

ZÁKLADNÉ PRINCÍPY A OPERÁCIE HARMONIZÁCIE PRIESTOROVÝCH DÁT

Kristína KROČKOVÁ

The basic principles and operations of spatial data harmonisation

Abstract: The spatial data harmonisation is theme that is closely related to the data interoperability. Interoperability means activities to eliminate data heterogeneity, which present a big problem nowadays. One of the offered solutions would be building the Spatial Data Infrastructures (SDI). On European level the main activity is concentrated in working on project INSPIRE, which defines required rules for building the infrastructure for Europe region. Not only in Europe but many other countries around the world investigate a lot of time to have their data harmonised in SDI. Defined processes, schemas, rules and other related aspects are standardised on international level to assign interoperability as much as possible. The main goal is to save funds and to prevent duplicity in data production. Somewhere in the middle of this complicated process is data harmonisation that provides transformation between source data sets and target schemas defined by SDI requirements. The most important is correct understanding of basic operations and all possible coming problems before starting the spatial data harmonisation process of data from any domain.

Keywords: interoperability, spatial data infrastructure, data harmonisation, schema mapping, reclassification, transformation, INSPIRE, Humboldt

Úvod

Problémom súčasnosti sa stáva obrovské množstvo priestorových dát, ktoré sú tvorené nejednotným spôsobom, bez definovaných aspoň základných pravidiel a špecifikácií. Laik by povedal, že každý si dáta tvorí „po svojom“ a tá istá informácia môže byť spracovaná rôznym spôsobom, často podľa subjektívneho prístupu tvorcu. Navyše dochádza k duplicitnej tvorbe a uchovávaní dát, čo vedie samozrejme k zbytočnému zvyšovaniu nákladov. Riešením eliminácie týchto duplicit je vypracovanie a účinné využívanie metadát, ktoré podrobne opíšu dáta a používateľ tak dostane informácie bez nutnosti priameho prístupu k dátovým vrstvám. V ďalšej fáze prichádza na rad proces harmonizácie, ktorý rieši vzájomnú heterogenitu dát a umožňuje tak vytváranie interoperability jednotlivých dátových štruktúr nielen v rámci jednej oblasti, regiónu, či krajiny, ale hlavne v medzinárodnom až globálnom rozsahu.

Harmonizácia dát je zložitý a komplikovaný proces závislý na faktoroch samotnej heterogenity dotknutej domény dát. Existuje a stále vzniká veľa projektov, ktoré pracujú na definovaní harmonizačných pravidiel a na ich implementácii, testovaní a validácii. Pracovné tímy sa v rámci nich snažia o optimalizáciu procesov a definovanie pravidiel tak, aby boli aplikovateľné pre danú doménu priestorových dát.

Článok sa skladá z dvoch častí, ktoré opisujú problematiku harmonizácie priestorových dát, jej základné princípy s vysvetlením základných operácií. Prvá časť je venovaná vysvetleniu pojmu harmonizácia dát, princípom a požiadavkám kladeným na tento proces. Vychádza z projektu INSPIRE, ktorý túto problematiku podrobne rozpracováva. Ďalšia časť definuje samotné základné pojmy harmonizácie dát a následne opisuje jednotlivé operácie harmonizácie priestorových dát a uvádza názorné príklady problematickej konverzie a reklasifikácie dát. Nechýba opis transformácie dát do Európskeho terestrického referenčného súradnicového systému 1989 (ETRS89) ustanoveného smernicou INSPIRE (2007), s odporúčanými kartografickými zobrazeniami používanými v závislosti od nárokov na presnosť výstupnej mapy a jej mierky.

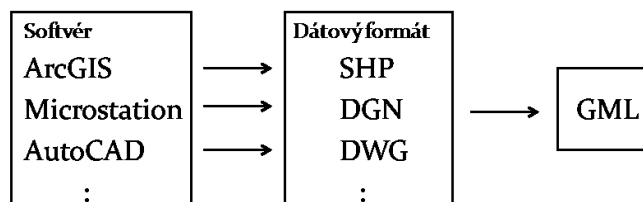
Ing. Kristína KROČKOVÁ, Slovenská technická univerzita v Bratislave, Stavebná fakulta, Katedra geodetických základov, Radlinského 11, 813 68 Bratislava, e-mail: kristina.krockova@stuba.sk

1. Harmonizácia priestorových dát

Existuje niekoľko dôvodov, prečo priestorové dáta nemožno považovať za homogénne. Sú to napr. rozdielne normy a štandardy použité pri výrobe dát, rôzne dátové modely, mierka, jazyk a symbolika dát, rôzne súradnicové systémy a jednotky a v neposlednom rade i rôzna úroveň kvality dát. Celkovo má proces odstránenia heterogenity dát za následok vznik ich vzájomnej interoperability (obr. 1). Proces dosiahnutia vzájomnej interoperability dát je zabezpečený pomocou harmonizácie dát.

Slovo harmonizácia je odvodené od gréckeho slova „*harmozo*“ - *zapadať, pripojiť sa*. To napovedá o základnom princípe harmonizácie dát a zmysle tohto procesu aplikovaného, v našom prípade, na oblasť priestorových dát.

Dôvodov na harmonizáciu priestorových dát je hneď niekoľko. Podrobne sú predstavené v nasledujúcej kapitole spolu s princípmi harmonizácie definovanými v smernici INSPIRE (2007), ktorá predkladá i základné požiadavky na harmonizáciu. Samotný proces harmonizácie je definovaný rámcovo a líši sa prípad od prípadu.



Obr. 1 Schéma harmonizácie rôznych dátových formátov

1.1 Princípy a požiadavky na harmonizáciu dát

Podľa toho, aký druh činnosti chceme s priestorovými dátami vykonať, rozlišujeme nasledujúce dôvody pre harmonizáciu dát. Najbežnejším dôvodom dnes je snaha o spájanie dát z rôznych zdrojov, pričom napríklad priestorové dáta navzájom geometricky nezapadajú, majú rôzne mierkové spracovanie, typ a reprezentácia prvku sú navzájom odlišné atď. Ďalším dôvodom je požiadavka na poskytovanie dát spĺňajúcich štandardy (napr. definované v smernici INSPIRE, 2007), alebo na importovanie dát do existujúcej aplikácie.

Jednou zo základných úloh definovaných v smernici INSPIRE (2007) je harmonizácia dát formou vývoja dátových špecifikácií. Tento typ harmonizácie je definovaný ako proces vývoja jednotnej množiny špecifikácií dátových produktov spôsobom, ktorý umožňuje poskytnutie prístupu k priestorovým dátam pomocou priestorových dátových služieb v reprezentácii, ktorá umožňuje uceleným spôsobom ich kombinovanie s ďalšími harmonizovanými dátami.

Dátové špecifikácie INSPIRE sú výsledkom harmonizačného procesu založeného na existujúcich dátových špecifikáciách, a kde je to možné na požiadavkách používateľov a prípadových štúdiách. Práca na špecifikáciách je založená na systéme, ktorý identifikuje relevantné komponenty procesu harmonizácie dát. V tomto kontexte sa uvažuje o požiadavkách harmonizácie dát na dvoch rôznych úrovniach, a to na úrovni aplikačných schém a dátovej úrovni (napr. napojenie dát v hraničných oblastiach).

Drafting Team „Data Specifications“ (2010), ktorý je základom pre vývoj tematických dátových špecifikácií, definuje päť hlavných harmonizačných požiadaviek pre:

- rôzne dátové formáty – poskytnúť interoperabilný prístup k heterogénnym dátovým zdrojom formou prevodu (obr. 1) do jazyka GML (Geography Markup Language),
- rôzne dátové modely – poskytnúť riešenia pre harmonizáciu dátových modelov formou dátových špecifikácií a konceptuálnych schém,
- chýbajúce, nekonzistentné a zastarané metadáta – poskytnúť riešenia na vyhľadanie a možnosti tvorby a publikovania metadát,
- problémy s významom (sémantikou) objektov – poskytnúť riešenia ako aplikačný doménový slovník a Thesaurus,

- rozdielne súradnicové systémy – poskytnutie súradnicovej transformačnej služby na prevod do Európskeho terestrického referenčného súradnicového systému 1989 (ETRS89).

1.2 Proces harmonizácie dát a jeho hlavné problémy

Do procesu harmonizácie dát vstupuje množstvo komponentov ako dátové formáty, referenčné systémy, dátové modely, konceptuálne a aplikačné schémy, ale i terminológia, jazyky či mierka. Okrem nich harmonizáciu dát sprevádzajú aj rôzne procesy ako spracovanie, transformácia, spájanie dát, generalizácia, reklasifikácia, kódovanie alebo dekódovanie dát. Z tohto množstva komponentov a procesov vyplýva, že sa jedná o zložitý proces. Jeho komplexnosť je možné a zároveň vhodné redukovať rozdelením harmonizačného procesu na menšie, viac špecifické procesy a ich reťazením dosiahnuť želaný finálny výsledok.

Na základe mnohých testovaní v rámci projektu Humboldt (HUMBOLDT Consortium, 2006 – 2011) sa proces harmonizácie dát rozdeľuje na dve základné fázy (Fichtinger et al., 2011). Prvou je definovanie cieľa, ktoré začína analýzou a identifikáciou reálnych požiadaviek používateľa, na ktorom spolupracujú doménoví experti s vývojármi. Požiadavky sa opisujú zväčša pomocou modelovacieho jazyka UML (Unified Modelling Language). Dôležitá je podrobná analýza dostupných dátových zdrojov a prípadné hľadanie možnosti ako nahraďiť chýbajúce. Nasleduje identifikácia požiadaviek na samotnú harmonizáciu dát a dizajnovanie konceptuálneho dátového modelu UML. Výsledkom celého procesu sú zdrojové a cieľové dátové špecifikácie prepojené aplikačnou schémou, ktorá sa následne implementuje, testuje a validuje. Druhou fázou harmonizácie je samotné spracovanie transformácie heterogénnych dát z rôznych dátových zdrojov tak, aby odpovedala cieľovej dátovej špecifikácii.

Na základe skúseností z doterajších projektov budovania infraštruktúr priestorových informácií (SDI – Spatial Data Infrastructure) a projektov zaoberajúcich sa harmonizáciou priestorových dát sa objavilo mnoho problémov, ktoré bránili hladkému priebehu nastavených procesov harmonizácie. Počas ich riešenia sa stále pracovalo a pracuje na zdokonaľovaní a vylepšovaní pôvodných návrhov. Počas tohto vývoja sa ustálilo niekoľko hlavných komplikácií, s ktorými sa stretne takmer každý pokus o dosiahnutie interoperability v SDI. Podľa rozdielnej úrovne komplikácií, ktoré môžu byť riešené pri dosahovaní interoperability, rozlišujeme tieto hlavné typy (SDI Cookbook, 2009):

- „cross-border“ – napájanie geometrie rôznych dátových vrstiev,
- „cross-sector“ – tvorba dátových štruktúr pre rôzne aplikačné domény,
- „cross-type“ – kombinovanie rozličných dátových typov (napr. rastrových a vektorových),
- „overlap“ – rovnaké prvky pochádzajúce z rôznych zdrojov a procesov.

2. Základné operácie harmonizácie priestorových dát

Harmonizácia dát je zložitý proces, ktorý si vyžiadal definovanie vlastnej terminológie a špecifické základné pojmy, ktoré sa používajú nielen v akademických publikáciách, ale aj v bežnej praxi. Je potrebné si ozrejmiť a vysvetliť jednotlivé pojmy, ktoré budeme ďalej uvádzať v ich pôvodnej anglickej forme tak, ako ich definuje projekt Humboldt (HUMBOLDT Consortium, 2006 – 2011).

Schema translation predstavuje formálny opis modelu, ktorý transformuje zdrojové dáta podľa požiadaviek výstupnej dátovej špecifikácie.

Schema matching je proces identifikácie dvoch sémanticky príbuzných objektov.

Schema mapping predstavuje transformácie medzi objektmi.

Edge matching predstavuje detekciu a následné spájanie navzájom prislúchajúcich si geometricky reprezentovaných objektov.

2.1 Schema translation

Za najdôležitejšiu operáciu harmonizácie dát býva považovaná *schema translation*, ktorá tvorí základný pilier harmonizačného modelu a rozoznáva tieto základné typy schém:

- konceptuálna – formálna špecifikácia dátového modelu, kde model je tvorený z dátových štruktúr, kódových listov (Code List) atď., najčastejšie vyjadrená jazykom UML.
- logická/fyzická – je zameraná viac na fyzickú štruktúru modelu. Ako vyjadrovací prostriedok sa používa tzv. XSD (XML schema definition) súbor definovaný pomocou jazyka XML (XML schema definition language).

- prevodné súbory (XML/GML) – svojím obsahom reprezentujú fyzickú štruktúru zdrojového a cieľového modelu.

Schema translation pozostáva z troch chronologicky nadväzujúcich častí; *schema matching*, *schema mapping* a *schema transformation*, ktorá je posledným krokom a predstavuje transformáciu dát (najčastejšie využívaným je proces ETL – Extract, transform and load) podľa definovaných transformačných pravidiel.

2.2 Schema matching

Ide o hľadanie sémanticky príbuzných objektov (tried prvkov, atribútov), teda tých, ktoré významovo spolu súvisia alebo navzájom korešpondujú, napr.:

Cesta – Ulica
Železnica – Dráha
Spevnený – Asfaltový

Existuje niekoľko reprezentácií, ako možno tento proces zabezpečiť a podrobne sú vysvetlené v HUMBOLDT Consortium (2006 – 2011):

- *Databáza pojmov* je najjednoduchšia a obsahuje základné sémantické jednotky pre rôzne oblasti. Tvorí ju zoznam jednoslovných podstatných mien.
- *Slovník* obsahuje definície a výslovnosť. Predstavuje spojenie medzi pojmami a ich rozličnými definíciami.
- *Thesaurus* je zoznam slov zoskupených podľa podobnosti a významu (napr. synonymá a antonymá).
- *Ontológie* slúžia na systematický (často i hierarchický) popis významov. Nás zaujíma iba počítačový prístup k ontológii, ale existuje i filozofický prístup.

2.3 Schema mapping

Schema mapping opisuje ako konvertovať dáta navzájom. Ide o zložitý a náročný proces, ktorý si vyžaduje nielen znalosti v oblasti harmonizácie dát, ale aj doménových expertov, ktorí poznajú zdrojové dáta a predkladajú svoje požiadavky na finálnu podobu transformovaných dát. Po vyhľadání vzťahov, podobnosti a sémantiky medzi zdrojom a cieľom nasleduje definovanie transformačných pravidiel. Medzi základné a najbežnejšie procesy tejto časti harmonizácie dát patria: *reklasifikácia*, *konverzia dátových typov* a *súradnicová konverzia a transformácia* medzi referenčnými systémami.

Reklasifikácia

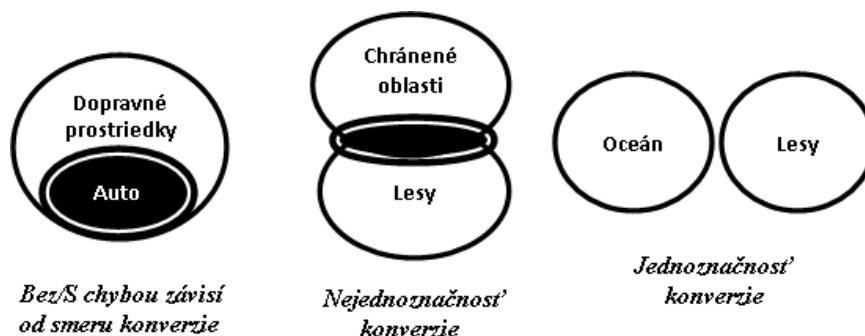
Pri reklasifikácii sú objekty zoskupované do kategórií. Kategórie predstavujú diskrétné entity, charakterizované množinou vlastností, ktoré jej členovia zdieľajú. Kategórie by mali byť jasne definované a ľubovoľná entita z klasifikovaného priestoru prislúcha jednoznačne jednej a jedinej z navrhovaných kategórií. Klasifikačné systémy pre prvky rovnakého typu môžu vychádzať napr. z dvoch faktorov „*Hodnota*“ a „*Farba*“. Prístup k definovaniu kategórií však nie úplne priamočiarly a nie vždy musí garantovať, že všetky konverzie sú prevedené správne. Tu rozoznávame niekoľko možných reklasifikačných chýb (obr. 2).

Konverzia dátových typov

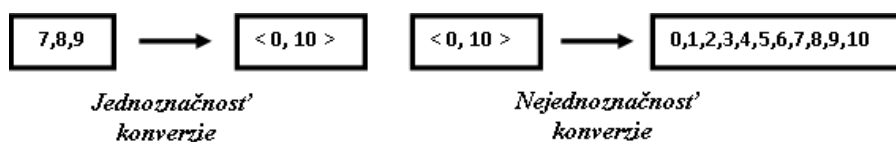
Konverzia dátových typov predstavuje konverziu jedného typu reprezentácie do iného. V mnohých prípadoch je možné použiť automatickú konverziu, avšak nie vždy je zaručená stopercentná bezchybnosť konverzie dát. Je veľmi dôležité analyzovať riziko degradácie kvality dát ešte pred samotným spustením konverzie. Riešením je v takýchto prípadoch vyladenie konverzie pomocou doplnenia krokov, tzv. „preprocessing“.

Najpoužívanejšími základnými dátovými typmi sú: Integer, Real, Date, Char, Varchar, Enum (interval), ďalej sú priestorové typy (napr. Point, Linestring, Polygon) a iné. Na ozrejmienie problematiky uvádzame vzájomné konverzie dátových typov na príkladoch (obr. 3 a 4).

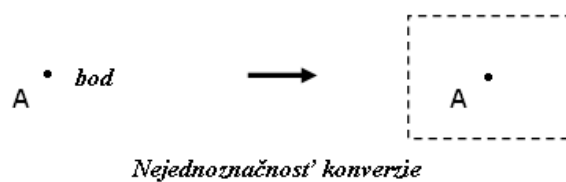
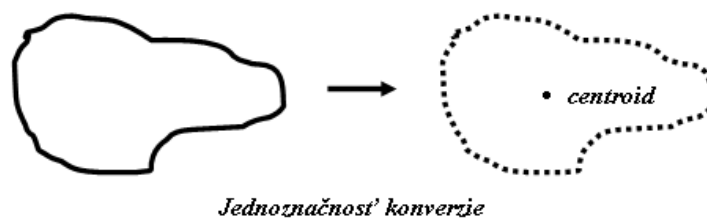
Pri konverzii bodu na polygón je bežným riešením použitie štandardizovanej predlohy polygónu určených rozmerov s centriodom umiestneným v bode.



Obr. 2 Chybovosť reklasifikácie



Obr. 3 Konverzia dátového typu „Integer“ a „Interval“



Obr. 4 Konverzia priestorových dátových typov „Bod“ a „Polygon“

Súradnicová konverzia a transformácia

Referenčným systémom ustanoveným pre horizontálnu zložku priestorových dát je podľa INSPIRE (2007) Európsky terestrický referenčný súradnicový systém 1989 (ETRS89). Pre oblasti mimo tohto referenčného rámca je preferovaný Medzinárodný terestrický referenčný systém (ITRS). Podrobne sa tejto téme venuje ISO norma pod názvom ISO 19111:2007 Geographic Information – Spatial referencing by coordinates (ISO 19111, 2007). Pre vertikálnu zložku je ustanovený Európsky vertikálny referenčný systém (EVRS) pre dáta spadajúce do geografickej oblasti EVRS.

Na výpočet elipsoidickej šírky a dĺžky sa používajú parametre elipsoidu Geodetic Reference System 1980 (GRS80) a na výpočet rovinných súradníc príslušné kartografické zobrazenie, v prípade Slovenska je to tzv. Křovákovo zobrazenie s použitím Besselovho elipsoidu.

Na základe požiadaviek a odporúčaní smernice INSPIRE (2007) použitie kartografických zobrazení závisí od nárokov na presnosť výstupnej mapy a jej mierku. Na základe toho pre systém ETRS89 sú odporúčané tri zobrazenia, dostupné v transformačných službách:

- „Lambert Azimuthal Equal Area“ (ETRS89-LAEA) vhodné pre priestorové analýzy a paneurópske zobrazenie dát bez plošného skreslenia,
- „Lambert Conformal Conic“ (ETRS89-LCC) pre konformné paneurópske zobrazenie v mierkach menších alebo rovných ako 1 : 500 000,
- „Transverse Mercator“ (ETRS89-TMzn) pre konformné paneurópske zobrazenie v mierkach väčších ako 1 : 500 000.

2.4 Edge matching

Nejedná sa o novú problematiku, ale o klasické napájanie geometrických dát používané napríklad pri digitalizácii mapových listov analógových máp. Teraz sa tento problém preniesol na hranice dátových vrstiev, kde je nutné dodržiavať nasledujúce podmienky:

- napojenia geometrie nemožno opravovať jednoduchým posunom (môžu vznikajú nové problémy),
- hladký priebeh geometrie,
- 90° uhly rohov (budovy).

Často sa už tieto problémy dajú vyriešiť automaticky v softvéroch alebo službách pomocou účelovo vyvinutých algoritmov. Project Humboldt (HUMBOLDT Consortium, 2006 – 2011) ponúka napr. „Edge Matching Service“, ktorú je možné implementovať ako službu WPS (Web Processing Service).

2.5 Ďalšie operácie používané v procese harmonizácie dát

Existuje mnoho ďalších operácií, ktoré vstupujú do procesu harmonizácie dát. Uvádzame zoznam operácií riešených v špecifických prípadoch počas transformácie zdrojových dát do požadovanej cieľovej štruktúry:

- filtrovanie dát – používa sa na dosiahnutie menšieho počtu bodov ich odstránením, keď je potrebné ich hustotu znížiť,
- prevzorkovanie dát – umelo zvýšená hustota bodov (opak filtrovania),
- „address matching (geo-coding)“ – konverzia medzi adresou a súradnicami,
- transformácia medzi časovými referenčnými systémami,
- viacmierková reprezentácia,
- topológia – rozdielne pravidlá a nové obmedzenia,
- viacjazyčnosť – preklady medzi rozdielnymi jazykmi,
- terminológia a taxonómia – rozdielne klasifikačné systémy.

Záver

Cieľom príspevku bolo priblížiť základné princípy, postupy a procesy v súčasnosti veľmi aktuálnej témy harmonizácie dát. Tejto oblasti spracovania priestorových dát sa dnes venuje veľké množstvo projektov a organizácií v zahraničí aj na Slovensku.

Je publikovaných množstvo dokumentov, noriem a špecifikácií, ktoré definujú a riadia prácu s priestorovými dátami tak, aby zabezpečili čo možno najväčšiu mieru interoperability medzi priestorovými dátami. Toto zhrnutie základných pojmov dáva predpoklad k správne pochopeniu problematiky, ktorá je veľmi široká a nie vždy jednoducho interpretovateľná, keďže každý prípad je nutné riešiť osobitne a pristupovať k nemu vždy podľa charakteru dotknutej dátovej domény pracujúcej s priestorovými dátami.

Téma harmonizácie dát je zložitý proces, o čom svedčí i množstvo aspektov, ktoré do tohto procesu vstupujú. Existuje viacero možností ako definovať koncept harmonizácie dát v závislosti na tom, ktoré aspekty sú v centre pozornosti. Harmonizácia dát nie je novým problémom. Mnohé riešenia boli vyvinuté vo všeobecnom svete informačných technológií a následne upravené na špecifické „priestorové“ riešenia. V súčasnej dobe existujú a sú k dispozícii mnohé nástroje na rieše-

nie jednotlivých krokov v procese harmonizácie dát. Napriek tomu nie je dnes vytvorený nástroj, ktorý by pokrýval všetky procesné kroky. Z tohto dôvodu sa už vôbec nedá hovoriť o automatizovanom procese zabezpečenom používateľsky jednoduchým a prívetivým riešením. Otvorenou otázkou ostáva aj harmonizácia dát „nových“ oblastí využívajúcich priestorové dáta, ktoré doposiaľ nemajú presnú definíciu v projekte INSPIRE. Problematika tvorby Národnej infraštruktúry pre priestorové informácie je zatiaľ úzko naviazaná na smernicu INSPIRE, avšak jej ďalšie budovanie a rozširovanie otvára možnosti implementácie širokého spektra ďalších dátových domén.

Všeobecnou tendenciou vo vývoji infraštruktúr priestorových informácií je rozširovanie dátovej základne o ďalšie typy dát, ktorých prepojenie s ich priestorovou zložkou dáva pridanú hodnotu dátam a umožňuje vykonávať priestorové dopyty a analýzy, čím napomáha samotnému využitiu dát a lepšiemu fungovaniu v rámci celej domény.

Literatúra

- Drafting Team "Data Specifications" (2010). *D2.5: INSPIRE Generic Conceptual Model, Version 3.3*. [cit. 2012-06-29]. Dostupné na: <http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3_3.pdf>
- FICHTINGER, A. et al. (2011). *Data Harmonisation Put into Practice by the HUMBOLDT Project*, International Journal of Spatial Data Infrastructures Research, Vol. 6, 234-260
- HUMBOLDT Consortium (2006-2011). *The Humboldt Framework, Humboldt Training Package*. [cit. 2012-06-29]. Dostupné na: <<http://www.gisig.it/humboldt/training/>>
- INSPIRE (2007). *SMERNICA EURÓPSKEHO PARLAMENTU A RADY 2007/2/ES* zo 14. marca 2007, ktorou sa zriaďuje Infraštruktúra pre priestorové informácie v Európskom spoločenstve (Inspire). [cit. 2012-06-29]. Dostupné na: <<http://inspire.gov.sk/Upload/documents/smernica.pdf>>
- ISO 19111 (2007). *ISO 19111:2007 Geographic Information – Spatial referencing by coordinates*.
- SDI Cookbook (2009). *Spatial data Infrastructure* (Last revision on 14th april 2009 by Greg Yetman). [cit. 2012-06-29]. Dostupné na: <http://www.gsdioc.org/GSDIWiki/index.php/Main_Page>

S u m m a r y

The basic principles and operations of spatial data harmonisation

This paper presents the basic information about actual theme related to spatial data and efforts to solve a big issue of nowadays, the requirement of their mutual interoperability. Theme and process of interoperability elimination is called spatial data harmonisation. One of the results of European Union (EU) and its requirement of sharing data and information is established project INSPIRE and its directive (INSPIRE, 2007) that covers all EU countries. In this directive are defined the basic principles of spatial data harmonisation which make a base for following activities on national and international level. Specialised project Humboldt for data harmonisation topics and development was created in 2006 and rise up to establishment of Humboldt consortium (2006-2011) associated with many countries from EU. Project was finalised in 2011 and results are fully free to access and download from web page (<http://www.gisig.it/humboldt/training/>) and include a lot of experiences of testing projects of spatial data harmonisation.

One of my main focuses was to outline basic operations as parts of spatial data harmonisation process. The main is "schema translation" which covers many partial operations like the feature reclassification and its fallibility as present of Fig. 2, the mutual conversion between data formats based on using GML exchange format (Fig. 1), the conversion between data types (Fig. 3 and Fig. 4), finding of semantically related objects, spatial data transformation between coordinates systems and the final result of physical data transformation based on defined schemas and transformation rules.

Operation that is not fully new and was used many years ago when information technologies were firstly used for matching of spatial datasets and digital maps is called "edge matching". Today we can choose between many tools that completely cover these matching issues automatically.

This paper gives a summary of basic concepts to correctly understand the issue of spatial data harmonisation, which is very broad and not always easy to interpret. Each case of spatial data harmonisation must be dealt with uniquely and always access it by nature of the domain working with spatial data. In coming days more and more domains are confronted to solve the elimination of their data heterogeneity what is the main purpose of data harmonisation. That is why this research is so attractive and we will continue with developing

of this theme especially for thematic domains not included in INSPIRE themes, as is the protection of culture-historical heritage, which "fights" also a lot with problem of heterogeneous spatial data.

Fig. 1 Schema of various data formats harmonisation

Fig. 2 Reclassification fallibility

Fig. 3 Conversion of data type "Integer" and "Interval"

Fig. 4 Conversion of spatial data types "Point" and "Polygon"