

Štandardizácia, formalizácia, katalógy - nevyhnutné predpoklady pre modelovanie topografickej témy a integráciu rôznorodých systémov

Autor: Ing. Marián Adamják¹

1. Úvod a ciele

Výskum histórie človeka, dejiny ľudstva nám ukazujú, že jedným z najvýznamnejších medzníkov vo vývoji človeka bol vznik artikulovanej reči a vznik písma. Kým vznik artikulovanej reči umožnil pravekému človeku efektívnu výmenu informácií, vznik písma mu umožnil uchovávanie týchto informácií dlhšiu dobu nezávisle od dĺžky života ľudských bytosťí. Písma však umožňuje človeku nielen informácie uchovávať ale zároveň ich aj odovzdávať bez osobného kontaktu medzi ľuďmi. Denne čítame noviny, denne sa stretávame s písaným textom ale neuvedomujeme si že vo svojej podstate sa jedná len o odovzdanie informácie medzi dvomi ľuďmi – autorom a čitateľom. Na pozadí tohto dejia – výmeny informácie – však autor „automaticky“ predpokladá že čitateľ bude poznať znaky ktorým je text napísaný a bude rozumieť významu skupín znakov (slovám). Autor sa o tento predpoklad môže opriest preto, lebo čitateľ bol s najväčšou pravdepodobnosťou zoznámený s jedným z najznámejších štandardov – s jazykom.

Jazyk má niektoré pozoruhodné vlastnosti. Jazyk obvykle nie je záväzne kodifikovaný štátou (ani obdobnou) autoritou a predsa je široko rozšírený a používaný. Jazyk sa vyvíja v čase, jeho prostriedky sa menia a kopírujú vývoj ľudí a ich spoločenstva.

Ludová múdrost hovorí, že je dobré poučiť sa z cudzích chybách a úspešné skutky opakovať. Preto sa najprv pokúsim nájsť isté analógie medzi dorozumievacím jazykom ľudí a vyjadrovacím aparátom, ktorý sa používa v informačných systémoch, najmä v systémoch ktorých úlohou je modelovať geosféru, alebo jej časť. Na tejto analógii sa pokúsim demonštrovať nevyhnutnosť štandardizácie. V ďalšej časti potom načrtнем niektoré možnosti a riešenia ako postupovať pri štandardizácii na národnej úrovni a v závere sa pokúsim zhodnotiť doterajšie úsilie skupiny expertov, ktorí sa podieľajú na štandardizácii prostredia v ktorom by sa mohla vybudovať Národná priestorová infraštruktúra.

2. Jazyk a informačné systémy

Informačný systém je možné charakterizovať ako súbor technických a technologických prostriedkov, programového vybavenia, údajov, postupov a personálu, ktorý slúži najmä na uchovávanie, spracúvanie a následné využívanie týchto informácií. Najčastejším účelom budovania informačných systémov je podpora riadenia a rozhodovania. Poznávacie a dokumentačné funkcie informačných systémov sú taktiež významné avšak obvykle sú obsiahnuté v podpore riadenia.

Jednou zo základných činností pri vytváraní informačného systému je návrh štruktúry jeho údajového skladu a zároveň návrh vyjadrovacích prostriedkov, ktorými budú informácie používateľovi systému zobrazované a interpretované. Tieto vyjadrovacie prostriedky obvykle vychádzajú zo zvyklostí a overených postupov, ktoré budú používatelia používať v dobe kedy

¹ Topografický ústav Banská Bystrica, <http://topu.army.sk>

je systém navrhovaný, alebo vyvíjaný. A práve tento fakt, že vyjadrovacie prostriedky sú prispôsobované zaužívaným postupom a metódam, že kopírujú existujúce vyjadrovacie prostriedky vytvára isté paralely medzi ľudským jazykom a vyjadrovacími prostriedkami informačných systémov.

Ak jazykom rozumieme sústavu vyjadrovacích, obvykle znakových, prostriedkov istého spoločenstva, ktorý slúži ako nástroj myslenia, dorozumievania a ukladania poznatkov, potom spôsob akým sa „dorozumieva“ informačný systém s používateľom nesie zrejme všetky znaky jazyka.

Vlastnosti ľudského jazyka	Spôsob komunikácie človek - IS
Vyjadrovacie prostriedky (znaky) tvoria konečnú množinu – znaky a slová bežne používaných abecied sú konečnou množinou.	Znaky na dorozumievanie medzi človekom a IS tvoria konečnú množinu prvkov. Ak sa nepoužíva ľudský jazyk, potom použité symboly sú dohovorené a tvoria konečnú množinu
Pravidlá akými sa vyjadrovacie prostriedky používajú na vyjadrenie myšlienky, poznatku sú ustálené, tvoria konečnú množinu pravidiel.	Pravidlá ktoré určujú ako a kedy sa používajú, jednotlivé symboly, znaky a dohovorené texty sú zrejme konečnou množinou, keďže sú súčasťou konečnej množiny inštrukcií programov, ktoré obsluhujú IS.
Jazyk slúži na dorozumievanie v komunite. Táto komunita je vymedzená okrem iného tým, že rozumie tomuto jazyku.	Dorozumievanie medzi človekom a IS je založené na tom, že človek bude rozumieť vyjadrovacím prostriedkom IS a IS systém bude „rozumieť“ človeku ak tento bude používať dohodnuté vyjadrovacie prostriedky.
Jazyk slúži na ukladanie poznatkov – myslí sa tým najmä písaný jazyk.	Technické prostriedky IS umožňujú uložiť poznatky dohovoreným spôsobom poznatky, späťne ich dohovoreným spôsobom zobraziť, trvalo zaznamenať v človekom vnímateľnej forme (výtačok na papieri).

Z predchádzajúceho porovnania je zrejmé, že spôsob dorozumievania medzi človekom a IS, ale aj v širšom chápaní medzi dvomi obecnými subjektami je založený na jazykoch.

Definícia jazyka sa opiera o spoločenstvo, ktoré prostredníctvom tohto jazyka komunikuje. Pojem spoločenstvo je pre ďalšie uvažovanie veľmi dôležitý, spoločenstvá sú obvykle nekonfliktné komunity v ktorých ľahko dochádza k zhode na vyjadrovacích prostriedkoch jazykov a na pravidlach jazykov. Príkladom takýchto spoločenstiev sú okrem národov, národností, či etnických skupín aj odborné komunity, ktoré okrem toho že používajú vlastnú odbornú terminológiu (vyjadrovacie prostriedky) ľudského jazyka, majú aj vlastné ustálené postupy a vyjadrovacie prostriedky informačných systémov.

Hoci v minulosti boli tieto odborné komunity len málo prepojené a v prípade ich vzájomnej komunikácie dochádzalo len k malej výmene informácií, dnes je to naopak. Vzájomná komunikácia medzi komunitami je nevyhnutnosťou a množstvo informácií, ktoré komunity navzájom zdieľajú sa veľmi rýchlo zväčšuje. Z tohto plynne posledná paralela medzi jazykmi a dorozumievacími prostriedkami informačných systémov – a tou je dorozumievanie medzi spoločenstvami a komunitami.

Dorozumievanie medzi kultúrnymi komunitami ľudí, t. j. medzi spoločenstvami s odlišným jazykom sa deje na základe pravidiel, ktoré umožňujú prekladať jednotlivé prvky

jedného jazyka do jazyka druhého. Pravidlá prekladu sú založené na spoločnom chápaní a vnímaní objektívnej reality. Čím je chápanie objektívnej reality medzi komunitami odlišnejšie tým zložitejšie pravidlá prekladu medzi jazykmi je potrebné vytvoriť – často sa stáva, že nie je ani dokonca možné vytvoriť takéto pravidlá a je možný len približný preklad významu prekladanej jazykovej štruktúry.

Obdobne pre dorozumievanie medzi odbornými komunitami a ich „jazykmi“ je potrebné vytvoriť súbor pravidiel ktorými je možné „preklad“ uskutočniť. Predpokladom k vytvoreniu takýchto pravidiel je rovnaké, alebo aspoň veľmi blízke chápanie objektívnej reality, resp. jej jednotlivých častí.

V prostredí informačných systémov, ktoré obsahujú priestorovú zložku, resp. opisujú geografickú sféru, alebo jej niektorú časť je pre úspešnú výmenu informácií medzi rôznymi informačnými systémami nevyhnutné na jednej strane rešpektovať zvyklosti a „odborný život“ komunít ktoré tieto systémy používajú, ale na druhej strane je potrebné, aby tieto komunity vnímali elementy geografickej sféry rovnako alebo aspoň významovo veľmi blízko.

Toto blízke chápanie jednotlivých elementov geografickej sféry považujem za nevyhnutný predpoklad štandardizačného procesu, ktorým sa upraví vzájomná výmena a zdieľanie informácií medzi informačnými systémami rôznych autorov, používateľov či vlastníkov.

3. Štandardizácia na národnej úrovni

Štandardizácia na národnej úrovni je založená na vydávaní normatívnych aktov, ktoré upravujú technické vlastnosti výrobkov, formu a obsah dokumentov a pod. Tieto normatívne akty sú rôznej povahy od zákonných noriem až po odporúčania a metodické pokyny. Existujú však aj dohody skupín špecialistov, najmä na medzinárodnej úrovni, ktoré sú považované za „de facto“ štandardy a sú rešpektované širšou odbornou komunitou. Typickým predstaviteľom takýchto „de facto“ štandardov je väčšina noriem upravujúcich komunikáciu na internete, v prostredí GIS sú to potom aktivity konzorcia OpenGIS.

Postup na národnej úrovni v oblasti informačných systémov, ktoré obsahujú priestorové informácie či skôr v geoinformačných technológiách (GIT), by mal vychádzať z existujúcich záväzných alebo uznávaných „de facto“ štandardov. Medzi záväzné štandardy patria najmä ISO normy, smernice EU a normatívne akty ústredných orgánov štátnej správy, v tomto prípade najmä Úradu geodézie kartografie a katastra. Medzi „de facto“ štandardy je možné zaradiť zverejnené normy a normatívne dokumenty týkajúce sa informačných systémov na úrovni rezortov, prípadne „de facto“ štandardy vytvorené širokým rozšírením niektorých aplikácií (Kokeš, Teplyn, ESRI shape file) a pod.

V ďalšom texte sa sústredím výhradne na štandardizáciu majúcu vzťah k informačným systémom s priestorovou zložkou a k zdieľaniu údajov, najmä vo vzťahu k moderným technologickým trendom v budovaní IS. Týmto trendom sú najmä servisne orientované architektúry. Zároveň vyslovím predpoklad, že IS ktoré obsahujú priestorovú zložku alebo chcú využívať informácie o objektoch reálneho sveta budú využívať viaceré informačné zdroje – viaceré databázy alebo viaceré služby poskytujúce údaje v reálnom čase či ich kombinácia. Pritom elementárnym predpokladom správneho využitia poskytnutej informácie je, aby bola obidvomi stranami – poskytovateľom aj používateľom – chápaná rovnako. Zhodné chápanie pojmov, ktoré reprezentujú realitu (objekty v reálnom svete) sa na prvý pohľad môže javiť triviálne, avšak vo väčšine prípadov tomu tak nie je. Príkladom nech je objekt ktorý laik pomenuje „železnica“ avšak špecialisti rozlišujú železničnú trať, vlečku, koľaj a mnoho ďalších

pojmov. Pritom odborník z iného odboru než „železničiar“ bude potrebovať generalizovať tieto odborné pojmy práve na pojem „železnica“. Ďalším príkladom nech je objekt ktorý väčšina laikov nazve trafostanica, pričom odborník rozlišuje prostý transformátor, zariadenie pre prepínanie a rozvod elektrickej energie a ďalšie pojmy. Toto sa môže komplikovať, ak tieto zariadenia nie sú na stĺpoch elektrického vedenia, ale sú v budovách. Potom urbanista nemusí vnímať zásadný rozdiel medzi budovou kde je umiestený transformátor, budovou požiarnej stanice a malou hospodárskou budovou.

Skúsenosti ukazujú, že vzájomné efektívne zdieľanie údajov môže byť komplikované ba až znemožňované rozdielnym vnímaním objektov reálneho sveta. Toto rozdielne vnímanie objektov reálneho sveta je potrebné vnímať ako prirodzený prvok odbornej a kultúrnej rozmanitosti a je potrebné vytvoriť také národné štandardy a normy, ktoré budú:

- rešpektovať špecifiká odborných komunít,
- umožňovať zblíženie chápaní objektov reálneho sveta,
- reflektovať jestvujúce štandardy a normy.

Tieto podmienky som sa pokúsil dodržať pri návrhu konceptu „Národného registra tried priestorových objektov“ (NRTPO). Tento koncept vznikol ako odpoveď na niektoré ľažko riešiteľné, a pritom oprávnené požiadavky zástupcov odborných komunít pri vytváraní „Katalógu tried objektov pre ZBGIS resp. CPD“.

4. Skúsenosti z tvorby štandardov pre ZBGIS a CPD a koncept NRTPO.

Koncept obsahu Centrálnej priestorovej databázy Vojenského informačného systému (CPD VISÚ) začal vznikať približne v roku 1998, kedy v Topografickom ústave padlo zásadné rozhodnutie vytvoriť takúto databázu na základe priameho poznania reálneho sveta – tzn. priamym mapovaním vhodnými metódami a nie prevzatím iných modelov reálneho sveta (vektorizáciou máp). Toto zásadné rozhodnutie postavilo riešiteľský tým pred náročný problém a tým bolo najmä vytvorenie zoznamu tried objektov a ich vlastností, ktoré majú tvoriť model reálneho sveta uloženého v CPD VISÚ. Prvú verziu tohto zoznamu tried a vlastností, tzv. Katalóg tried objektov CPD sa podarilo ustáliť na prelome rokov 2000 a 2001. Vychádzali sme pri tom zo štandardu NATO STANAG 7074, ktorý kopíruje „de facto“ štandard DIGEST². V súbežne s prípravou katalógu bol vytvorený aj prototyp technologických postupov. V rokoch 2001 a 2002 bol vykonaný pilotný zber údajov, na základe ktorého riešiteľský tým pristúpil k zásadnej korekcii obsahu Katalógu tried objektov CPD a zároveň k spresneniu technologických postupov. Táto zásadná korekcia bola vynútená najmä tým, že model územia vytváraný na základe prvej verzie katalógu neumožňoval výstižne a verne zachytiť všetky požadované fenomény. Od roku 2003 dochádza len k malým úpravám Katalógu tried objektov CPD.

Významným medzníkom v krátkej histórii Katalógu tried objektov CPD bolo pristúpenie Geodetického kartografického ústavu k zásadám ktoré riešiteľský tým CPD prijal v dobe projektovania CPD. Osvojenie si klúčových zásad zo strany GKÚ umožnilo vytvoriť spoločný katalóg tried objektov, ktorý má byť katalógom pre ZB GIS. Toto zblíženie stanovísk umožňuje okrem iného vytvárať veľmi podobné modely reálneho sveta a tak efektívne vytvoriť referenčný základ pre IS obsahujúce alebo využívajúce informácie s priestorovou zložkou.

Posledným významným faktorom ktorý má vplyv na obsah Katalógu tried objektov CPD (a ZBGIS) sú príspevky členov Pracovnej skupiny pre GIS v štátnej správe (PS GIS), ktorá je

² www.digest.org

vytvorená Radou vlády pre informatiku. Návrhy členov PS GIS na zmeny a úpravy katalógov boli však typickým prejavom odlišného chápania reality, hoci tieto odlišnosti boli členmi PS GIS vzájomne veľmi tolerované so snahou o dosiahnutie zhody. Na účely posúdenia všetkých návrhov a požiadaviek si PS GIS vytvorila pracovný tým, ktorý posúdil všetky návrhy a požiadavky jednotlivých členov PS GIS. Časť z týchto návrhov a požiadaviek bola do katalógu zapracovaná, ale časť bola zamietnutá. Najčastejším dôvodom zamietnutia návrhu alebo požiadavky bola prílišná odborná špecifickosť navrhovanej úpravy katalógu, ktorá sa obvykle týkala len veľmi malej odbornej komunity. PS GIS sa však rozhodla pravidelne vyhodnocovať všetky návrhy a námety a periodicky vydávať úpravy katalógu.

Značné spektrum a odlišnosti v návrhoch členov PS GIS, špecifické požiadavky na rozšírenie katalógu ma viedli k vypracovaniu konceptu „Národného registra tried priestorových objektov“, ktorý bol členom PS GIS predstavený na zasadnutí PS GIS v máji 2005. Cieľom tohto konceptu je na jednej strane zachytiť všetky požiadavky na úpravu katalógu a na druhej strane vytvoriť taký štandard, ktorý umožní chrániť doteraz vyhotovené údajové množiny a databázy bez zásadných zmien týchto množín a databáz.

4.1. Koncept NRTPO a jeho zložky

Účelom „Národného registra tried priestorových objektov“ NRTPO by malo byť registrovať všetky používané triedy objektov, všetky používané atribúty a ich domény hodnôt, ktoré sa používajú v registrovaných IS s priestorovou zložkou. Súčasťou konceptu je aj odporúčanie na vytvorenie vhodnej inštitucionálnej a technickej infraštruktúry, ktorá by umožnila takýto register viesť.

NRTPO by sa mal skladáť z nasledovných súčastí:

1. Register tried objektov,
2. Register atribútov,
3. Register IS s priestorovou zložkou,

Medzi regisrami tried a atribútov nebude existovať žiadny vzťah. Účel a obsah jednotlivých regisrov je uvedený v nasledovnom teste.

Register tried objektov

Obsahom registra je zoznam a opis tried objektov. O každej triede objektov by sa mali viesť nasledovné informácie:

1. Názov triedy (jedinečný)
2. Výstižný popis, definícia triedy
3. Odporúčaná skratka, alebo identifikátor
4. Údaje o platnosti registrácie

Registrovať by sa mala každá trieda, ktorá bude používaná v i len jednom z registrovaných IS. Pri registrácii sa posúdi, či neexistuje už predtým registrovaná trieda, ktorá reprezentuje rovnaký objekt, alebo jav.

Register atribútov

Obsahom registra je zoznam a opis atribútov (vlastností), ktorými sú opisované objekty v registrovaných IS. O každom atribúte sa vedú nasledujúce informácie:

1. Názov atribútu (jedinečný)
2. Výstižný popis, definícia atribútu
3. Odporúčaná skratka, alebo identifikátor

4. Doména hodnôt
5. Údaje o platnosti registrácie

Registrovať by sa mal každý atribút, ktorý bude používaný v i len jednom z registrovaných IS. Pri registrácii sa posúdi, či neexistuje už predtým registrovaný atribút rovnakého významu. Obdobne sa bude postupovať pri dopĺňaní, resp. rozširovaní domén hodnôt.

Register IS

Obsahom registra by mal byť zoznam registrovaných IS, ktoré obsahujú priestorové údaje. Pod pojmom „obsahuje priestorový údaj“ sa rozumie skutočnosť, že sa v ňom vyskytujú, alebo môžu vyskytovať entity reprezentujúce priestorové objekty (alebo javy) vrátane opisu geometrie týchto objektov a priestorových vzťahov. Pre každý registrovaný systém sa budú viest' nasledovné údaje:

1. Názov IS, skratka, označenie
2. Prevádzkovateľ
3. Informácia o prevádzkovom stave
4. Zoznam všetkých tried, ktoré sú v ňom používané, pričom v zozname môžu byť len registrované triedy. Pre každú triedu sa vedia zoznam atribútov, ktoré sa k danej triede v IS používajú. Môžu sa použiť len tie atribúty, ktoré sú registrované. Pre každú triedu a atribút sa vedú informácie o názve (kóde) týchto tried, resp. atribútov v konkrétnom IS. Prevádzkovateľ môže zverejniť o každej triede (atribúte) ďalšie údaje – najmä o kvalite
5. Údaje o platnosti registrácie

4.2. Registračné procedúry

Prevádzkovateľ, ktorý prevádzkuje IS s priestorovými údajmi, bude môcť (prípadne musieť) požiadať o registráciu svojho IS. V registrácii bude musieť vytvoriť väzbu medzi „svojimi“ triedami a atribútmi a registrovanými triedami, resp. atribútmi. Ak sa v registri nebude nachádzať vhodná trieda resp. atribút, ktorý bude zodpovedať jeho potrebám, požiada o zaregistrovanie novej triedy, resp. atribútu.

4.3. Vzťah k dnešnému katalógu objektov

Prototypy registrov spomenutých v časti 4.1 môžu byť vytvorené zo zoznamu tried, atribútov a ich domén ktoré sú dnes obsiahnuté v Katalógu tried objektov CPD resp. katalógu tried objektov používaných pre ZBGIS. Registre IS môže byť vytvorený dvomi regisitračnými záznamami a to záznamami registrujúcimi ZBGIS a CPD. Dnešný Katalóg tried objektov CPD sa potom stane súčasťou registrácie CPD, pretože bude formálne mapovať triedy objektov a atribútov CPD do tried novovo vzniknutých registrov (obdobne sa bude zrejme postupovať pri ZB GIS).

V prípade, že „iný“ IS sa bude chcieť registrovať nemusí vo svojej štruktúre striktne používať názvy a označenie triedy z dnešných katalógov používaných pre ZBGIS (alebo CPD), ale stačí ak svoje triedy „namapuje“ na registrované triedy. Ak potenciálny „tretí“ používateľ tohto „iného“ IS bude z neho chcieť využívať údaje bude môcť v registroch NRTPO zistiť všetky potrebné údaje o triedach objektov, atribútoch a doménach tohto „iného“ IS, čo mu umožní spoľahlivé využitie údajov a informácií z tohto „iného“ IS.

4.4. Odporúčania k vytvoreniu NRTPO

Pre vytvorenie NRTPO, bude potrebné zhodnúť sa na autorite, ktorá bude register viesť. Podľa platných zákonov by takouto autoritou nesporne mal byť Úrad geodézie, kartografie a katastra (UGKK). Pri ÚGKK by však mala pracovať "stála komisia", ktorá bude posudzovať žiadosti o registrovanie v jednotlivých častiach registra. Hlavnou úlohou tejto komisie by malo byť najmä zamedzenie duplicit v jednotlivých častiach registra a vyžadovanie úplnosti v registračných údajoch. Na základe návrhu tejto stálej komisie bude ÚGKK vydávať vhodnou formou (Smernica, Oznámenie a pod.) platnú verziu NRTPO. Samozrejmost'ou by mala byť elektronická dostupnosť aktuálne platnej a aj predchádzajúcich verzií NRTPO.

Zriadenie autority a detaile implementácie takého registra upravuje aj norma ISO 19135, ktorá je v súčasnosti pripravovaná skupinou ISO/TC 211.

5. Zhrnutie

V príspevku som sa snažil poukázať na dôležitosť štandardizácie v oblasti informačných systémov ktoré obsahujú aj priestorové informácie. Dokumentoval som paralely medzi jazykom človeka a dorozumievaním, či výmenou informácií medzi rôznymi informačnými systémami. Preklady medzi rôznymi ľudskými jazykmi považujem za paralely s rozhraniami medzi informačnými systémami. Tieto rozhrania, cez ktoré si informačné systémy vymieňajú informácie, sú sice vytvárané na základe určitých štandardov avšak tieto štandardy len málokedy štandardizujú obsah, syntax a sémantiku vymieňaných informácií. Potreba a rozšírenie štandardov obsahu vymieňaných informácií nie je vynucovaná aj preto, lebo priestorové informačné systémy budú neposkytujú informácie v reálnom čase – napr. formou služieb, alebo výmena informácií sa vykonáva len prostredníctvom nosičov a to často jednorázovo. Formát a štruktúra sa pri takejto výmene obvykle dohodne na úrovni používateľov takýchto systémov.

Tento stav, kedy budú štandardy neexistujú, alebo sa neaplikujú zrejme nebude mať dlhé trvanie, pretože je možné predpokladat výrazné rozšírenie služieb ktoré informácie budú poskytovať v reálnom čase. Ďalším faktorom, ktorý si vynúti štandardizáciu aj na úrovni údajových štruktúr a obsahu bude zvýšená citlivosť používateľov na aktuálnosť informácií. Ak sa informácie budú aktualizovať priebežne, potom sa výmena údajov na nosičoch stane prakticky nepoužiteľnou a k údajom bude potrebné pristupovať v reálnom čase. Pristupovanie k informáciám v reálnom čase viacerými používateľmi by mohli mať za následok že poskytovateľ údajov by musel pre každý pristupujúci subjekt zriaďovať osobitné rozhrania, čo zrejme nebude efektívne.

Štandardizácia umožní výrazne zefektívniť procesy pri poskytovaní a vzájomnej výmene a zdieľaní informácií. Autorizované registrovanie informačných systémov a metaúdajov o nimi poskytovaných údajoch a službách vytvorí prostredie pre efektívny a najmä ekonomický rozvoj informačných systémov v SR.