

# APLIKÁCIA GEOINFORMAČNÝCH SYSTÉMOV V ŠTATISTIKE

Dagmar Kusendová

## Súhrn

Cieľom článku je ozrejmiť vybrané aspekty spojenia štatistiky a geoinformačných technológií resp. geoinformačných systémov a priblížiť relevantné pojmy, koncepcie a aplikácie spolu s výberom aktuálnych problémov a možností ich riešenia u nás.

Prvá časť uvádza do problematiky geoinformačných technológií a ich služieb, približuje základné geoinformačné termíny spolu s náčrtom histórie vývoja GIS a potenciálu geografických informácií v kontexte potrieb štatistickej praxe a výskumu.

Ďalšia časť sa zaoberá priestorovými štatistickými jednotkami a ich základnými vlastnosťami, metódami geografického referencovania a kódovania štatistických údajov v kontexte interoperability, štandardizácie a unifikácie priestorových štatistických systémov a ich aplikácie v prostredí geoinformačných technológií. Záver poskytuje stručný prehľad vybraných geografických údajových báz a medzinárodných geoinformačných projektov, ktoré majú väzbu na súčasné potreby a aktivity európskej (EUROSTAT) a našej štatistiky.

## Úvod

Informácie o priestorovej situácii sú významné pre rôzne rozhodnutia spojené s priestorovým kontextom. Súčasná informačná technológia poskytuje nové možnosti pre zhromažďovanie, riadenie a prezentáciu geografických informácií, a to najmä vďaka rozvoju špecializovaných *geoinformačných technológií* (GIT), resp. ich súčasti — *geografických informačných systémov* (GIS). Spojenie štatistiky s GIT a GIS je veľmi perspektívne vzhľadom na ich prienik v oblasti zberu, spracovania, správy, analýzy a distribúcie štatistických geografických údajov. Cieľom článku je ozrejmiť vybrané aspekty tohto spojenia a priblížiť relevantné pojmy, koncepcie a aplikácie spolu s výberom aktuálnych problémov a možností ich riešenia u nás.

## 1. GEOINFORMAČNÉ SYSTÉMY A GEOGRAFICKÉ INFORMÁCIE

Priestorové, resp. *geografické informácie* **poznámka 1** (GI) sa v praxi najčastejšie spájajú s mapami, ktoré znázorňujú rôzne objekty a javy lokalizované na povrchu Zeme. Tradičnú tvorbu analógových (papierových) máp postupne nahrádza automatizovaná digitálna tvorba s použitím GIT, pričom napr. výsledky populačných cenzov sú často prezentované pomocou aplikačných programových systémov *desktop mapping GIS* (programy ArcView, MapInfo, GeoMedia apod.). Prostredníctvom siete Internet sú „on-line“ vizualizované napr. výsledky volieb s využitím geoinformačných funkcií *World Wide Web Map Serverov*. Virtuálna počítačová sieť spolu s verejnými geografickými bázami poskytuje širokej používateľskej obci geoinformačné služby pomocou ktorých sa dajú generovať vlastné tematické mapy, získavať rôzne GI atď. Vzniká nová vedná disciplína — *Geographic Information Science*, ktorá sa venuje teoretickým a metodickým aspektom spracovania GI pomocou GIS [1]. GIS sa najskôr vyvíjali v rámci geovedných a technických disciplín (geodézie kartografie, geografie, matematiky, ...) ako ich technologický nástroj. Dnes sa využívajú, resp. sami vytvárajú systémové spojenie geografických metód výskumu a informačných technológií. V najširšom význame sú GIS systémovo organizovanou informačnou technológiou pre tvorbu, správu, integráciu a poskytovanie GI.

Účel každého GIS vymedzuje jeho štruktúru, ktorá sa skladá z troch hlavných súčastí, a to: *technickej, údajovej a organizačnej*, pričom údaje tvoria kľúčovú súčasť a organizačná predstavuje ľudský faktor reprezentovaný riadiacimi, obslužnými a používateľskými zložkami. Organizácie, resp. podniky zvyčajne používajú GIS na zlepšenie organizácie vlastných operácií, pričom sa geografické údaje/geodaje zbierajú, udržiavajú, aktualizujú centrálné a poskytujú sa rôznym používateľom v rámci alebo aj mimo organizácie. Rôzne štúdie ukázali, že lokálna správa znamená viacnásobné udržiavanie veľmi podobných údajov za cenu veľkých nákladov, ktoré sa dajú redukovať práve pomocou technológií GIS pri aplikácii nutných organizačných pravidiel. Všetky súčasti GIS musia byť

vzájomne prepojené do jedného celku, ktorého funkčnosť podmieňuje vhodné technické (hardvérové a softvérové) vybavenie všetkých úrovní informačného systému, dôsledná analýza vlastností údajov vo vzťahu k objektu a funkciám systému, špecifikácia integrácie údajov medzi jednotlivými úrovňami systému spolu s jeho organizačným zabezpečením.

### Vývoj a interoperabilita GIS a GI

V praxi sa aplikujú GIS v troch hlavných oblastiach, a to: *evidencie, správy a plánovania*, pričom každá z nich má vlastný vývoj počítačového spracovania priestorových dát.

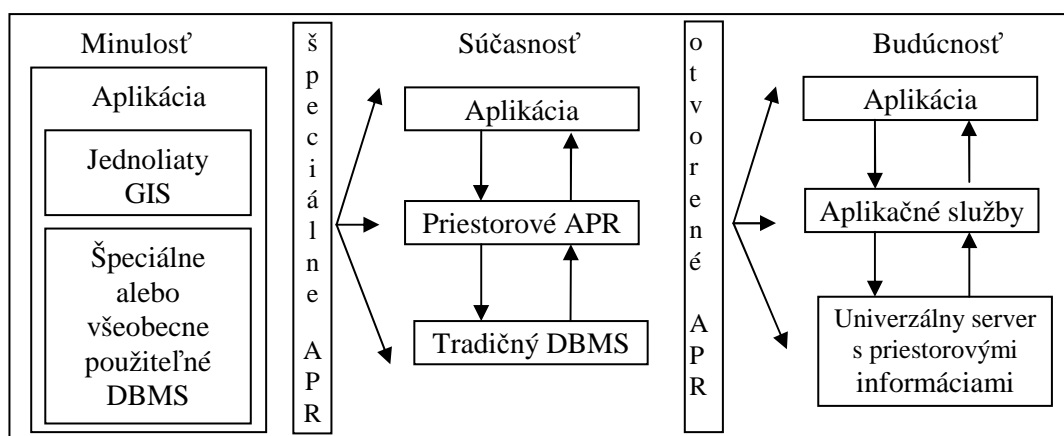
Dnes sú GIS zrelou informačnou technológiou, ktoré predávajú, resp. zavádzajú špecializované spoločnosti vo forme viacúčelových modulových počítačových systémoch (ESRI, Intergraph, Siemens, Unisys a iné). V počiatkoch sa trh s GIS orientoval na univerzálne GIS (splňajúce rôzne požiadavky používateľov, integrujúce množstvo funkcií z iných aplikačných oblastí) a národné trhy so špecifickými požiadavkami na administratívne úkony, lokálnu podporu (jazyka, školení, ...) so zameraním sa na jednocúčelové aplikácie (správa katastra, logistika apod.). Vývoj monolitických viacúčelových GIS vyrábaných jednou spoločnosťou dosiahol svoj zenit v 90. rokoch 20. storočia, pričom pridaná funkcionálna nebola bezproblémová — vyžadovala zvýšený podiel organizačnej a programátorskej práce.

Súčasný trend vývoja programov GIS je zameraný na vývoj a predaj komponentov splňajúce špecifikácie „otvoreného GIS – OPEN GIS“. Väčšina z nich má obmedzenú množinu funkcií, ktoré sú však veľmi dobre prepracované, pričom organizáciu dát, operácií a vizualizáciu výsledkov ponechávajú naďalej prostrediu komplexných GIS.

V súlade s aktuálnym vývojom informačných technológií sa v GIS presadzujú otvorené programové platformy, ktoré sú významným integrátorom geograficky referencovaných štatistických údajov a geoštatistických metód a techník. Výrazný podiel na rozširovaní trhu s digitálnymi geoúdajmi má rozvoj interoperability GIS – t. j. údajová a funkčná zdieľateľnosť a prepojitelnosť rôznych programových aplikácií založená na voľnej výmene údajov a ich spracovaní v reálnom čase bez ohrozenia straty informačnej hodnoty. Rozpracúvajú sa spôsoby práce v počítačových sieťach, prístupu a zdieľania údajov, tvoria sa nové programové aplikačné rozhrania a metaúdajové štruktúry.

obr.1

Interoperabilita a s ňou spojená štandardizácia a normalizácia geografických údajových štruktúr je dlhodobý cieľ celej geoinformačnej komunity (databázových distribútorov, národných mapovacích a štatistických služieb, vojenskej administratívy, ...), ktorý sa nemôže dosiahnuť bez vzájomnej spolupráce všetkých zainteresovaných strán, či už vo verejnom alebo súkromnom sektore.



Obr. 1 Vývoj technológií GIS

(APR - Aplikačné programové prostredie, DBMS – systém riadenia bázy dát, GIS –geografický informačný systém)

Vývoj GIT smeruje ku GIS, ktoré sa dajú integrovať s inými prostriedkami na spracovanie dát, kde GIS tvoria len jeden komponent, ktorý spracúva a exportuje GI do iných subsystémov alebo procesov, resp. komunikujú cez otvorené počítačové rozhrania s inými (expertnými alebo znalostnými) systémami.

Ekonomická sila a potenciál GI sa len odhaduje. Predpokladá sa, že napr. v rokoch 2005 až 2010 by sa malo vynaložiť na tvorbu a správu GI v prostredí GIT približne 100 až 500 euro/osobu/rok, pričom niektoré odhady uvádzajú, že v tomto období by mala byť jedna osoba z tisíc vzdelaná a školená ako profesionál v GIS [2].

## 2. GEOGRAFICKÉ INFORMAČNÉ SYSTÉMY A ŠTATISTIKA

GIS prenikajú do štatistiky, ale aj iných odborov, v dvoch základných formách, a to ako súhrn geoinformačných technológií a metód, ktoré sa uplatňujú najmä pri tvorbe, údržbe a distribúcii geografických údajových báz najčastejšie s humánogeografickým, demogeografickým alebo environmentálnym zameraním. Ide v podstate o tvorbu štandardizovaných geoinformačných, resp. geografických distribučných údajových báz v medzinárodnom a národnom meradle spolu s tvorbou príslušných obslužných programov s vybranými funkciami programov GIS typu *desktop mapping* alebo *mapping browser* (MapInfo, ArView, GeoMedia, ...), ktoré umožňujú používateľom získať relevantnú a aktuálnu GI. Ide zväčša o programové aplikácie zamerané na jednoduchú štatisticko-priestorovú analýzu a kartografickú (resp. multimediálnu) prezentáciu geografických údajov a informácií. Predpokladom použitia tohto typu programov a aplikácií GIS v štatistickej praxi je existencia vhodných geografických báz štatistických údajov.

Vďaka priestorovo-spojitému charakteru fyzikogeografických javov, spolu s rozvojom *dial'kového prieskumu Zeme* (DPZ), vzniklo množstvo digitálnych geografických báz predovšetkým pre správu prírodných (environmentálnych) zdrojov. Správa ľudských zdrojov a "geoinformatizácia" v sociálno-ekonomických vedách má svoju relatívne oddelenú cestu vývoja, ktorá je prirodzeným dôsledkom rozdielného charakteru a pojatia, a tým aj modelovania objektov a javov ich výskumu [1]. Je snaha prepojiť obidve cesty vývoja formou identifikácie priestorových interakcií humánnych a environmentálnych aktivít, ktoré by mohli riešiť problém získavania aktuálnejších a (polohovo) presnejších GI pri analýze rozmiestnenia populácie na základe interpretácie údajov z DPZ [3] apod.

Štatistické údaje sa často ponúkajú spolu s programami *desktop GIS* vo forme "ukážkových" údajových súborov vhodných pre spracovanie a analyzovanie v danom programe. Nedostatkom týchto súborov, ak nemajú veľmi generalizovaný obsah, je ich aktualizácia. Tú zabezpečuje buď predajca, alebo sám používateľ programu. Bohužiaľ táto forma aktualizácie často nie je zárukou korektných (skutočne pravdivých a aktuálnych) údajov. Výnimku tvoria distribútori, ktorí sú správcami alebo spolupracujú so štátnymi garantmi demogeografických, resp. sociálno-ekonomických štatistických údajov. V ostatnom čase prenikajú GIT aj do štatistických programov, akým je napríklad *System SAS* (<http://www.sas.com/>), formou špecializovaných modulov (SAS/GIS), ktoré poskytujú jednoduché a efektívne spojenie geoinformačných a štatistických postupov a metód v jednom programovom systéme [4]. Odstraňuje sa tak problém s výmenou (importom/exportom) a geovizualizáciou údajov, pri ktorom dochádza k degradácii ich kvality.

Rozvoj GIT významne zasiahol do tvorby a aplikácie najmä demografických štatistických údajov. Svoju históriu začal v 70. rokoch implementáciou výsledkov cenzu USA do geoinformačnej údajovej štruktúry známej pod akronymom *TIGER* (Topologically Integrated Geographic Encoding and Referencing). Tento referenčný kódovací systém Federálneho štatistického úradu pri ministerstve obchodu v USA – *United States Bureau of the Census* (USBC) sa stal často používaným prototypom cenžálnej štatistickej geografickej údajovej bázy v GIT [5].

Všeobecne má trh s GI inú dynamiku v Európe ako v USA, kde *United States Geological Survey* (USGS [poznámka 2](#)) a USBC majú mandát na zhromažďovanie a distribúciu primárnych geografických (polohopisných a výškopisných) a k nim referencovaných štatistických (demo, socio, ...) údajov. Americký trh s GI bol založený na tom, že obe uvedené organizácie transformovali svoje údaje do elektronickej podoby a sú zdarma dostupné prostredníctvom internetu len za cenu reprodukcie. Výškopis, uličná sieť, demografické a ďalšie geoštatistické údaje sú integrované do báz,

ktoré pokrývajú celý štát od malých až po veľké mierky. To všetko umožňuje experimentovanie a vývoj rôznych aplikácií GIS.

V Európe tieto údaje zhromažďujú rôzne národné agentúry, ktoré pracujú na základe iných zákonov a rôznych mandátov, pričom svoje údaje poskytujú v prevažnej miere za úplatu. Nedoriešené otázky autorstva vytvárajú bariéry ich širšieho použitia spolu s ďalšími príčinami, z ktorých podľa [6] sú najdôležitejšie tieto:

- 1) *nedostatok povedomia potencionálnych používateľov GI,*
- 2) *nedostupnosť geografických dát a informácií,*
- 3) *rozdrobenosť trhu s GI.*

U nás poskytujú digitálne geograficky referencované štatistické údaje komerčné organizácie, ktoré buď predávajú programy GIS (napr. firma ArcGEO program ARCGIS, <http://www.arcgeo.sk>) alebo špecializované produkty GIS (InfoMapa – PJSOFT, <http://www.pjsoft.cz>). Najväčšie predpoklady na obsadenie trhu s „použiteľnými“ geografickými údajmi pre geoštatistické analýzy a prezentácie majú aj v Európe najmä zástupcovia amerických výrobcov GIT, a to ESRI (The Environmental Systems Research Institute, in Redlands, California) a Intergraph (Huntsville, Alabama).

### 3. GEOGRAFICKÉ REFERENCOVANIE A KÓDOVANIE ŠTATISTICKÝCH DÁT

#### Priestorové štatistické jednotky

V štatistike je **štatistická jednotka** definovaná ako „základný a presne vecne, časovo a priestorovo vymedzený objekt pozorovania prejavu určitého štatistického javu. Priestorové (miestne) vymedzenie štatistických jednotiek znamená určenie miesta alebo územia, za ktoré sa vykonáva štatistické skúmanie“ [7].

Sociálno-ekonomické a demografické štatistické dáta sa obvykle priestorovo vzťahujú k určitému systému územných jednotiek, v ktorom má každá jednotka formálne vymedzenú polohu a rozlohu (napr. kraj, obec, urbanistický obvod apod.). Použiteľnosť referenčných systémov v praxi súvisí s ich územným vymedzením, mierkou alebo rozlišovacou úrovňou, štruktúrou a konzistenciou.

#### Územné vymedzenie

V územno-správnych systémoch, ktoré tvoria *priestorovú referenčnú bázu* väčšiny sociálno-ekonomickým štatistickým údajom, je každá štatistická jednotka jednoznačne identifikovaná štandardným názvom alebo kódom s formálne vymedzenou polohou (hranicou) a rozlohou obvykle v rámci národného kartografického systému, a to buď v analógovom tvare (papierová mapa) alebo v digitálnom tvare (digitálna mapa). Okrem územno-správnych systémov sa používajú priestorové referenčné systémy založené na zemepisných sieťach (s rôznou dĺžkou strán), centrálnych bodoch (centroidoch), poštových smerových číslach apod. V mnohých štátoch sú štatistické dáta o obyvateľstve, domoch, nehnuteľnostiach vzťahované k poštovým priestorovým referenčným systémom. V Dánsku ho napr. vytvorili na báze pravidelnej hierarchickej zemepisnej siete s rozmerom strany od 100 km do 100 m [8].

Poštové smerové čísla (PSČ) u nás nie sú zachytené v grafickej forme na žiadnom verejnosti prístupnom grafickom operáte (mape) pre celé územie Slovenskej republiky. Sú evidované len v písomnej forme v tvare kódovníkov (číselníkov) za jednotlivé obce, resp. uličné úseky väčších obcí a miest a využívajú nepriame určenie geografickej polohy (na rozdiel od priameho pomocou geodetických alebo zemepisných súradníc), ktoré je však z hľadiska určovania referencie adries nejednoznačné a neumožňuje korektne stanoviť ich geografickú lokalizáciu. Ak by sa uvažovalo u nás o zavedení štatistického PRS na báze poštových adries, napr. v kontexte tvorby *základných územných registrov verejnej správy*, tak ideálnym riešením geografickej „referencie“ štatistických údajov je zavedenie systému centroidov a traktov s geografickou lokalizáciou (súradnice bodových a čiarových segmentov vo vybranom súradnicovom systéme určené pomocou geodetických prístrojov alebo prístrojov GPS).

Charakter územných referencií štatistických údajov je rôzny, tak ako sú rôzne sledované témy. a vytvoriť univerzálny PRS, ktorý by vyhovoval väčšine u nás sledovaných štatistických ukazovateľov nie je jednoduchá úloha. Príklady z iných štátov Európy naznačujú, že ide riešiteľný problém.

### **Mierka**

V kontexte použitia GIT by mala byť mierka priestorových jednotiek v súlade s rozlišovacou úrovňou štatistických zisťovaní, resp. analýz. Rozlíšenie charakteristík obyvateľstva na úrovni zázemia obce je pre aplikácie v hraniciach sčítacích obvodov nedostatočné, ale v rámci regionálnej mierky (kraja) dostatočné. Z hľadiska mierkového rozlíšenia v územnom rozsahu Slovenska sa javí mierka 1:10 000 ako najvhodnejšia na realizáciu štatistických analýz. Väčšina sociálno-ekonomických a demogeografických aplikácií je obmedzená územným rozlíšením na úrovni obcí, ktorej táto mierka postačuje. Rozlíšenie 1:10 000 má aj digitálna základná báza (tzv. ZB GIS) jednotnej lokalizácie údajov územne orientovaných štátnych informačných systémov SR. Pomocou nej by sa mala zabezpečiť jednotná geografická lokalizácia a identifikácia taktiež štatistických údajov vytváraných jednotlivými rezortmi hospodárstva Slovenska. Prechod na väčšiu rozlišovaciu úroveň je problematické z dôvodu nevyhovujúceho stavu v oblasti aktualizácie, koordinácie a poskytovania/šírenia — diseminácie ZB GIS pre potreby praxe.

### **Štruktúra**

S mierkou úzko súvisí zväčša hierarchická štruktúra *priestorových referenčných jednotiek*, ktorá by mala zabezpečiť ich jednoznačnú skladobnosť. Teoreticky to znamená, že každá jednotka by mala byť bezo zvyšku zložená z najbližších hierarchicky nižších jednotiek (napr. územie Slovenska na najvyššej úrovni územno-správneho delenia vydeľujú hranice krajov, na nižšej okresy atď.). Podmienka skladobnosti v praxi nie je vždy splnená, v dôsledku čoho vznikajú veľké problémy pri tvorbe a aktualizácii geografickej bázy GIS [9].

### **Konzistencia**

Konzistencia priestorových referenčných jednotiek znamená, že sú významovo rovnocenné, teda ich obsah je porovnateľný, resp. sú obsahovo ekvivalentné. Pri analýze hustoty obyvateľstva sa kladie dôraz na zachovanie plošnej ekvivalentnosti územných jednotiek, zatiaľ čo pri analýze vekovej štruktúry by mal byť počet obyvateľov približne rovnaký pre každé územie. Je viacero geoštatistických metód, ktoré tieto otázky pomáhajú riešiť [10], z ktorých majú perspektívne uplatnenie najmä metódy *pravdepodobnostného mapovania* s využitím GIS, ktoré sa využívajú napr. v Českej republike v sociálnej štatistike pri riešení problematiky trhu práce [11].

### **Európska priestorová údajová infraštruktúra**

Až na niektoré malé a stredné mierky v rámci Európy donedávna neexistovala žiadna funkčná priestorová údajová báza, ktorá by poskytovala potrebné GI v dôsledku rôznorodosti vytvorených a vytváraných národných geoúdajových báz. Z toho vyplývala ich medzinárodná neporovnateľnosť a obtiažna aplikácia mimo svojho národného rámca (územia). To nastolilo aktuálnu potrebu vytvorenia kartograficky a obsahovo "bezošvej" priestorovej údajovej infraštruktúry (*Spatial Data Infrastructure* — SDI) v mierke Európy. Európske spoločenstvo (ES) už niekoľko rokov vyvíja rôzne iniciatívy vedúce k tomuto cieľu.

### **Národné územné štatistické jednotky**

Tak, ako sa vytvára európska SDI, sa postupne vytvára aj spoločná celoeurópska priestorová referenčná báza, ktorá by mala poskytovať aktuálne a porovnateľné priestorové štatistické jednotky. Riešia sa otázky medzinárodnej štandardizácie a unifikácie územných jednotiek pre regionálne štatistické analýzy, ktoré v rámci Európskej únie (EÚ) vyústili do tvorby jednotnej klasifikácie územno-štatistických jednotiek — NUTS (*Nomenclature des Unités Territoriales Statistiques*).

Národné územné štatistické jednotky, ktoré vznikli v rámci NUTS, predstavujú medzinárodný štandard súčasných priestorových referenčných jednotiek. Vznikli ako rámec pre jednotnú európsku regionálnu štatistiku a sú súčasťou rôznych štandardizovaných údajových báz, ktoré využíva EÚ pre svoje štatistické monitorovanie a sociálno-ekonomické analýzy na úrovni európskych regiónov a pre potreby prípravy, realizácie a hodnotenia regionálnej politiky [12]. Od roku 1988 je klasifikácia NUTS používaná v legislatíve EÚ hlavne na riešenie úloh spojených s dotačnou politikou zo štrukturálnych fondov únie. Vymedzenie jednotlivých úrovní NUTS, ktoré majú čisto štatistický charakter, by malo odpovedať metodickým princípom a štandardom Európskeho štatistického úradu

— EUROSTAT, t. j. rešpektovanie veľkosti územných jednotiek podľa odporúčaní EÚ a princípov skladobnosti vo vzťahu k administratívno-správnemu deleniu štátu, podobnosť jednotiek nižšieho rádu vytvárajúcich jednotku vyššieho rádu a uplatňovanie právomocí zastupiteľských zborov a orgánov štátnej správy na danej hierarchickej úrovni.

Z pôvodných troch úrovní NUTS sa počet úrovní NUTS rozšíril na šesť úrovní (NUTS 0 až NUTS 5 a LAU 1 a 2 **Poznámka 3**), ktoré predstavujú jednotlivé veľkostné skupiny vyhovujúce väčšine požiadaviek kladených na takýto typ systému. Ide o skupiny presne identifikovaných a vydelených územných jednotiek v príslušnej mierke (od úrovne štátu až po obce) s hierarchickou štruktúrou. Počet obyvateľov a rozloha priestorovej jednotky, prislúchajúce danej úrovni NUTS, sú dobrými indikátormi správnosti ich vymedzenia na nižších úrovniach. V praxi sa však metodické princípy klasifikácie dôsledne neuplatňovali, pretože sa pri jej tvorbe vychádzalo z už existujúcich (národných) systémov. To viedlo, spolu s rozšírením úrovní NUTS, v niektorých prípadoch k výraznej diferenciacii vyčlenených územných jednotiek a k ich nekonzistencii. Príkladom nejednotnosti úrovní NUTS bolo napr. ich prvotne navrhnuté členenie u nás a v Českej republike pre potreby projektu *Seamless Administrative Boundaries of Europe* — SABE [13], kde na úrovni NUTS 3 by boli u nás okresy, a v ČR kraje [14] s využitím aktuálne platných národných priestorových jednotiek. V dôsledku neustálych rozdielov v NUTS bolo iniciované nariadenie Európskeho parlamentu o zavedení všeobecnej klasifikácie územno-štatistických jednotiek NUTS [15], kde je určené rozmedzie počtu obyvateľstva pre NUTS 3 v rozmedzí 150 – 800 tisíc, NUTS 2 do 3 mil. a NUTS 1 do 7 mil. obyvateľov. Hoci sa problém s klasifikovanou definíciou sústavy územných jednotiek pre štatistiku zdá okrajový, môže zohrať dôležitú úlohu v procese pripravenosti Slovenska začleniť sa do európskych priestorových štruktúr.

#### **4. GEOGRAFICKÉ INFORMAČNÉ SYSTÉMY V EURÓPSKEJ ŠTATISTIKE**

Používanie GIS a GI v Komisii EÚ začalo environmentálnym projektom CORINE a po vytvorení Európskej agentúry pre životné prostredie (EPA) sa objavila reálna potreba GI a požiadavka na služby GIS ako nástrojov integrácie údajov z rôznych zdrojov. Postupne dochádza k integrácii nielen štatistických údajových báz na celoeurópskej úrovni v rámci Európskeho štatistického úradu, ale v spolupráci s národnými mapovými službami aj ich lokalizačných (terestrických, geodetických, kartografických) báz prostredníctvom rôznych celoeurópskych geoinformačných projektov **Poznámka 4** (*GISCO*, *SABE*, *INSPIRE* a ďalších), ktorých účelom je zaistenie pohotovej dostupnosti k relevantnej informácii pre ľubovoľnú poznávaciu alebo rozhodovaciu činnosť pomocou komplexu technologických, ekonomických, právnych a ďalších nástrojov [16].

##### **GISCO (Geographic Information System of the European Commission)**

Rozdiely v postupoch aktualizácie nomenklatúr, mierok, zobrazovacích systémov a používania metaúdajov boli hlavné dôvody, ktoré viedli k integrácii geografických údajov v európskej štatistike. Z hľadiska európskej štatistiky je GISCO najvýznamnejším projektom GIS, ktorého cieľom je tvorba a aktualizácia GI pre potreby ES a EUROSTAT [17]. Základ projektu vytvorili údaje z bázy REGIO, v ktorej sú štatistické informácie (demografické, ekonomické, o (ne)zamestnanosti, poľnohospodárske, ...) mapované na regionálnej (národnej) úrovni jednotlivých členov EÚ. Geografická báza GISCO má relačno-entitovú štruktúru údajov digitálne formátovaných pre priestorový model programu ARC/INFO, resp. ArcGIS. Súčasťou projektu bolo vytvorenie “štatistického topografického programu REGIOMAP“, ktorý umožňuje zobrazovanie a komparáciu demogeografických a sociálno-ekonomických údajov od roku 1960. Slovensko do REGIO databázy je zahrnuté od roku 1996.

Dnes je údajová báza GISCO používaná vo funkcii priestorovej referenčnej bázy, ktorá obsahuje dva typy základných údajov, a to: topografické a tematické prepojené relačnými identifikátormi. Nosným programom projektu je tvorba a údržba geografickej referenčnej bázy, priestorové analýzy, vývoj architektúry GISCO, aplikácie a spolupráca najmä s projektom *INSPIRE* (Infrastructure for Spatial Data in Europe) s cieľom vytvoriť prototyp informácií nielen o životnom prostredí. V súčasnosti je asi 12 tém rozdelených do 40 tematických vrstiev v mierke od 1:100 000 do 1:10 mil.

Doposiaľ má GISCO limity, ktoré vyplývajú z nedostatku špecialistov GIS, neprepojenia údajových vrstiev, nepriameho prístupu k údajom a ich nedostatočnej a problematickej aktualizácie. Novinkou je nová technologická architektúra GISCO podmienená zavedením nového softvéru (Oracle 8i 6+) a objektového údajového modelu, zaznamenávaním časových radov, prezentáciou a rozširovaním údajov prostredníctvom internetu atď. V softvéri sú však nedoriešené špeciálne interakcie medzi rozdielnymi komponentmi programu ArcGIS a databázou Oracle. V aplikáciách sa rieši diseminácia údajov, tvorba tematických máp v rôznych mierkach, integrácia a generalizácia údajov (NUTS a LAU) pre potreby mapovej vizualizácie.

V priebehu roka 2002 realizačný tím GISCO dokončoval aplikáciu WWW pre NUTS, ([http://europa.eu.int/comm/eurostat/ramon/nuts/splash\\_regions.html](http://europa.eu.int/comm/eurostat/ramon/nuts/splash_regions.html)) a spolu s ďalšími partnermi (EUROGeographics **Poznámka 5**) realizoval seminár pre krajiny EÚ a kandidátske krajiny PHARE zameraný na hodnotenie prieskumu o používaní GI a GIS v jednotlivých krajinách, ich pripravenosti integrovať sa do európskeho štatistického systému a na rozvoj spolupráce a výmenu skúseností medzi národnými štatistickými úradmi, národnými mapovými agentúrami a EUROSTAT [18].

### **INSPIRE (Infrastructure for Spatial Data in Europe)**

Vytvorenie koherentnej infraštruktúry priestorových údajov v Európe je hlavným cieľom iniciatívy INSPIRE založenej na presadzovaní týchto princípov:

- geografické údaje, resp. informácie by sa mali získavať/zberať len raz a efektívnym spôsobom,
- mala by byť možná ich kombinácia z rôznych zdrojov a z celého územia Európy s možnosťou ich šírenia (distribúcie/poskytovania) mnohým používateľom a aplikáciám,
- GI by sa mali zbierať na jednej úrovni s možnosťou generalizácie a detailizácie, t. j. poskytovania v rôznych úrovniach (mierkach) pre výskum, strategické účely apod.
- GI sú potrebné pre efektívne riadenie na všetkých úrovniach, pre transparentnú, otvorenú a zrozumiteľnú prezentáciu (vizualizáciu) v používateľsky prívetivom prostredí.

Dôležité sú procesy vedúce k harmonizácii údajov, ich definovaniu, tvorbe procedúr hodnotiacich ich kvalitu a metaúdajových systémov, štandardizácii, vývoju a výskumu a licencovaniu GI. Hlavným cieľom INSPIRE je dosiahnuť taký stav, ktorý by umožnil ľahký bezbariérový prístup ku geografickým informáciám, ktoré by sa dali „zdieľať“ a kombinovať kontinuálne (bez švov). Aktuálny vývoj v tvorbe celoeurópskej priestorovej infraštruktúry prezentujú prvé kroky k vývoju implementačnej stratégie, zahájenie legislatívnych procedúr začiatkom roka 2003, prijatie legislatívy s predpokladom v roku 2005 a štart implementácie v roku 2006, zapojenie národných štatistických úradov do pripomienkovania projektu prostredníctvom konzultácií cez Internet [19].

### **Súčasná aktivity Európskeho štatistického úradu v oblasti GIS**

Rozvoj GIS v EUROSTAT siaha do obdobia okolo roku 1990 (projekt GISCO), od roku 1998 začína nadobúdať reálnu podobu. Zavedenie GIS v EUROSTAT je prioritou, ktorá by mala priniesť pridanú hodnotu podobne ako je tomu v Severnej Amerike.

V roku 2003 sa v sídle EUROSTAT konalo od roku 1998 už piate pracovné zasadenie s názvom „*Geografické informačné systémy pre štatistiku*“. Ide o spoločné zasadanie národných štatistických úradov a národných mapových agentúr, ktoré hodnotia vývoj projektov GIS za uplynulý rok a podrobne diskutujú ďalší plánovaný vývoj. Porovnávajú sa aktivity v oblasti GIS v rámci ES. V správe zo zasadnutia sa okrem iného napr. konštatuje, že očakávania EUROSTAT ohľadne zapájania nášho štatistického úradu (ale aj ďalších) do jednotlivých projektov GIS sa postupne naplňujú.

Z tohto hľadiska je zaujímavé sa pozrieť na aktivity *Českého štatistického úradu* (ČSÚ) v tejto oblasti. Tie začínajú začiatkom 90. rokov minulého storočia, pričom ich druhá polovica je spojená s intenzívnou všestrannou medzirezortnou prípravou a spoluprácou so štatistickými inštitúciami Kanady. Súčasťou budovania technologického, personálneho a údajového jadra GIS ČSÚ ČR bola rozsiahla údajová príprava tvorby (v rokoch 1999/2000) a aktualizácie (r. 2000/2001) tematických mapových vrstiev [20]. ČSÚ sa v roku 2001 pripojil k verejnemu metainformačnému systému MIDAS zameraného na geoinformácie, ktorý má v súčasnosti štatút *odborového katalógu zdrojov verejnej správy ČR* [21]. V kontexte tvorby *Informačného systému verejnej správy ČR* realizuje tvorbu

rôznych základných štatistických priestorových registrov [22]. ČSÚ používajú rovnakú technológiu produktov GIS ako EUROSTAT, a to produkty firmy ESRI.

Obsahovo je príprava GI a GIS na pôde EUROSTAT charakteristická tým, že je rozvinutá do:

- b) vlastnej vedecko-výskumnej bázy a spolupráce na externých projektoch v rámci EÚ a ich koordinácie najmä s národnými mapovými agentúrami,
- a) vedecko-výskumnej oblasti v podobe desiatok realizovaných projektov, ktoré zahrňujú oblasti výskumu štatistických jednotiek (NUTS, LAU) a ich priestorových sietí, porovnateľnosti a individualizácie štatistických údajov, adresovania, generalizácie mapových výstupov, rozširovania GI (informačných výstupov) a „mapserverov“, mapovania Európy, metaúdajov, legislatívy a ďalších oblastí.

Do oblasti NUTS je v súčasnosti najviac geoinformačne orientovaných aktivít. Dôležité sú aktivity v oblasti konverzie metadát podľa normy ISO 19115. Od januára 2002 sú informácie týkajúce sa NUTS a štatistických regiónov v sieti Internet. Používateľský prístup k NUTS požaduje maximálnu stabilitu číselníkov NUTS. Preto sa dajú meniť len každé tri roky s dvojročným prechodným obdobím. Cieľom je redukovať rozdielne veľkosti jednotlivých stupňov NUTS čo do rozlohy a počtu obyvateľov. To si vyžiada zmeny v NUTS niektorých krajín (Nemecko, Taliansko, Fínsko, Portugalsko, Španielsko). Vo vertikálnom územnom reze sú štatistické údaje v EUROSTAT plošne spracovávané do úrovne NUTS 3 (rozlíšenie 1:4 mil). Podrobnejšie úrovne sa testujú na pilotných územiach, pričom najvýznamnejším projektom v tomto smere je projekt TANDEM.

### **TANDEM (Towards a Common Geographical Base for Statistics across Europe)**

Projekt realizuje konzorcium zástupcov Švédska, Fínsko a Veľká Británie. Program výskumu reaguje na skutočnosť, že územím malé štatistické jednotky sú veľmi dôležité pre členské štáty EÚ. Problém diseminácie individuálnych údajov je tak rozdielny, ako sú rozdielne jednotlivé členské krajiny v súvislosti s agendou priestorových/geografických dát a analýz. Územne malé štatistické jednotky by sa mali stať tiež súčasťou referenčných dát projektu INSPIRE.

Potreba porovnateľnosti/konzistencie priestorových štatistických dát v Európe rieši výskum v dvoch súbežných vývojových smeroch, a to formou areálových/polygón (nepravidelných) a sieťových/grid (pravidelných) prístupov. Záverečná správa štúdie TANDEM o malých štatistických územiach [17] svedčí o potrebe vývoja prototypu systému, ktorý by mali používať všetky krajiny Európy, a to vo forme pravidelných aj nepravidelných priestorových jednotiek.

Konzorcium v súlade s touto stratégiou sa zameriava na vývoj oboch typov priestorových modelov, ich referenčných systémov, na tvorbu konzistentných a harmonizovaných dátových súborov, sídelných obvodov, rieši špeciálne relevantné problémy (malé štatistické obvody), realizuje štúdie vývoja urbánnej štatistiky a priestorových analýz urbanistických ukazovateľov, disemináciu dát založených na pravidelných priestorových sieťach apod.

Osobitne aktuálna je otázka spôsobov diseminácie dát/informácií prostredníctvom sektorov verejnej správy v kontexte politického programu elektronizácie spoločnosti „eEurope 2005“ s dôrazom nariadenie otázok spojených s moderným poskytovaním služieb občanom a schopnosti šírenia GI v trhových cenách a chránenej informačnej infraštruktúre.

### **Systém územných štatistických jednotiek v SR**

Otázky klasifikácie priestorových štandardov a registrov vhodných pre stanovovanie polohy a územnej platnosti štatistických údajov u nás rieši *Štátny informačný systém SR (ŠIS SR)* prostredníctvom koncepcie a realizácie *Národnej priestorovej informačnej infraštruktúry (NGII) SR*. Cieľom je vytvoriť štandardnú štruktúru územných jednotiek pre tvorbu a podporu geoinformačných systémov štátu, samosprávy a iných subjektov. Prvým krokom v tejto oblasti bolo najmä schválenie dátových štandardov ŠIS SR pre prvky identifikujúce a popisujúce územie (názov štátu, okresu, obce, katastrálneho územia a ulice, číslo popisné, počítačové a orientačné objektu, kód katastra, parcelné číslo) *Výnosom Štatistického úradu SR č. 372/1998-830 z 13.10.1998*. Vyhlásené štandardy tvoria legislatívny rámec pre vzťahnú štatistickú sieť na území SR, v rámci ktorej by nemal byť problém definovať ďalšie priestorové prvky vytvárajúce účelové štatistické siete, ktorým by sa mohol napr. stať pripravovaný *systém poľnohospodárskych produkčných blokov* a iné.



Základný rámec pre štandardizáciu priestorových jednotiek bol vytvorený už v roku 1979 *Vyhláškou 120/1979 Zb. o priestorovej identifikácii informácií*. Priestorová identifikácia v ŠIS sa definuje ako jednoznačné a nezameniteľné priestorové určenie prvkov určitej triedy údajového modelu reality pomocou územných identifikátorov. Štandardné priestorové referenčné jednotky sú označené identifikačným číslom (spravidla vo forme kódu), popísané názvom a ďalšími charakteristikami. Priestorové štruktúry a priestorové informačné jednotky ŠIS SR sa funkčne delia na: *správne, technické a sídelné* s maximálnym rozlíšením na úrovni základných územných jednotiek, územno-technických a základných sídelných jednotiek, ktorých georeferencia je zaznamenaná aj digitálne (ZB GIS a VKM – vektorové katastrálne mapy). Ide o naše najstabilnejšie referenčné jednotky v priestore a čase s reálnou možnosťou preskupovania štatistických jednotiek do úrovne LAU (obce). Pri temporálnych štatistických a lokálnych analýzach však táto úroveň nie je dostatočná najmä v dôsledku častých zmien administratívneho a správneho členenia SR.

## 5. CENTRÁLNY GEOINFORMAČNÝ ÚZEMNÝ REGISTER Štatistického úradu SR

Štatistický úrad (ŠÚ) SR, ako vrcholový úrad štátnej správy v oblasti zberu, spracovania a distribúcie štatistických údajov, by mal disponovať vlastnou centrálnou geografickou bázou dát štatistických územných jednotiek, ktorej súčasťou by bola univerzálna referenčná štatistická sieť pre stanovovanie polohy a územnej platnosti štatistických údajov — *Centrálny geoinformačný územný register ŠÚ SR (CGÚR)*. Pre tvorbu, použitie a prevádzku CGÚR je nevyhnutné vytvoriť geoinformačný systém, ktorý by systematicky zjednotil technologické prostriedky, údaje a organizáciu v rámci ŠÚ spolu s ďalšími kooperujúcimi rezortmi, a to najmä Úradom geodézie, kartografie a katastra. CGÚR vo funkcii univerzálnej štatistickej siete pre územia Slovenska by bol prvým krokom k vytvoreniu funkčného GIS v našej štatistike.

Korektná štatistická sieť je nevyhnutným predpokladom efektívneho využitia geoinformačných technológií a systémového prístupu k spracovaniu, správe, geografickej analýze a distribúcii štatistických dát, ktoré sa viažu na geografický priestor. Potreba vzniku CGÚR súvisí aj so súčasným trendom integrácie európskych geografických informačných štruktúr (INSPIRE, SABE), vývoja regionálnych štatistík Európskej únie (GISCO), ako aj implementácie metód a nástrojov priestorovej štatistiky do bežnej štatistickej praxe.

Nedostatočná pozornosť otázkam tvorby, dôslednej aktualizácie a dostupnosti národných priestorových štatistických údajových báz, kvalifikovanej definícii sústavy územných jednotiek pre štatistiku spolu s inými problémami môže viesť k zaostávaniu Slovenskej republiky v pripravenosti riešiť aktuálne územné problémy a byť rovnocenným partnerom novej Európy.

## Literatúra

- [1] Goodchild, M. F. Geographic Information Systems. In. International Encyclopedia of the Social & Behavioral Sciences, Elsevier Science B.V., 2001. ISBN 0-08-043076-7, p. 6175 – 6182
- [2] Frank, A. U., Raubal, M., Van der Vlugt, M. eds. Panel-GI Compendium A guide to GI and GIS. Vienna, GeoInfo Series Nr. 21, Institute for Geoinformation Technical University of Vienna, 2000, p. 9 – 42.
- [3] Martin, D., Bracken, I. The interpretation of socioeconomic and physical resource data for applied land management information system. Applied Geography, Vol. 13, No. 1, 1993, p. 45 – 53.
- [4] Stehlíková, B. Priestorová štatistika a software pre analýzu priestorových údajov. In. Zborník príspevkov z 10. medzinárodného seminára Výpočtová štatistika. Bratislava, Slovenská štatistická a demografická spoločnosť, 2001, s.114 – 116.
- [5] Dent, R. D. Cartography Thematic Map Desing. USA, Wm.C.Brown Publishers, 1996, p. 86 – 90.
- [6] Meixner, H., Frank, A. U. GI Policy – Study on Policy Issues Relating to Geographic Information in Europe. Brussels, The European Commission (EC) DG XIII, Technical Report, 1997, <http://forum.europa.eu.int/Public/irs/dsis/giss/library>
- [7] Bakytová, H., Ugron, M., Kontšeková, O. Základy štatistiky. Bratislava, ALFA, 1979.
- [8] Statistics Denmark: The National Square Grid – Denmark and the development of Information Bearing Cells in this system. Document EC No. F/GIS/61/EN – Part B, Meeting of the Working

- Party "Geographical Information Systems for Statistics". Luxembourg, October 23 – 24, 2002, <http://forum.europa.eu.int/Public/irs/dsis/giss/library>
- [9] Vavrínek, B. Priestorové jednotky v informačnom systéme. Bratislava, Pedagogické listy, č. 4, Stavebná fakulta Technickej univerzity, 1997, s. 91 – 99.
- [10] Horák, J. Využití pravděpodobnostního mapování v analýze trhu práce. Ostrava, Sborník vědeckých prací Vysoké školy báňské, Technická universita Ostrava, roč. 48, 2002, s. 131 – 139.
- [11] Prostorová analýza nezaměstnanosti. Grant GA ČR č. 402/99/0022, [http://gis.vsb.cz/Vyzkum/GA\\_CR/GrantPAN/pan.htm](http://gis.vsb.cz/Vyzkum/GA_CR/GrantPAN/pan.htm)
- [12] EUROSTAT 2003, <http://europa.eu.int/comm/eurostat/public>
- [13] SABE User Guide. MERGIN, 1998, <http://www.eurogeographics.org/products/SABE/index.htm>
- [14] Kusendová, D. Problematika geokódovania štatistických údajov. In. Výpočtová štatistika. Zborník referátov. Bratislava, Slovenská štatistická spoločnosť, 1999, s. 46 – 52.
- [15] Proposal for a regulation of the European Parliament and of the Council of the European Union On the establishment of a common classification of Territorial Units for Statistics – NUTS – 2001/0046 COD. Directorate EC F – 4, Brussels, EUROSTAT, 2001.
- [16] The European Commission's GI and GIS WebPortal, <http://www.ec-gis.org/>
- [17] Geographic Information System of the Commission – Progress Report 2001/2002. Document EC Unit E – 4, Luxembourg, EUROSTAT, September 2002, <http://forum.europa.eu.int/Public/irs/dsis/giss/library>
- [18] Final report „Meeting of the Workshop on GI/GIS matters for Phare Countries“ Joint meeting with National Statistical Offices and National Mapping Agencies of the Phare Candidate Countries. Document WS/PHARE, Luxembourg, EUROSTAT, October 24, 2001, <http://forum.europa.eu.int/Public/irc/dsis/giss/library>
- [19] INSPIRE, <http://www.ec-gis.org/inspire>.
- [20] Udržalová, K. Územní příprava Sčítání lidu, domů a bytů v roce 2001 a metainformace o GIS produktech. In. GIS Ostrava 2002. Ostrava, Institut ekonomiky a systémů řízení, Hornickogeologické fakulty VŠB – Technická univerzita Ostrava, 2002. (CD nosič ISSN 1213 – 2454)
- [21] Metainformace o tématické vrstvě definičních bodů stavebních objektů a o tématické vrstvě hranic sčítacích obvodů v České republice, Praha, Český statistický úřad, 2002, <http://gis.vsb.cz/midas>
- [22] Udržalová, K. Územní registr budov ČR jako GIS řešení. In. GIS Ostrava 2003. Ostrava, Institut ekonomiky a systémů řízení, Hornickogeologické fakulty VŠB – Technická univerzita Ostrava, 2003. (CD nosič ISSN 1213 – 2454)

## Poznámky

### Poznámka 1

Podľa noriem ISO 19101 a 19107 ide o informácie týkajúce sa javu implicitne alebo explicitne priradeného k polohe na Zemi. Geografické informácie tvoria údaje o polohe, tvare a reláciách medzi geografickými javmi, zapamätané spravidla v tvare súradníc a topológie.

### Poznámka 2

U nás má v súčasnosti tento mandát Geodetický a kartografický ústav (GKÚ), resp. Topografický ústav (TOPU ASR).

### Poznámka 3

Local Administrative Units zodpovedajú NUTS 4 a 5, ich zoznam národnej štatistickej úrady aktualizujú pre EUROSTAT každý rok.

### Poznámka 4

Podrobnejšie k jednotlivým projektom pozri [16],[13],[19].

### Poznámka 5

EuroGeographics je strategickým partnerom iniciatívy INSPIRE, jeho úloha spočíva v harmonizácii a produkcii referenčných geografických údajov v spolupráci s európskymi národnými mapovými agentúrami.

Doc. RNDr. Dagmar Kusendová, PhD., Prírodovedecká fakulta Univerzity Komenského, Katedra humánnej geografie a demogeografie, Mlynská dolina, 842 15 Bratislava, e-mail: kusendova@fns.uniba.sk

### Citácia:

Kusendová D. (2003). Aplikácia geoinformačných systémov v štatistike. Slovenská štatistika a demogeografia. Bratislava, Štatistický úrad Slovenskej republiky, roč. 13, č. 4, s.39-54.

## OBSAH

<b>SÚHRN.....</b>	<b>1</b>
<b>ÚVOD.....</b>	<b>1</b>
<b>1. GEOINFORMAČNÉ SYSTÉMY A GEOGRAFICKÉ INFORMÁCIE .....</b>	<b>1</b>
VÝVOJ A INTEROPERABILITA GIS A GI.....	2
<b>2. GEOGRAFICKÉ INFORMAČNÉ SYSTÉMY A ŠTATISTIKA .....</b>	<b>3</b>
<b>3. GEOGRAFICKÉ REFERENCovanie A KÓDOVANIE ŠTATISTICKÝCH DÁT.....</b>	<b>4</b>
PRIESTOROVÉ ŠTATISTICKÉ JEDNOTKY .....	4
<i>Územné vymedzenie.....</i>	4
<i>Mierka .....</i>	5
<i>Štruktúra.....</i>	5
<i>Konzistencia .....</i>	5
EURÓPSKA PRIESTOROVÁ ÚDAJOVÁ INFRAŠTRUKTÚRA .....	5
NÁRODNÉ ÚZEMNÉ ŠTATISTICKÉ JEDNOTKY.....	5
<b>4. GEOGRAFICKÉ INFORMAČNÉ SYSTÉMY V EURÓPSKEJ ŠTATISTIKE.....</b>	<b>6</b>
GISCO (GEOGRAPHIC INFORMATION SYSTEM OF THE EUROPEAN COMMISSION) .....	6
INSPIRE (INFRASTRUCTURE FOR SPATIAL DATA IN EUROPE) .....	7
SÚČASNÉ AKTIVITY EURÓPSKEHO ŠTATISTICKÉHO ÚRADU V OBLASTI GIS .....	7
TANDEM (TOWARDS A COMMON GEOGRAPHICAL BASE FOR STATISTICS ACROSS EUROPE).....	8
SYSTÉM ÚZEMNÝCH ŠTATISTICKÝCH JEDNOTIEK V SR .....	8
<b>5. CENTRÁLNY GEOINFORMAČNÝ ÚZEMNÝ REGISTER ŠTATISTICKÉHO ÚRADU SR.....</b>	<b>9</b>
<b>LITERATÚRA .....</b>	<b>9</b>