 607-620, ISSN
- [31] Longley, P.A., et al, (1999): Geographical Information Systems. Vol.1, 2 Second Edition, New York: John Wiley and Sons,
- [32] Mark, D. M. (1993): Human spatial cognition. In Medyckyj-Scott, D., and Hearnshaw, H. M., editors, Human Factors in Geographical Information Systems, Belhaven Press, p.51-60, <http://www.geog.buffalo.edu/~dmark/DMScottchapter.html> (30.1.2006)
- [33] Mars, N.J.I. (1995): What is an ontology? In Proceedings Seminar on the impact of ontologies on reuse, interoperability and distributed processing (editor: A. Goodall), p. 9-19, London, November 7, 1995, Unicom, Uxbridge, Middlesex, UK
- [34] Mayberry, M. (2002): Data Quality: Before The Map is Produced, [http://www.directionsmag.com/article.php?article\\_id=250](http://www.directionsmag.com/article.php?article_id=250) (30.1.2006)
- [35] Meško, D., Katusčák, D. a kol. (2004): Akademická príručka, Vydavateľstvo Osveta, spol. s r.o., 316p., ISBN 80-8063-150-6
- [36] Moellering, H., Aalders, H.J.G.L, Crane, A. (2005) – World Spatial Metadata Standards, © ICA, Elsevier Science/Pergamon USA, ISBN - 0080439497
- [37] Molenaar, M. (2003a): Spatial Data Modelling and The Specification of Semantics and its Uncertainty Aspects, In Proceedings of The 2<sup>nd</sup> International Symposium of SDQ'03, editors: Shi, W., et al, ARC SIT, dept. Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, ISBN 962-367-374-4
- [38] Molenaar, M. (2003b) Spatial data modelling and the specification of semantics and its uncertainty aspects: In Proceedings of the 2nd international symposium on spatial data quality: 19 - 20 March 2003, Hong Kong / ed. by Shi, W., Goodchild, M.F., Fisher, P.F – Hong Kong, The Hong Kong Polytechnic University. p. 10-16
- [39] Molenaar, M (1994): A Syntactic Approach for Handling the Semantics of Fuzzy Spatial Objects. Advanced Geographic Data Modelling 1994 (AGDM'94), Spatial Data Modelling and Query Languages for 2D and 3D Applications, Proceedings, Delft, The Netherlands, 12-14 September 1994. Published by the Netherlands Geodetic Commission, Publications on Geodesy No. 40, p. 155-

- [41] Nenadál, J. (2001): *Meření v systémech managementu jakosti*, Management Press, Praha, ISBN 80-7261-052-6
- [42] Page-Jones, M. (2001): Page-Jones, M.: *Základy objektově orientovaného návrhu v UML*, 368p. Grada, 2001, ISBN: 80-247-0210-X.
- [43] Qiu, J. and Hunter, G.J. (2003): A GIS with the Capacity for Managing Data Quality Information, In *Spatial Data Quality*, edited by W. Shi, P.F. Fisher, and M.F. Goodchild, Taylor&Francis, London, UK, ISBN 0-415-25835-9, p. 230-250
- [44] Quinne, W. V.(1977): *Ontological Relativity and Other Essays*, p.165, Columbia University Press, USA, ISBN 0231083572
- [45] Rektorys, K. a kol. (1973): *Přehled užití matematiky*, SUTL – Nakladatelství technické literatury, Praha, ČR, ISBN 04-001-73, 1136 p.
- [46] Salgé, F., (1995): *Semantic Accuracy*, in *Elements of Spatial Data Quality*, Elsevier Science Ltd Oxford, UK, ISBN 0-08-042432-5, p.139 – 151.
- [47] Schmuller, J. (2001).: *Myslíme v jazyku UML*, Grada Publishing a.s., Praha, ČR, ISBN 80-247-0029-8
- [48] Shi, W., Fisher P.F. and Goodchild M.F. (2002): *Spatial Data Quality*, Taylor & Francis, London, UK, ISBN 0-415-25835-9, 336 p.
- [49] Shi, W. (1998): A Generic Statistical Approach for Modelling Error of Geometric Features in GIS. In *International Journal of Geographical Information Science* Volume 12, Number 2, p.: 131-143, Taylor & Francis, London, UK, ISSN 1365-8816
- [50] Stefanakis, E., Sellis, T. (1997): Towards the design of a DBMS repository for the application domain of GIS: requirements of users and applications. In: *Proceedings of the 18th International Cartographic Conference*, Stockholm, Sweden, p. 2030-2037.
- [51] Stromček,V. (2005) – Methodic advance for GIS analysis, In *Proceedings of the International symposium GIS Ostrava 2006*, held in Ostrava 23<sup>rd</sup> – 25<sup>th</sup> January 2006, [http://gis.vsb.cz/GISEngl/Conferences/GIS\\_Ova/GIS\\_Ova\\_2006/Proceedings/Referat/default.htm](http://gis.vsb.cz/GISEngl/Conferences/GIS_Ova/GIS_Ova_2006/Proceedings/Referat/default.htm) (6.2.2006)
- [52] Tiainen, E. (2004): Directions in modeling Land Registration and Cadastre Domain – Aspects of EULIS glossary approach, semantics and information services, In *Proceedings of the Workshop Standardisation in the Cadastral Domain*, Bamberg, Germany, 9-10 December 2004, FIG, Denmark, 2004, <http://www.oicrf.org/searchdocument.asp?e=20> (1.2.2006)
- [53] Tinnachote, Ch., Chen, X. (2003): Implementation of Positional Error Model Using an Object-Oriented Approach, *The 2nd International Symposium on Spatial Data Quality*, Hong Kong, p. 483-491
- [54] Uitermark, H. T., (2001): *Ontology-Based Geographic Data Set Integration*, Deventer, NL, ISBN 90-365-1617-X
- [55] van Oosterom, P.J.M., et al (2004) Remarks and Observations related to the further development of the Core Cadastral Domain Model. In *Proceedings of the Workshop Standardisation in the Cadastral Domain*, Bamberg, Germany, 9-10 December 2004, FIG, Denmark, 2004, <http://www.oicrf.org/searchdocument.asp?e=20> (1.2.2006)
- [56] Veregin, H. (1999): Data quality parameters. In: *Geographical Information Systems*. Second Edition, John Wiley and Sons, New York, p. 177–189.
- [57] Veregin, H. (1998): *Data Quality Measurement an Assessment*, NCGIA Core Curriculum in GIScience, [http://www.ncgia.ucsb.edu/education/curricula/gisc/units/u100/u100\\_f.html](http://www.ncgia.ucsb.edu/education/curricula/gisc/units/u100/u100_f.html)

(1.2.2006)

- [58] Wang I Zhou, et al (2003): Research and Design on The Sharing-Oriented Framework of Geospatial Metadata Management, In Spatial Data Quality, edited by W. Shi, P.F. Fisher, and M.F. Goodchild, Taylor&Francis, London, UK, ISBN 0-415-25835-9, p. 222-226
- [59] Webster's – Third New International Dictionary, <http://www.m-w.com/home.htm> - (5.9.2005)
- [60] Williamson, I., et al, (2003): Developing Spatial Data Infrastructures. From concept to reality, Taylor & Francis, London, 316 p., ISBN 0-415-30265-X
- [61] Worboys, M. (1995): GIS – a computing perspective, Taylor & Francis Ltd, London, ISBN 0-7484-0065-6
- [62] Encyclopaedia Britannica <http://www.britannica.com> (1.2.2006)
- [63] EuroGeographics Expert Group on Quality, [http://www.eurogeographics.org/eng/05\\_quality.asp](http://www.eurogeographics.org/eng/05_quality.asp) (1.2.2006)
- [64] UML Resource Page <http://www.uml.org/> (1.2.2006)
- [65] Open Geospatial Consortium <http://www.opengeospatial.org/> (1.2.2006)
- [66] Lofti Zadeh's Home page: <http://www.cs.berkeley.edu/~zadeh/> (3.2.2006)
- [67] LBMS, s.r.o.: <http://www.lbms.cz/> (10.2.2006)

#### Standards, norms and specifications

- [68] ISO 19103:2005, Geographic information – Conceptual schema language
- [69] ISO 19107:2003, Geographic information – Spatial schema
- [70] ISO 19109:2005, Geographic information — Rules for application schema
- [71] ISO 19110:2005, Geographic information — Methodology for feature cataloguing
- [72] ISO 19111:2003, Geographic information — Spatial referencing by coordinates
- [73] ISO 19112:2003, Geographic information — Spatial referencing by geographic identifiers
- [74] ISO 19113:2002, Geographic information — Quality principles
- [75] ISO 19114:2003, Geographic information — Quality evaluation procedures
- [76] ISO 19115:2003, Geographic information — Metadata
- [77] DIS 19131, Geographic information — Data product specifications, 2005
- [78] FDTS 19138, Geographic information — Data quality measures, 2005
- [79] ISO 2859:1995, Sampling procedures for inspection by attributes
- [80] ISO 3951:1989, Sampling procedures and charts for inspection by variables for percent nonconforming
- [81] DIGEST 2.1:2000, DGIWG, September 2000
- [82] STN P ENV 12 657:2001, Geografická informácia. Opis dát. Metadáta
- [83] STN P ENV 12 656:2001, Geografická informácia. Opis dát. Kvalita
- [84] ISO 9000 series of standards: Translated into plain English, <http://www.praxiom.com/#ISO%209001%20LIBRARY> (1.2.2006)
- [85] Zákon č. 162/1995 Z.z. o katastri nehnuteľností a o zápise vlastníckych a iných práv k nehnuteľnostiam (katastrálny zákon) [www.geodesy.gov.sk](http://www.geodesy.gov.sk) (6.2.2006)
- [86] Zákon č. 215/1995 Z.z. o geodézii a kartografii [www.geodesy.gov.sk](http://www.geodesy.gov.sk) (6.2.2006)
- [87] Vyhláška č.79/1996 Z.z. ktorou sa vykonáva zákon Národnej rady Slovenskej republiky o katastri nehnuteľností a o zápise vlastníckych a iných práv k nehnuteľnostiam (katastrálny zákon) [www.geodesy.gov.sk](http://www.geodesy.gov.sk) (6.2.2006)
- [88] Vyhláška č.178/1996 Z.z. ktorou sa vykonáva zákon o geodézii a kartografii [www.geodesy.gov.sk](http://www.geodesy.gov.sk) (6.2.2006)

- [89] Kolektív autorov (2004):Katalóg objektov ZB GIS verzia 10/2004, ÚGKK SR, TOPÚ Banská Bystrica, 2004, [www.geodesy.gov.sk](http://www.geodesy.gov.sk) (6.2.2006)
- [90] Koncepcia tvorby, aktualizácie a správy Základnej bázy geografického informačného systému do roku 2005 (2002), Úrad geodézie, kartografie a katastra slovenskej republiky, Číslo P-506/2202 z 1.2.2002 [www.geodesy.gov.sk](http://www.geodesy.gov.sk) (6.2.2006)

## SUMMARY

Aim of this thesis is to introduce a view on the spatial data quality from abstraction of the real world into a spatial dataset in order to inform the user of the spatial data via metadata. The spatial data quality is described as the difference between the specification of the dataset and the dataset itself. The quality of a dataset is evaluated against the universe of discourse, described by specification of a dataset. Definition of the spatial data quality helps the data users in their decision-making if the dataset corresponds with their requirements. The quality information reported for a spatial dataset will also depend on a consistent application of standardized methods for measuring the quality of spatial data. The standardized set of evaluation criteria and procedures will guarantee the quality of a spatial dataset. Quality evaluation can be an expensive procedure. Although it is valuable, costs must be weighted against the benefits of quality information – the expenditure of the quality evaluation must be matched to the consequences of errors.

Expert on quality (who is not else as manager) keeps on explaining what quality means in a certain organization rather than making formulas and producing documents – it is very difficult talking about spatial data quality in general terms. The subject of the thesis is too wide to treat it within the extent of the PhD thesis. It requires much more time, pages – sometimes it can be a hole-life research. Many of the data quality parameters vary with the intended use of the spatial dataset or with technology used. Other data quality parameters are not valid anymore or are not recognized yet. During the writing of this thesis the number of valid data quality parameters increased from known 6 (in the year 2000) up to 16 (in the year 2005)!

To conclude the research on the subject chosen for the PhD. thesis – “Quality of the data in spatial datasets” there are following results arising from the thesis:

- defined data quality parameters (chosen according to the standard from the ISO 19100 series of standards) are applied for the model of Slovak Core Cadastral Domain Model – chapter 8.1 and Annex 3.
- suggested methodology for creation of a spatial dataset suggested with respect to the principles of the spatial data quality definition and evaluation. The methodology is suggested for spatial dataset intended (by producer of the dataset) to produce with respect to the international standards from the ISO 19100 series of standards.

Within a context of a theory of spatial data quality following general requirements for the application of the principles of data quality concepts in practice arose from this thesis:

- There is a necessity to change a complex view on a spatial dataset through all the process of its creation – to be able to set a proper data quality model;
- Strengthen the necessity of data product specification – in a case it is not done, it could mean the idea of dataset’s purpose is not very clear, therefore it is very difficult to talk about quality.

Recommendations for development of quality on practice:

- fundamental principle is to be aware who are the users of spatial dataset created, what are their requirements and at what level they demand the data quality evaluation
- producer sets forth the data quality requirements, usually more detailed than user's requirements demand, in a data product specification together with the ontology of the spatial dataset. This specification stays fixed for all process of creation of a spatial dataset. Any changes have strong rules to be done and only relevant exception should be accepted.
- to ensure the set level of data quality of a product, producer sets the abstraction and surveying rules to be kept in order to hold the level of quality.
- the spatial data quality control is done in context with quality management system of the particular organization.

*RESULTS FOR SCIENCE DEVELOPMENT IN THE FIELD OF GEODESY, GEOINFORMATICS AND GEODETIC CARTOGRAPHY:*

- *First self-contained text on the topic of spatial data quality in the local (Slovak) GI community – from abstraction of the reality up to spatial data quality evaluation and control;*
- *Definition of the selected data quality parameters – selection is coordinated by the ISO 19100 series of standards*
- *List of methods for evaluation of data quality in the terms of international standard*
- *List of data quality measures*

*RESULTS FOR GEODETIC, GEOINFORMATIONAL AND CARTOGRAPHICAL PRACTICE:*

- *Application of the data quality principles through the quality model for SkCCDM*
- *Contribution to the research on the multipurpose cadastral domain model of Slovak republic – by the application (in Annex 3)*
- *Proposal of the list recent methods for evaluation of data quality parameters selected in accordance of standard specification. Author's had chosen those parameters relevant (intended to document by producer, or discussed during the process of creation) for the application within a selected spatial dataset (ZB GIS);*
- *Suggestion of methodology for creation of spatial dataset in context of principles for spatial data quality evaluation*