

INFORMAČNÍ BULLETIN



České statistické společnosti

Ročník 21, číslo 2, březen 2010

PŘÍKLADY POUŽITÍ GIS PŘI HODNOCENÍ ZDRAVOTNÍCH DAT

**Hana Šlachtová, Hana Tomášková, Pavla Polaufová,
Ivan Tomášek, Jiří Michalík, Anna Šplíchalová**

Adresa: Zdravotní ústav se sídlem v Ostravě,
Partyzánské nám. 7, 728 00 Ostrava

E-mail: hana.slachtova@zuova.cz

Abstrakt

In 2005 the National Reference Laboratory for using GIS in Public Health was established by the Czech Ministry of Health. Since that time a number of studies has been carried out. The presented paper shows some examples of using GIS for analysis of health data – e.g. impact of air pollution on prevalence of allergy, response-rate differences by area deprivation, space-time analysis of traffic air pollution impact on population, analysis of mortality in small areas using Bayesian methods. The outputs of such analyses are not only presented in scientific journals but also offered to public health authorities.

Key words: intersect of coverages, GIS selection, Bayesian methods.

1. Úvod

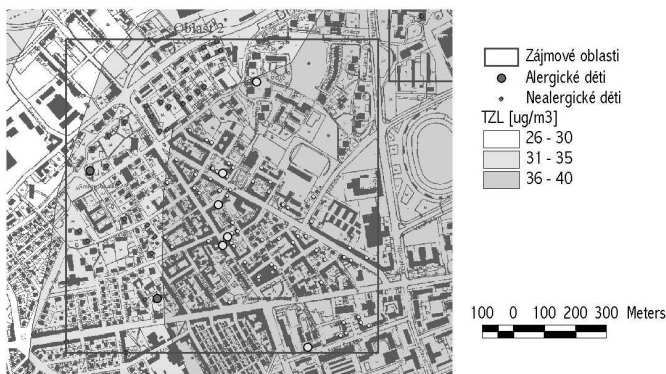
Věstníkem MZ ČR byla v dubnu 2005 zřízena Národní referenční laboratoř (NRL) pro využití GIS v ochraně a podpoře veřejného zdraví ve Zdravotním ústavu se sídlem v Ostravě (ZÚ). Na tomto pracovišti byly již před zřízením NRL realizovány studie používající GIS jako nástroj pro statistické vyhodnocení dat, které umožnily právě geoinformační technologie, a v těchto aktivitách pokračuje NRL i nadále. V článku jsou uvedeny příklady použití GIS při hodnocení zdravotních dat.

2. Analýza znečištění ovzduší ve vztahu k výsledkům prevalenčního šetření

V letech 1996–1998 byla na ZÚ vyhodnocována ostravská data z opakované průřezové studie výskytu alergických onemocnění ve věkové skupině 5, 9, a 13 let z celostátního monitoringu. Údaje byly získány z dotazníků vyplněných rodiči dětí a příslušnými pediatry. Pro statistickou analýzu byly použity obvyklé postupy používané v epidemiologii – základní vztahy mezi proměnnými

byly vyhodnoceny χ^2 -testem na hladině významnosti 5 % a výsledné faktory ovlivňující výskyt alergií (adjustované na vliv zavádějících faktorů) byly analyzovány logistickou regresí.

Pro vyhodnocení vlivu znečištěného ovzduší na výskyt alergií byli pacienti lokalizováni dle adresy bydliště do GIS ve čtvercích o straně 1 km ve spádové oblasti příslušných pediatrií. Průnikem vrstvy adresních bodů a rozptylových modelů průměrných ročních koncentrací základních polutantů bylo možno rozdělit děti (jak alergické, tak nealergické) podle příslušných ploch izokoncentrací (obr. 1).



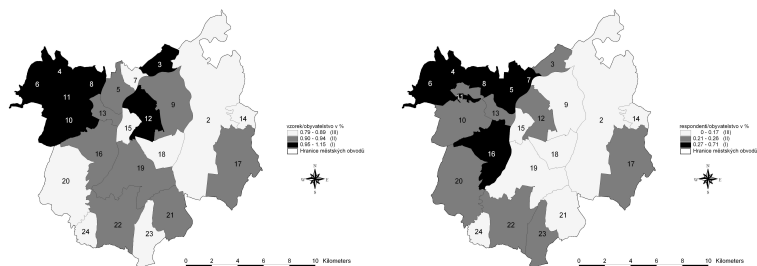
Obrázek 1: Výběr alergických dětí podle izokoncentrací TZL.

Navazovala logistická regrese, která v adjustovaných faktorech zohledňovala také vliv různých koncentrací škodlivin v ovzduší. Vliv znečištěného ovzduší na výskyt alergických onemocnění však nebyl prokázán [7].

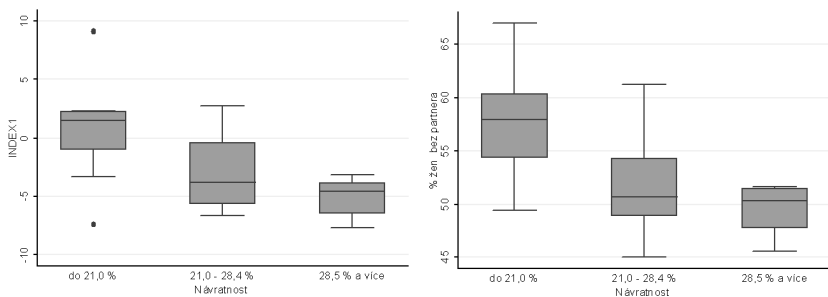
3. Vyhodnocení rozdílů návratnosti dotazníkové studie vzhledem k deprivaci území

Odborná literatura uvádí nižší návratnost dotazníků v oblastech s vysokou deprivací a nízkou sociální kohezí. Data z dotazníkového šetření zdravotního stavu, životního stylu a socioekonomických faktorů (grant IGA MZ ČR č. NJ6139 [8]) byla použita k analýze rozdílů návratnosti podle typu zástavby (indikátor sociální soudržnosti sídla) a podle deprivace jednotlivých ostravských městských obvodů. Deprivace byla vyjádřena indexem deprivace SESDI (IGA MZ ČR č. NR/8480 [9]).

Analýza návratnosti dotazníkového šetření užitím GIS byla provedena porovnáním počtu obyvatel v jednotlivých 23 městských obvodech Ostravy s počty osob vybraných náhodným výběrem v městských obvodech (členění podle městských obvodů nepatřilo ke kritériím náhodného výběru). Do GIS bylo lokalizováno všech 3 000 oslovených (vzorek) s atributem o návratnosti. Porovnáním počtu osob ve vzorku k počtu obyvatel jednotlivých obvodů χ^2 -testem bylo zjištěno, že v procentním zastoupení vzorku v městských obvodech není statisticky významný rozdíl. GIS analýzou bylo porovnáno procento respondentů, kteří vrátili vyplněný dotazník, ke vzorku, tedy celkovou návratnost v obvodech. Podle míry návratnosti byly městské obvody rozděleny do 3 kategorií. Vyhodnocením očekávané a skutečné návratnosti (opět ve 3 kategoriích – obr. 2) byly GIS analýzou zjištěny rozdíly v návratnosti [10].



Obrázek 2: Vizualizace očekávané a skutečné návratnosti v městských obvodech.



Obrázek 3: Znázornění vztahu míry návratnosti a SESDI a podílu žen samostatně žijících.

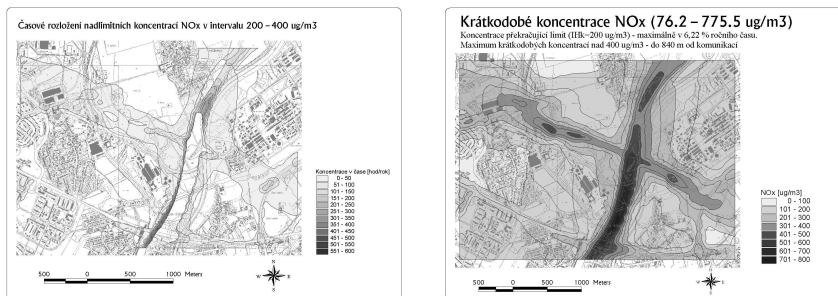
Vyšší návratnost byla zaznamenána v obvodech s převažující zástavbou rodinných domků, opačně nižší v obvodech s převládající zástavbou činžovních a panelových domů. V další analýze byl prokázán inverzní vztah depri-
vace k návratnosti podle SESDI ($r = -0,60$), míry nezaměstnanosti ($r = -0,44$), podílu osob se základním vzděláním ($r = -0,48$) a podílu žen-samoživitelek ($r = -0,61$). Na obr. 3 jsou znázorněny průměrné hodnoty socioekonomických ukazatelů ve 3 skupinách podle návratnosti. Analýza roz-
ptylu ANOVA potvrdila statisticky významný vztah mezi mírou návratnosti a SESDI ($p < 0,05$) a podílem žen bez partnera ($p < 0,01$). Vztah míry návratnosti a podílu nezaměstnaných a podílu osob se základním vzděláním nebyl statisticky významný.

4. Odhad vlivu na zdraví obyvatel na základě časového rozložení imisní zátěže z dopravy

Cílem analýzy bylo prověřit lokálně v detailech posuzovaný úsek komunikace Severní spoj v Ostravě s ohledem na imisní zatížení podél jejího průběhu a rovněž v souvislosti s dálnicí D 47 a návaznými komunikacemi ve výhle-
dovém roce 2020 [6]. Posouzení imisní zátěže bylo zpracováno formou krát-
kodobých, maximálních denních a průměrných ročních koncentrací NO_x ve vztahu ke konkrétní populaci v různých úrovních imisního zatížení a v případě
krátkodobých koncentrací i se zohledněním jejich časového průběhu.

Pro stanovení výhledové imisní zátěže byly použity rozptylové modely znečišťujících látek v ovzduší (výhledové dopravní intenzity, emisní faktory vozidel, meteorologická data – větrná růžice, lokalizace zdrojů a výpočtových bodů v GIS). Pro míru ovlivnění obyvatel byly použity GIS analýzy výběru počtů ovlivněných obyvatel v blízkosti posuzovaných komunikací pro různé
úrovně vybraných imisních charakteristik NO_x . Nově byl vytvořen model doby trvání krátkodobých koncentrací pro jednotlivé intervaly imisní zátěže – model vznikl překryvem informačních vrstev nadlimitních krátkodobých koncentrací a doby jejich trvání během roku.

	Rozpětí konc. ($\mu\text{g}\cdot\text{m}^{-3}$)		
Čas. interval (hod)	135–200	201–400	nad 400
0–50	3423	809	93
51–100	144	8	0
101–150	0	0	0
151–200	0	0	0
201–250	0	3	0
Σ Obyvatel	3567	820	93



Obrázek 4: Časové rozložení nadlimitních koncentrací NO_x v intervalu $200\text{--}400\ \mu\text{g.m}^{-3}$ (vlevo), Krátkodobé koncentrace NO_x ($76,2\text{--}775,5\ \mu\text{g.m}^{-3}$) (vpravo).

Tab. 1 – Výběr počtu obyvatel pro krátkodobé koncentrace NO_x .

Nadlimitními koncentracemi bude v posuzované oblasti ovlivňováno 913 obyvatel z celkového počtu 14313 obyvatel bydlících v zájmovém území, což činí cca 6,4 %. Detailní GIS analýzou bylo predikováno, že 5% hranice ročního času (dle tehdejší platné legislativy) nebude překračována v trvale obydlených oblastech. Pouze s pomocí GIS analýz bylo možno zjistit, zda a v tomto případě, i nově, v jakém časovém intervalu bude nadlimitními koncentracemi ovlivňováno trvale bydlící obyvatelstvo podél posuzovaných komunikací [6].

5. Analýza úmrtnosti v malých sídelních jednotkách užitím bayesovských metod

V současné době jsou rutinní zdravotní ukazatele v České republice sledovány a vyhodnocovány na úrovni ČR, krajů a okresů. V letech 2007–2009 zpracovala NRL pro GIS v rámci projektu OE236 – EUREKA/ISTAHIS/ analýzu úmrtnosti na úrovni 22 obcí s rozšířenou působností (ORP) v Moravskoslezském kraji (MSK) [11]. Byla analyzována data o počtech zemřelých na nejvýznamnější příčiny úmrtí za roky 2003–2006.

Vzhledem k malým počtům úmrtí na specifické příčiny v jednotlivých letech nebylo možno použít rutinní statistické postupy – SMR (standardizovaný úmrtnostní index) či SIR (standardizovaný incidenční index), vypočtené vzhledem k referenčním hodnotám. V případě územních jednotek s malými počty obyvatel jsou tyto charakteristiky nestabilní. Bayesovské metody umožňují tuto nestabilitu rizika korigovat a případně i zhodnotit vliv dalších faktorů. Pro zpracování byl použit program WinBUGS [5].

Empirické bayesovské metody spočívají na principu sdílení informací mezi oblastmi. Půjčováním informací ze sousedních oblastí nebo ze všech oblastí se stabilizuje variance vyhlazených odhadů směrem k průměru sousedních oblastí (local smoothing) nebo k celkovému průměru všech oblastí (global smoothing).

Základním nástrojem pro stanovení pravděpodobnosti je *Bayesova věta*, která umožňuje kombinovat současné znalosti s experimentálně získanými daty

$$p(\theta|y) \propto p(\theta)p(y|\theta)$$

kde θ jsou neznámé parametry, y jsou pozorovaná data, $p(\theta)$ je apriorní (nepodmíněná) pravděpodobnost pro parametry, $p(y|\theta)$ je pravděpodobnost dat podmíněná parametry a $p(\theta|y)$ je aposteriorní pravděpodobnost daná daty. V bayesovské statistice je parametr θ považován za náhodnou veličinu, na rozdíl od klasických metod, ve kterých je θ fixní veličina. Apriorní informace vyjadřuje znalosti o parametru nezávisle na pozorovaných datech. Ta může být informativní, pokud existuje dostatek informací nebo neinformativní v případě, že informace o parametru nejsou k dispozici nebo jsou nedostatečné. Zahrnutí apriorní informace do výpočtů je jednou z hlavních předností bayesovské statistiky [11].

Pro výpočet relativního rizika pro každou ORP vztaženého k průměrné hodnotě rizika v ČR (R_{CR}) a k výpočtu relativního rizika vztaženého k průměrné hodnotě rizika pro MSK (R_{MSK}) byl použit log-normální model.

Log-normální model [4] vychází z Poissonova rozložení pro pozorované počty O a rozdělení náhodných efektů na log-relativní škále rizika.

$$O_i \sim \text{Poisson}(E_i R_i),$$

$$\log R_i = \alpha + \theta_i,$$

kde:

- α – průměrná hodnota log relativního rizika (intercept),
- i – oblast
- R_i – riziko pro každou jednotku vztahu k referenčnímu regionu (České republice)
- E_i – očekávaný počet případů v oblasti i
- θ_i – náhodné jevy (nekorelovaná heterogenita),

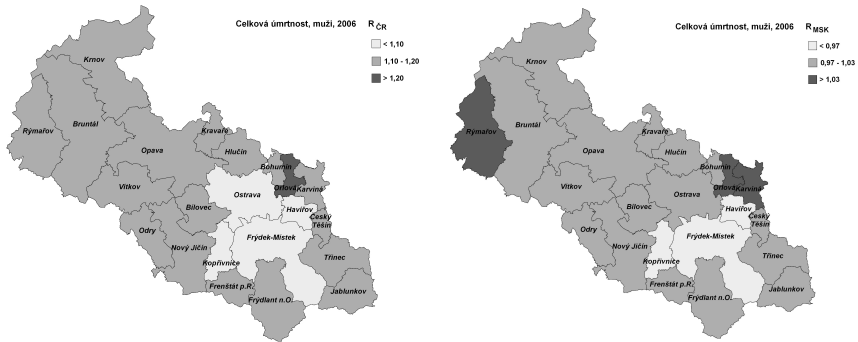
Náhodným jevům je v tomto modelu přiřazeno normální rozdělení

$$\theta_i \sim N(0, v^2),$$

kde:

- v^2 je rozptyl (vyjadřuje množství extrapoissonovských variací v datech [11]).
- Parametr α a v^2 je přiřazena odpovídající apriorní distribuce [11].
- R_i vyjadřuje riziko pro každou jednotku ve vztahu k průměrné hodnotě rizika pro Českou republiku (R_{CR}).
- Adjustované R_i vyjadřuje riziko pro každou jednotku ve vztahu k celkovému průměru rizika pro zvolenou oblast (R_{MSK}).

Nevýhodou tohoto modelu je, že nezohledňuje prostorovou strukturu, tj. vliv sousedních jednotek.



Obrázek 5: Ukázka použití log-normálního modelu – celková úmrtnost mužů R_{CR} (vlevo) a R_{MSK} (vpravo) v roce 2006.

Pro výpočet relativního rizika pro každé ORP vztahené k průměrné hodnotě rizika v MSK a s ohledem na hodnoty sousedních ORP ($R_{MSK+okoli}$) byl použit konvoluční model.

Konvoluční model [2, 3] vychází z Poissonova rozložení pro pozorované počty O

$$O_i \sim \text{Poisson}(E_i R_i)$$

Náhodné efekty θ_i jsou rozděleny na dvě složky: 1) nestrukturované jevy H_i a 2) vlivy struktury S_i .

$$\log R_i = \alpha + S_i + H_i$$

Nestrukturované jevy, kterým je přiřazeno normální apriorní rozdělení, jsou modelovány nezávisle pro každou jednotku.

$$H_i \sim N(0, v^2)$$

Pro prostorovou strukturu byl použit kondicionální autoregresní prior (CAR) [2], který vychází z předpokladu, že v sousedních jednotkách je riziko podobné. Riziko pro každou jednotku je podmíněně závislé na rizicích v sousedních jednotkách. Podmíněný průměr apriorního rozložení pro vlivy struktury je dán aritmetickým průměrem vlivů struktury sousedních oblastí a podmíněná variance apriorního rozložení je proporcionální k počtu sousedních oblastí.

$$S_i | S_{j \neq i} \sim N(m_i, s_i^2),$$

$$m_i = \frac{\sum_{j \neq i} w_{ij} S_j}{\sum_{j \neq i} w_{ij}}, \quad s_i^2 = \frac{s^2}{\sum_{j \neq i} w_{ij}},$$

kde:

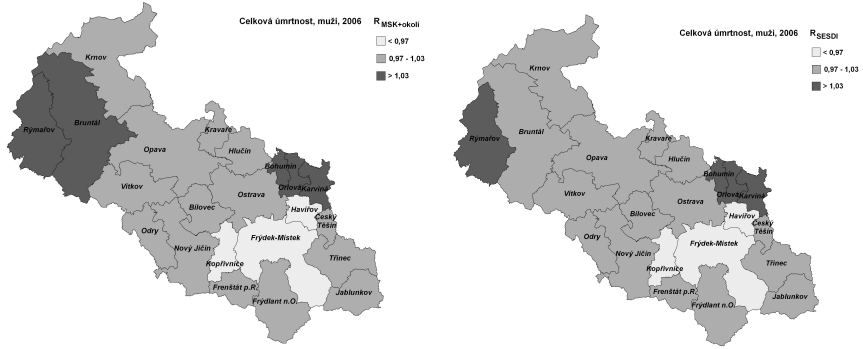
- m_i – střední hodnota Poissonova rozdělení pro pozorované počty ($m_i = E_i R_i$)
- w_{ij} – váha, která definuje vztah mezi jednotkou i a sousední jednotkou j (nabývá hodnot $w_{ij} = 1$, pokud oblasti i a j spolu vzájemně sousedí, tj. sdílejí společnou hranici, a $w_{ij} = 0$ v případě kdy spolu nesousedí),
- $\sum_{j \neq i} w_{ij}$ – počet jednotek, které sousedí s oblastí j ,
- s^2 – podmíněný rozptyl vlivu struktury v jednotce i podmíněný struktuře okolních jednotek,
- v^2 – marginální rozptyl nestrukturovaných jevů log-relativního rizika mezi oblastmi
- I – počet všech jednotek

Přímé srovnání s^2 a v^2 není možné. Stanovení podílu vlivu prostorové struktury na celkové variabilitě log-R lze provést na základě výpočtu tzv. prostorové frakce ($frac_{spatial}$), po předcházejícím empirickém výpočtu hodnoty marginálu rozptylu struktury log-relativního rizika mezi jednotkami $s_{marginal}^2$

$$s_{marginal}^2 = \sum_i (S_i - \bar{S})^2 / (I - 1)$$

$$frac_{spatial} = s_{marginal}^2 / (s_{marginal}^2 + v^2)$$

Pro výpočet relativního rizika (R_{INDEX1}) vztaheného k průměrné hodnotě rizika v *MSK* a s ohledem na hodnotu rizika okolních ORP a se zohledněním úrovně socioekonomické deprivace v jednotlivých ORP byl použit ekologický regresní model.



Obrázek 6: Vlevo je ukázka použití konvolučního modelu – celková úmrtnost mužů $R_{MSK+okoli}$ 2006, vpravo je ukázka použití ekologického regresního modelu – celková úmrtnost mužů R_{SESDI} .

Ekologický regresní model [5] vychází z konvolučního modelu a navíc zahrnuje vliv dalšího faktoru – socioekonomické deprivace (INDEX1) [9].

$$O_i \sim \text{Poisson}(E_i R_i),$$

$$\log R_i = \alpha + \beta X_i + S_i + H_i,$$

kde:

- β – regresní koeficient vyjadřuje vliv hodnoty SESDI na riziko v jednotce i
- X – kovariáta SESDI v oblasti i

Pro regresní koeficient β je použito uniformní apriorní rozdělení

$$\beta \sim flat().$$

Časoprostorové modely

Hierarchické modely [1] umožňují také modelovat variace onemocnění v čase a prostorové variace časových trendů v jednotlivých geografických jednotkách. Tyto modely většinou vycházejí z výše popsaného konvolučního modelu, který je doplněn o parametry charakterizující časový trend náhodných efektů a časoprostorové interakce (rozdíl mezi prostorovým a časovým trendem).

Časoprostorový model byl použit pro hodnocení rizika v jednotlivých letech a ke zhodnocení trendu úrovně úmrtí za období 2003–2006.

Výpočet relativního rizika u časoprostorových modelů vychází z Poissonova rozložení pro zjištěné počty O_{it} v jednotlivých geografických oblastech

(ORP) i ($i = 1, \dots, 22$) a v jednotlivých letech t ($t = 2003, 2004, 2005, 2006$) s průměrem $E_{it}R_{it}$, kde E_{it} jsou očekávané počty případů v oblastech i , čase t a R_{it} je relativní riziko pro oblast i , čas t :

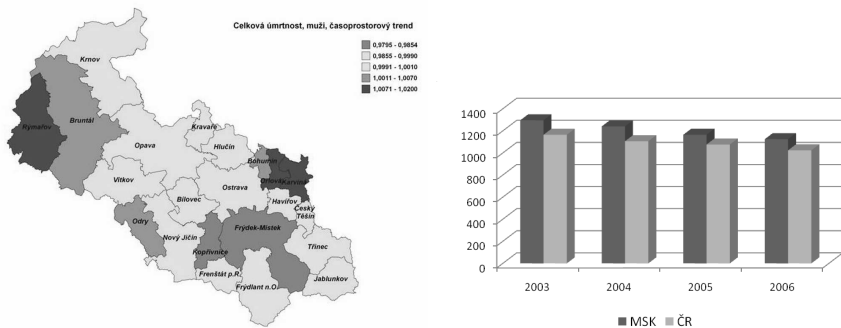
$$\log R_{it} = (\alpha + \theta_i) + (\beta + \delta_i)t$$

$$\theta_i \sim u_i + v_i$$

kde:

- α – průměrná hodnota log-relativního rizika za všechny sledované oblasti s uniformním rozdělením $\alpha \sim flat()$.
- θ – vyjadřuje efekt oblastí, zahrnuje vlivy heterogenity (korelované i nekorelované)
- β – průměrný lineární trend ze všech oblastí v čase t s normálním rozložením $N(\bar{\delta}_i, \tau.\delta)$
- δ_i – vyjadřuje časoprostorová interakce představující rozdíl mezi vlivem času a oblasti (obojí s vlastním Gaussovským CAR apriorním rozdělením s průměrem závislejícím na okolních oblastech [2]).

Výsledný časový trend (TT) byl vypočítán jako součet $(\beta + \delta_i)$.



Obrázek 7: Vlevo je ukázka časoprostorového modelu – celková úmrtnost mužů v letech 2003–2006, vpravo je celková standardizovaná úmrtnost mužů v MSK a ČR z rutinních dat.

6. Závěr

Prezentované ukázky užití GIS v hodnocení zdravotních dat a zdravotních rizik byly nejen publikovány v odborných časopisech, ale především poskytnuty orgánům ochrany veřejného zdraví k praktickému využití pro zpřesnění údajů z rutinních statistik.

Reference

- [1] Bernardinelli, L. et al. (1995). *Bayesian analysis of space-time variation in disease risk*. Statistics in Medicine, 14, p. 2433–2443.
- [2] Besag, J., York, J., Mollié, A. (1991). *Bayesian image restoration, with applications in spatial statistics*. Annals of the Institute of Statistics and Mathematics, 43, p. 1–59.
- [3] Clayton, D., Bernardinelli, L. (1992). *Bayesian methods for mapping disease risk*. In Elliott, P. Cuzick, J., English, D., Stern, R. eds. Geographical and environmental epidemiology: methods for small-area studies. Oxford: Oxford University Press, p. 205–220.
- [4] Clayton, D., Kaldor, J. (1987). *Empirical Bayes estimates of age-standardized relative risks for use in disease mapping*. Biometrics, 43(3), p. 671–681.
- [5] Lawson A. B., Browne W. J., Vidal Rodeiro C. L. (2003). *Disease Mapping with WinBUGS and MLwiN*. John Wiley & Sons Inc.; 2003.
- [6] Michalík, J. (2002). *GIS analýza časového rozložení imisního zatížení užitím modelu rozptylu znečišťujících látek z dopravy pro hodnocení zdravotních rizik*. Computer design, 1, s. 44–47.
- [7] Michalík, J., Šlachťová, H., Tomášková, H. (2001). *GIS v medicíně*. GeoInfo, 8(1), s. 20–22.
- [8] Slachťová, H., Splichalová, A., Tomášková, H. (2005). *Lifestyle, Socio-Economic Status and Health*. In: Focus on Lifestyle and Health Research. Ed. Kinger, L. New York: NOVA Science Publishers, p. 53–110.
- [9] Slachťová, H., Tomášková, H., Splichalová, A., Polaufová, P., Fejtková, P. (2009). *Czech socio-economic deprivation index and its use for analysis of mortality data*. Int J Public Health. 2009;54(4):267–73.
- [10] Šlachťová, H., Machová, T., Tomášková, H., Michalík, J. (2001). *GIS a dotazníkové šetření*. GeoInfo, 8(4), s. 59–60.
- [11] Tomášek, I., Tomášková, H. (2009). *Použití bayesovských metod pro mapování onemocnění*. Demografie, 51:190–201.

PROSTOROVÁ DATA

Jan Ružička, František Klímek

1. Úvod

Prostorová data jsou data, která popisují objekty reálného (eventuálně i nereálného) světa tak, aby byly zahrnuty časoprostorové aspekty objektů, jako je poloha v prostoru a čas, tvar (geometrie) a topologické vztahy mezi objekty, jako je sousednost, křížení, mimoúrovňové křížení. Prostorová data, která jsou vztažena k zemi, se někdy také nazývají geodata.

S prostorovými daty se setkáváme běžně v reálném životě, nejznámější formou reprezentace prostorových dat je mapa. V digitálním světě pak hovoříme o elektronické mapě nebo o elektronicky publikované mapě, případně o digitální mapě. Ke správě, analýze a vizualizaci digitálních prostorových dat slouží specializované systémy označované obvykle jako Geografické informační systémy (GIS). GIS je informačním systémem, jehož hlavním účelem je pracovat s informacemi, a to zejména odvozovat nové informace z dat a metadata. Na rozdíl od obvyklých IS však nabízí algoritmy pro práci s prostorovými daty.

2. Formáty

Digitální prostorová data mohou být ukládána v mnoha různých formátech. Klasifikace formátů by snesla hlubší a rozsáhlejší rozbor, zde se však omezíme jen na několik typických příkladů ukládání digitálních prostorových dat.

V zásadě se můžeme setkat se souborovým přístupem k ukládání prostorových dat, a s přístupem, kde je využívána nějaká z forem systému řízení báze dat (obvykle relačního nebo objektově relačního typu, ale již se objevily některé první pokusy o využití nativních XML databází).

V případě souborového přístupu jsou data ukládána do jednoho nebo skupiny více souborů, kdy v jedné logické souborové jednotce jsou obvykle uložena tématicky příbuzná prostorová data a také obvykle data z jednoho území. Některé souborové formáty mají omezení v ukládání geometrické složky popisu objektů na pouze jeden datový typ (např. pouze bod nebo pouze linie). Typickým příkladem je oblíbený a dobře dokumentovaný formát ESRI Shapefile, který je zároveň typickým zástupcem formátů, kde je pro uložení využíváno více souborů (3 a více). Základní soubor s příponou SHP obsahuje geometrii, soubor s příponou SHX je indexovým souborem pro efektivní

načítání souboru SHP a poslední ze základních tří souborů je formát DBF (obvykle DBase III).

Oproti tomu jiné formáty umožňují kombinovat v rámci jednoho souboru více geometrických typů a ty vhodně organizovat v různých vrstvách. Typickými příklady jsou proprietární formáty DGN (nativní formát CAD aplikace Microstation) a DWG (nativní program aplikace Autocad). Méně typickými formáty (jazyky) jsou GML (Geography Markup Language) a KML (Keyhole Markup Language), které jsou otevřenými jazyky založenými na XML. GML je standardizován významným sdružením Open Geospatial Consortium (OGC). KML je specifikován společností Google a je dnes jedním z nejprogresivnějších otevřených formátů na poli 3D dat.

Některé formáty umožňují zaznamenávat pouze 2D (nebo 2.5D data) některé 3D. Výjimečně se můžeme setkat s možností zaznamenávat 4D data (3D + čas), např. GML. Možnosti pracovat ve 4D prostoru jsou obvykle spojeny s konkrétní implementací v některém z programových prostředků a obvykle ve spojení se systémem řízení báze dat.

Doposud byly zmíněny pouze formáty, kde základním přístupem v popisu geometrie objektů reálného světa je využití vektorů (např. body, linie, polygony, TIN (triangular irregular network), B-Spline křivky, B-Spline plochy). V oblasti GIS se však také ve velké míře setkáváme s rastrovými daty, kde se obvykle nepopisují diskrétně jednotlivé objekty, ale pracuje se s prostorem jako celkem. Prostor je rozdělen pravidelnou mřížkou (v případě 3D pak prostorovou mřížkou, kde jsou buňky nazývány voxely) a každé buňce mřížky je přidělena určitá informace popisující dané místo. Hodnoty mohou být různého charakteru, např. spektrální projev při snímání z družice (letadla), identifikátor využití území (les, louka, pole, ...), nadmořská výška nebo třeba koncentrace kadmia v podzemní vodě. Takováto data jsou ukládána buď ve specializovaných formátech (např. IDRISI raster, GRASS raster) nebo jsou využívány běžné bitmapové formáty (např. TIFF (GeoTIFF), PNG, JPEG, ECW, MrSID).

Souborových formátů pro uložení geodat jsou dnes stovky. Naneštěstí zřejmě neexistuje reprezentativní přehled všech formátů, na který bychom mohli čtenáře odkázat. Jediným použitelným zdrojem by mohla být wikipedie, kde snad postupně bude takovýto ucelený přehled vytvořen:

[<http://en.wikipedia.org/wiki/GIS_file_formats>](http://en.wikipedia.org/wiki/GIS_file_formats)

Případně je možné využít přehled na stránkách www.geocomm.com:

[<http://data.geocomm.com/helpdesk/formats.html>](http://data.geocomm.com/helpdesk/formats.html)

V souvislosti s těmito daty je nutno zmínit problematiku prostorových referenčních systémů. V oblasti prostorových dat existují tisíce různých systémů pro zaznamenávání souřadnic, tak aby byla realita vhodně převedena do modelu. Např. v ČR je nařízením vlády povoleno využívat několik systémů pro nové mapování (ETRS 89, WGS 84, UTM 33, S-JTSK), ale stále přezívají data zaznamenaná s využitím starších systémů (např. S-42, Sv. Štěpán). Některé formáty umožňují přímo zaznamenat informaci o použitém systému (např. GML, GeoTIFF, GRASS raster) některé k tomuto účelu využívají další soubor s metadaty (např. ESRI Shapefile).

Přístup založený na využití systémů řízení báze dat (SŘBD) přináší mnoho výhod, proto bývá v nově vznikajících GIS obvykle využíván a soubory slouží pouze pro vstup a výstup dat. V současné době se často setkáme s tím, že běžný SŘBD již určitou podporu správy prostorových dat nabízí (např. MySQL, Oracle Spatial, Informix). Často je však vhodnější využít specializované řešení (ArcSDE/Geodatabase, PostgreSQL/PostGIS). Oblast správy prostorových dat v SŘBD relačního typu je standardizována OGC. K dispozici je tak verze SQL jazyka zaměřená na práci s prostorovými daty.

Jako příklad uveďme jednoduchou ukázkou využití SQL pro vypsání součtu délek řek pro jednotlivé kraje ČR:

```
SELECT k.nazkr, SUM(LENGTH(r.the_geom))/1000 as reky_km
FROM reky AS r, kraje_pseudo AS k
WHERE r.the_geom && k.the_geom
AND CONTAINS (k.the_geom, r.the_geom)
GROUP BY k.nazkr
ORDER BY reky_km;
```

3. Zdroje dat

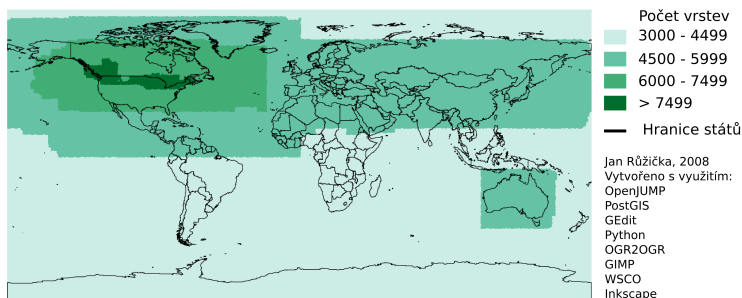
Digitálních prostorových dat je v současné době nepřeberné množství. S jejich organizací v katalogích (metainformačních systémech) je to však horší. Také dostupnost dat je komplikovaná, protože jsou často dostupná pouze za poplatek.

Není v možnostech tohoto článku podat vyčerpávající přehled všech dat dostupných v této oblasti. Z tohoto důvodu jsou uvedeny pouze odkazy na katalogy dat, které je možné využít pro vyhledání dat z různých oblastí. Pro demonstraci rozsahu uveďme následující příklad.

V současné době nejznámějším kanálem pro distribuci digitálních prostorových dat je „protokol“ Web Map Service (WMS). V katalogu, který byl vyvinut pro účely orchestrace služeb pro GeoWeb (v rámci GAČR 205/07/0797)

Počet vrstev publikovaných prostřednictvím WMS

Svět, Údaje z katalogu WSCO k 9.1.2008



Obrázek 1: Hustota pokrytí území světa vrstvami WMS služeb

a je nazván WSCO (Web Services Catalogue for Orchestration Open Source) je evidováno cca 200 takových služeb. V rámci těchto 200 služeb je zpřístupněno více než 10 tisíc datových vrstev v rozsahu od desítek objektů do deseti-tisíců objektů. Pro ilustraci byl připraven obrázek (Obr. 1) ukazující hustotu pokrytí území vrstvami geodat nabízenými prostřednictvím WMS služeb, které jsou evidovány v tomto katalogu.

Velmi pěkný seznam volně dostupných zdrojů pro zahraničí je na

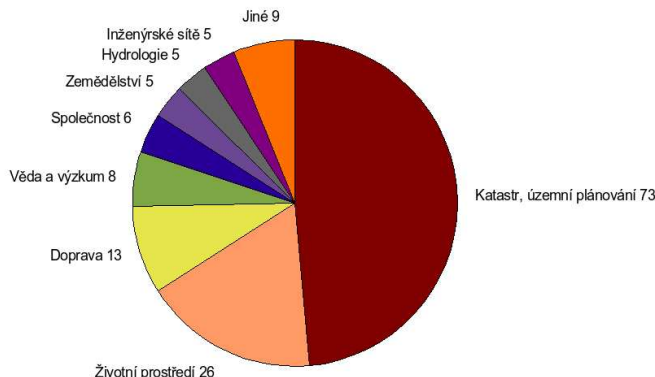
<<http://freegis.org/database/?cat=1>>

Pro představu, jaké typy dat jsou k dispozici v českém prostředí, jsme zpracovali graf (Obr. 2) ukazující počty datových sad evidovaných v metaportálu MIS CENIA (<http://mis.cenia.cz/>), který je provozován pod patronací Ministerstva životního prostředí. Vzhledem k evropské direktivě INSPIRE, která je zaměřena právě do oblasti životního prostředí je možno portál CENIA považovat za jeden z nejvýznamnějších v oblasti prostorových dat v ČR.

Velmi pěkně zpracovaný přehled českých zdrojů je možné nalézt ve výukovém materiálu připraveném Bronislavou Horákovou pro studenty oboru Geoinformatika na VŠB-TUO [Horáková 2007].

Uvedený zdroj však popisuje zejména data, která nejsou dostupná zdarma nebo jsou dostupná za určitých omezení. Oproti tomu existuje pro Českou republiku zajímavá kolekce geodat, která je distribuována zdarma. Některé datové sady z této kolekce sice nesmí být komerčně využívány, ale pro většinu z těchto dat to neplatí. Tato kolekce byla sesbírána Martinem Landou z ČVUT

Metaportál CENIA



Obrázek 2: Kategorie geodat evidovaných v metaportálu CENIA

a je nazvána FreeGeodataCZ. Tato data najdete i na GISAK LiveCD a LiveDVD, o kterém se zmiňuje druhý článek.

„Projekt FreeGeodataCZ reaguje na neutěšenou situaci ohledně volně dostupných geografických dat ČR. Primárně se snaží tato data shromažďovat a prezentovat je v kompaktní formě.“ [Landa 2007]

K dispozici jsou rastrová data digitálního modelu terénu ve dvou variantách, kde jedna z variant je vhodná zejména pro vizualizaci (dem_srtm) a druhá pro analýzu (dem_gtopo30).

Dále pak jsou k dispozici vektorová data:

- casti_obce (centroidy částí obcí)
- cesty (turistické stezky v okrese Žďár nad Sázavou)
- cfm (různé zájmové body)
- cr (hranice republiky)
- czfree_nodes (nody sítě CZFree)
- d_snimky (rozsah družicových snímků)
- klad_zm10 (klad listů Základní mapy 1:10 000)
- kraje_pseudo (hranice krajů)
- mes_casti (centroidy částí měst)
- mesta_b (města bodově)
- mesta_p (města polygony)

- obce (obce bodově)
- okresy_pseudo (hranice okresů)
- reky (řeky)
- silnice (silniční síť)
- silnice_pasport (pasport silniční sítě)
- silnice_useky (silniční úseky)
- silnice_uzly (uzly silniční sítě)
- voda (vodní plochy)
- zeleznice (železniční síť)

V neposlední řadě pak rastrová data reprezentující průměrnou roční teplotu (klimadata (bio_X)). Kompletní popis dat najdete na stránkách projektu FreeGeodataCZ [Landa 2007].

4. Metadata k datům

Metadata jsou ve stručnosti „data o datech“, a jako taková zefektivňují uchování a zpřístupňování informací.

Stálým problémem je velmi neutěšený stav metadat k prostorovým datům. K některým datům, která se běžně využívají pro rozhodování, nejsou metadata žádná nebo jsou metadata neúplná. Nepomáhá ani to, že v současné době je již téměř ustálena standardizace v této oblasti (existují standardy ISO), ani to, že již existují nástroje pro tvorbu metadat, a to i zdarma dostupné (např. GeoNetwork Open Source, CatMDEdit).

Literatura

[**Horáková 2007**] Horáková, B. *Zdroje prostorových dat*. Syllabus.

Dostupné na WWW (ZdrojeGeodat.pdf):

http://gisak.vsb.cz/~hor63/ZPD_06-07/ZPD_Sylaby06-07.zip

[**Landa 2007**] Landa, M. *FreeGeodataCZ*.

Dostupné na WWW: <http://grass.fsv.cvut.cz/>

GISAK LiveCD

František Klímek, Jan Růžička

1. Úvod

Před lety začal shromažďovat Martin Landa z ČVUT volně použitelná geografická data z oblasti České republiky s cílem nabídnout ucelenou sadu dat pro základní analýzy zejména pro studijní účely. Vznikla tak kolekce datových sad nazvaná FreeGeodataCZ. Tato kolekce je samozřejmě volně dostupná na Internetu a každý uživatel, který má o tato data zájem a má k dispozici nástroj pro vizualizaci formátu ESRI Shapefile, si může tato data prohlédnout.

Můžete se však dostat do situace, že chcete data ukázat někomu, kdo nemá ani prohlížečku ESRI Shapefile formátu a ani není ochoten (nemá právo) instalovat nějaký program. V takovém případě je možné využít výsledků projektu GISAK LiveCD. S využitím GISAK LiveCD si může kdokoli data prohlédnout (analyzovat) bez nutnosti instalace jakéhokoli programu na jeho PC. Navíc se prostřednictvím GISAK LiveCD může seznámit prostřednictvím tutoriálů s volně dostupnými nástroji pro vizualizaci a analýzu geodat.

LiveCD je operační systém (obvykle obsahující další software) uložený na bootovatelném CD, jenž z něj může být spuštěn bez nutnosti jeho instalace do pevné paměti, jakou je např. pevný disk. Počítač se znovu vrátí ke svému původnímu OS když je LiveCD vyjmuta z mechaniky a počítač je restartován.

GISAK LiveCD je projekt, který se snaží o podporu kombinované a distanční formy studia zkompletováním CD pro podporu několika předmětů vyučovaných institutem geoinformatiky. Hlavním cílem projektu je zjednodušit přípravu studentů v některých předmětech kombinovaného studia.

Dílčím cílem projektu je příprava tutoriálů pro výuku, tak aby bylo možné z nich studovat přímo v prostředí GISAK LiveCD. Dalším cílem je vytvoření šablony a návodu pro přípravu dalších tutoriálů z řad pedagogických kolegů. V neposlední řadě je to příprava dokumentace k samotnému LiveCD a procesu jeho přípravy tak, aby mohla vznikat podobná, avšak jinak tématicky zaměřená LiveCD.

GISAK LiveCD může být využito pro výuku, výzkum a testování. Uživatelé jsou z řad studentů kombinovaného studia na Vysoké škole báňské – Technické univerzitě Ostrava (VŠB-TUO), případně i denních studentů, dále pak studenti středních škol, kteří se o problematiku GIS zajímají, a v neposlední řadě také studenti celoživotního vzdělávání.

LiveCD již bylo využito přibližně 100 studenty VŠB-TUO (pravidelné využívání pro studium), asi 40 studenty středních škol (nemáme informace

o pravidelnosti či nepravidelnosti využívání), a přibližně 10 osobami z jiných škol (pedagogové i studenti). Počet stažení ze stránek projektu již přesáhl tři sta (zde však mohou být zahrnuti i uživatelé uvedení dříve).

LiveCD již bylo využito a využívá se pro výuku předmětů Programové vybavení pro GIS II, Modelování terénu a krajiny, Zpracování dat v GIS, Publikování prostorových dat v prostředí WWW, Modelování a simulace v geovědách, Multimediální systémy, v nejbližší době pak pro Značkovací jazyky v Geoinformatice, Tvorba WWW stránek, Tvorba WWW serverů a zvažuje se využití pro Úvod do geoinformatiky, GIS.

2. Nástroje

CD obsahuje následující nainstalované a nakonfigurované nástroje:

- GRASS (komplexní nástroj pro analýzu geodat, včetně práce ve 3D, 4D),
- QGIS (zatím jednoduchá prohlížečka geodat, postupně však svoji funkcionalitu rozšiřuje),
- Thuban (velmi jednoduchá prohlížečka),
- JUMP (komplexní nástroj pro vizualizaci a analýzu geodat),
- UMN MapServer (nástroj pro publikování prostorových dat v prostředí WWW),
- PHP/MapScript (nástroj pro vývoj aplikací pro publikování prostorových dat v prostředí WWW),
- PostGIS (nástroj pro správu geodat v prostředí relační databáze),
- Blender (nástroj pro 3D modelování a animace),
- GPS Drive (nástroj pro přípravu měření s GPS).

DVD přidává:

- GeoNetwork Open Source (nástroj pro správu metadat datových zdrojů a služeb),
- uDIG (zatím jednoduchá prohlížečka geodat, postupně však svoji funkcionalitu rozšiřuje),
- gvSIG (zatím jednoduchá prohlížečka geodat, postupně však svoji funkcionalitu rozšiřuje a v současné době nabízí velmi pěkné rozhraní pro tvorbu map),
- JGrass (rozšíření nástroje GRASS o uživatelsky přívětivější prostředí a nástroje pro hydrologické analýzy a modelování),
- LandSerf (nástroj pro 3D vizualizaci a analýzu digitálního modelu terénu, původně vyvinut pro potřeby archeologie),
- gpsbabel (nástroj pro manipulaci s daty z GPS přístroje).

3. Data

GISAK LiveCD rovněž obsahuje zmiňovaná prostorová data a další data určená pro testování. Je zde umístěn soubor cvičných dat z USA pro GRASS. Dále cvičná data pro celulární automaty. A především rozsáhlý soubor dat pro oblast České republiky sestavený Martinem Landou z ČVUT. Tento soubor dat je k dispozici nejen ve formátu ESRI Shapefile, ale i formátu pro GRASS a navíc tři datové soubory byly importovány do systému PostGIS. Všechny tyto datové soubory jsou po startu systému uživateli k dispozici a může s nimi pracovat.

4. Výukové materiály

Na zmiňovaném CD má uživatel k dispozici tutoriály, díky kterým se může ihned začít ponořovat do tajů GIS s využitím open source nástrojů. Mezi tutoriály, které uživatel na CD nalezne, je např.:

- GRASS – Prvé kroky, Autor: Klímek F.
- GRASS vizuálně, Autor: Klímek F.
- GRASS – Hledání vhodných území, Autor: Růžička J.
- GRASS – Animace šíření, Autor: Růžička J.
- JUMP – Načtení dat, Autor: Růžička J.
- JUMP – Rozdělení geoprvků do tříd, Autor: Růžička J.
- JUMP – Přidání ArcIMS služby, Autor: Růžička J.
- JUMP – Přidání vrstvy ze systému PostGIS, Autor: Růžička J.
- Quantum GIS – Načtení dat, Autor: Klímek F.
- Thuban – Úvod do práce s prohlížečkou, Autor: Klímek F.
- PostGIS – Začínáme, Autor: Růžička J.
- UMN MapServer – Začínáme, Autor: Růžička J.
- GISAK LiveCD – Různé, Autor: Růžička J.

Výše zmiňované tutoriály jsou přímo od autorů CD, nicméně zde lze najít i další výukové materiály dodané třetími stranami. Mezi těmito můžeme jmenovat:

- UMN MapServer – snadno a rychle, Autor: Čepický J.
- GRASS 6 VISUAL TUTORIAL, Autor: Moretti L.
- An introduction to the practical use of the Free Geographical Information System GRASS 6.0, Autoři: GDF Hannover bR
- Quantum GIS User Guide, Autoři: Sherman G. E., Sutton T., Blazek R., Luthman L.

- User's Manual for Thuban 1.0, Autoři: Coles J., Wagner J. O., Koormann F.
- JUMP Workbench User's Guide, Autoři: Vivid Solutions

Jedním z textů, který zde uživatel může najít, je i návod jak vytvářet tutoriály, obsahující i šablonu pro jejich tvorbu. Možná není překvapením že i tento návod je ve formě tutoriálů.

5. Přínos

Studijní obor geoinformatika takto získal další nástroj, kterým může usnadnit studentům kombinovaného (případně celoživotního) studia jejich přípravu formou samostudia. LiveCD však lze využít i pro studenty denního studia, pro zpracování samostatných úkolů nebo i pro pravidelnou výuku. Dokumentace k CD dále umožní vytvářet podobná CD i pro jiné obory a zaměření (což je však podmíněno vzděláváním pedagogů v této oblasti). Na jaře 2007 vzniklo CD zaměřené na výuku Multimediálních systémů pro studenty kombinovaného studia oboru Informační a systémový management v oblasti surovin. Návod a šablona umožní vytváření dalších tutoriálů. LiveCD již pomáhá v budoucnu dále pomůž zjednodušit výuku oboru geoinformatika.

6. Další informace

WWW stránky jsou dostupné na <http://gisak.vsb.cz/livecd/>. WWW stránky obsahují všechny tutoriály, návod pro tvorbu tutoriálů, šablonu pro tvorbu tutoriálů, dokumentaci k tvorbě LiveCD a v neposlední řadě ISO image samotného LiveCD a LiveDVD ke stažení. Postupně bychom rádi tyto stránky nahradili nějakým redakčním systémem např. MediaWiki.

Dokumentace k LiveCD je ve formě konfiguračního skriptu, který dokumentuje všechny kroky pro vytvoření vlastního LiveCD. Tvorba, a to si uvědomujeme, vyžaduje jisté pokročilé znalosti OS Linux. V dalším vývoji LiveCD se proto zkusíme zaměřit i na vytvoření tutoriálu pro tvorbu vlastního CD nebo ve vytvoření nástroje pro automatizovanou úpravu CD. V současné době je možné postupovat dle kombinace návodu na stránkách projektu Knoppix a dle konfiguračního skriptu. V případě nejasností je možné s autory konzultovat na jan.ruzicka@vsb.cz. Jinou možností je založit systém na distribuci DANIX, která nabízí zajímavý systém úpravy LiveCD.

LiveCD vyžaduje pro běh minimálně 512 MB RAM (případně umí využít swap oddíl na pevném disku). Minimální procesor, na kterém bylo testováno (a to doslova – studenti v něm dělali zápočtový test v prostředí nástroje PostGIS) byl Athlon 1 GHz. Oproti LiveCD je významně rychlejší provoz

systému a aplikací z LiveDVD. V obou případech je však nutná co nejrychlejší čtecí mechanika (24× pro CD-ROM již bylo k nevydržení).

7. Závěr

Práci s GISAKem demonstrují obrázky 1 až 4 v závěru článku.

Vytvořené LiveCD je již dnes využíváno řadou studentů. Předpokládáme další vývoj projektu, buď s podporou Fondu rozvoje vysokých škol (FRVŠ) nebo z vlastních zdrojů, tak jak tomu bylo do roku 2006. Rádi bychom také vytvořili speciální edici pro časopis GEO Business, a to s přispěním některého významného producenta dat, který by takto mohl svá data propagovat vzorkem dat přímo na LiveDVD.

Dále zvažujeme možnosti zařazení některých volně dostupných nástrojů, které však nejsou distribuovány jako open source. Zde je však distribuce problematická, a to jak z hlediska licenčního tak z hlediska provozu samotných aplikací, protože některé z nich neumožňují běh pod OS Linux. Takové případy bychom rádi řešili emulací prostředí OS Windows s využitím nástroje wine.

Literatura

[Růžička 2007] Růžička, J.; Klímek, F. *Závěrečná zpráva z projektu GISAK LiveCD*. 2007.

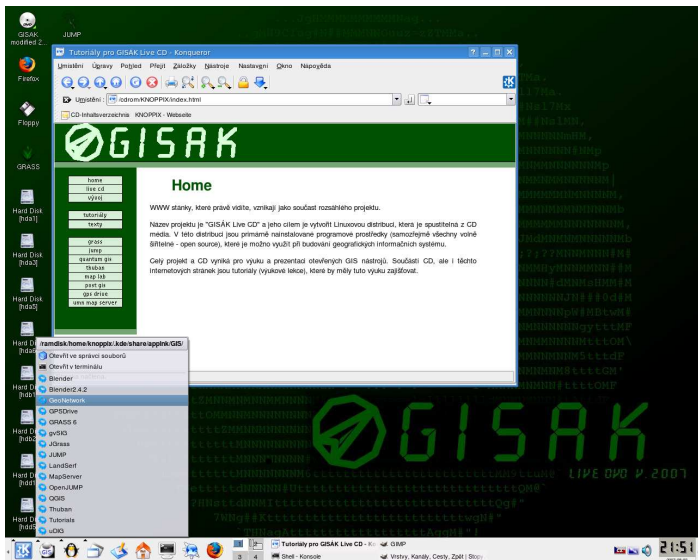
[Landa 2007] Landa, M. *FreeGeodataCZ*.

Dostupné na WWW: <http://grass.fsv.cvut.cz/>

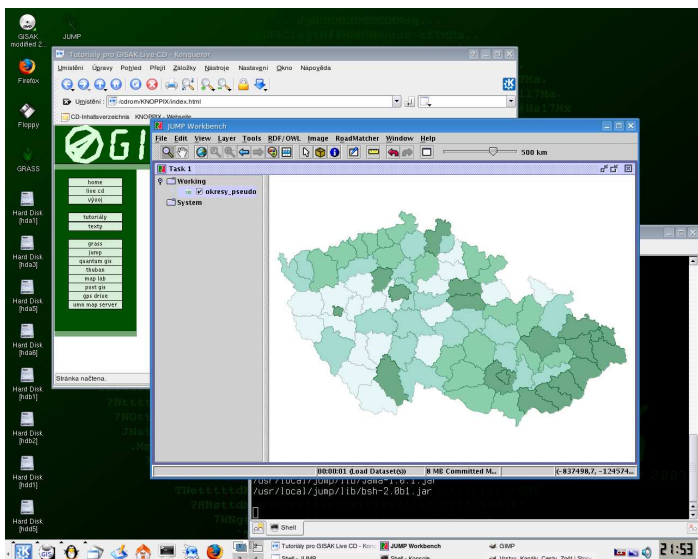
[Klímek 2007] Klímek, F.; Růžička, J. GISAK LiveCD – nebojte se open source GIS. In sborník z 1. národní kongres v Česku – *Geoinformatika pro každého*, Mikulov, 29.–31. 5. 2007.

Podpora

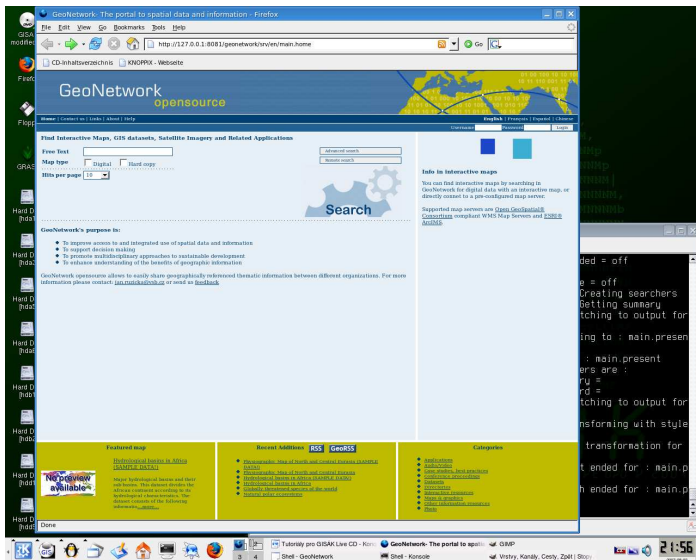
Projekt GISAK LiveCD se rozšířil za přispění FRVŠ. Významnou měrou rovněž přispěla open source komunita. Pro zpracování článku (projektu) byly využity open source projekty KNOPPIX, GRASS GIS, JUMP, OpenJUMP, JUMP Plugins from GISAK.VSB.CZ, Blender, QGIS, uDIG, GeoNetwork Open Source, UMN MapServer, GPS Drive, Open Office, GIMP, PostGIS, PHP, PostgreSQL, Apache HTTP Server, GNU/Linux Ubuntu, GNU/Linux Debian, X11, Freefont, a mnoho dalších.



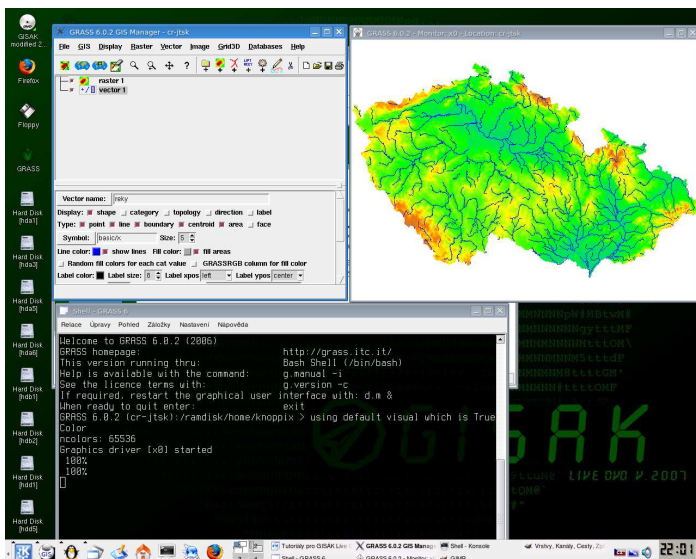
Obrázek 1: Pracovní prostředí GISAK LiveCD



Obrázek 2: Práce v programu JUMP



Obrázek 3: Úvodní obrazovka systému GeoNetwork Open Source



Obrázek 4: Práce v prostředí GRASS GIS

ZAŘAZENÍ GEOGRAFICKÝCH INFORMAČNÍCH SYSTÉMŮ DO VÝUKY PŘEDMĚTU INFORMATIKA VE VEŘEJNÉ SPRÁVĚ

Miroslava Dolejšová

Abstract: This paper deals with a new conception of the subject “Informatics in the public administration” taught in the second year of the field of study “Public administration and the regional development” and guaranteed by the Institute of the Informatics and Statistics at the Tomas Bata University in Zlín. The main purpose is to point out the significance of the geographical information systems in the public administration, to describe and clearly demonstrate the ways of their application both from the eye of citizens and the public authorities and to indicate the further development in this area.

Úvod

V současné době informační a znalostní společnosti se bez kvalitních informací prakticky neobejdeme. Na jedné straně existuje dostatek informací, ale stačí nám to k rozhodování? Víme, které informace máme použít, abychom učinili správné rozhodnutí? Máme k dispozici nástroje, které by nám byly schopny pomoci rozhodování usnadnit?

Odpověď je poměrně jednoduchá. Tyto nástroje k dispozici jsou a nazývají se geografické informační systémy.

Geografické informační systémy mají uplatnění v řadě oblastí. Jednou z nejvyužívanějších oblastí je oblast veřejné správy. Cílem tohoto příspěvku je popsat možnosti zavedení geografických informačních systémů do předmětu Informatika ve veřejné správě, který je vyučován na Fakultě ekonomiky a managementu Univerzity Tomáše Bati ve Zlíně.

Stručné představení předmětu IVS

Ústav informatiky a statistiky zajišťuje výuku povinného předmětu Informatika ve veřejné správě ve 2. ročníku zimního semestru studijního programu Hospodářská politika a správa ve studijním oboru Veřejná správa a regionální rozvoj, a to ve všech formách studia. Výuka v prezenční formě studia probíhá v rozsahu dvou cvičení po dobu 14 týdnů na počítačových učebnách, pro kombinovanou a celoživotní formu studia je vymezeno 9 konzultací.

Předmět je v prvé řadě zaměřen na práci s informačními zdroji, které jsou potřebné v oblasti veřejné správy. Neméně důležitým cílem tohoto předmětu je rozšíření počítačové gramotnosti studentů o kreslení diagramů v programu Microsoft Word, rozšíření znalostí programu Excel (grafy, jednoduché databázové funkce), případně další problémy, které budou potřeba pro vypracování bakalářské nebo diplomové práce.

Předmět je ukončen klasifikovaným zápočtem. Podmínkou pro jeho získání je vypracování seminární práce a úspěšné absolvování znalostního testu. Charakter seminární práce je zcela tvůrčí. Jejím cílem je provést analýzu toku informací ve studentem vybrané organizaci veřejné správy. V případě, že se studentům vypracování seminární práce zalíbí, mohou pokračovat v jejím podstatném rozšíření i v podobě bakalářské nebo diplomové práce.

Proč právě GIS v předmětu IVS?

Zásadním nedostatkem oboru Veřejná správa a regionální rozvoj jsou chybějící předměty, které jsou vyučovány Ústavem informatiky a statistiky ve studijním oboru Management a ekonomika. V oboru Management a ekonomika jsou navíc vyučovány předměty Databáze a programování (letní semestr 1. ročníku) a Aplikovaná informatika (zimní semestr 2. ročníku), který je věnován řešení konkrétních manažerských úloh v programu Excel. Studenti oboru Veřejná správa a regionální rozvoj mají pouze předmět Informatika pro ekonomy (pro oba obory je předmět společný) a ve druhém ročníku pak předmět Informatika ve veřejné správě.

Pouhé vyhledávání informačních zdrojů, rozšíření počítačové gramotnosti o možnosti Wordu a Excelu podle názoru autorky nestačí. Návaznost na další předměty oboru Veřejná správa a regionální rozvoj (především předmět Regionální analýza) vyvolala potřebu úpravy obsahu tohoto předmětu, který by propojil obsah obou předmětů z různých úhlů pohledu. Předmět Informatika ve veřejné správě by se mohl stát doplňkem předmětu Regionální analýza, případně dalších předmětů, které jsou vyučovány Ústavem veřejné správy a regionálního rozvoje. V každém případě by se jednalo o obohacení obsahu obou uvedených předmětů, rozšíření znalostí studentů a navázání úzké pedagogické i vědeckovýzkumné spolupráce mezi oběma ústavy.

Proč využívat GIS?

K nejčastějším typům informací, které potřebuje každý člověk i organizace, patří geografické informace. Pravidelně hledáme cestu do konkrétního místa, rozhodujeme se, kde umístíme další provozovnu, chceme vědět, jak dlouho

nám přeprava do zvoleného místa bude trvat, a také chceme znát vzdálenost mezi oběma místy. Geografické informační systémy nám tyto (a nejenom tyto) typy informací poskytují.

Možná se na první pohled bude zdát, že geografické informační systémy jsou jen software. GIS však možnosti klasického programu převyšují. Nejen, že nám usnadňují rozhodování, ale především se jedná o kvalitní grafický nástroj pro analýzu a modelování dat, která jsou ve většině případů volně k dispozici na Internetu ke stažení.

Geografické informační systémy jsou obvykle spojovány s mapami a navigačním systémem. Mapy jsou však výstupem geografických informačních systémů a navigační systém (myšleno systém globální navigace GPS) nebývá zahrnut do těchto systémů.

Základem jsou však data

V minulosti i dnes lidé získávali znalosti a sdíleli je v různých podobách. Dorozumívali se prostřednictvím slov, textu, symbolů v podobě hieroglyfů, hudby, obrazů a kreseb. V dnešní době digitalizace sdílíme znalosti prostřednictvím různých sítí (především World Wide Web). Postupně se možnosti rozšiřují o počítačové modelování a simulace, digitální zpracování obrazu i textu, správu obsahu a zejména o využívání metod statistické analýzy.

Tři pohledy na GIS

Na geografické informační systémy lze pohlížet ze tří hledisek: mapového, databázového a modelového.

Mapový pohled je nejjednodušší. Pomocí geografických informačních systémů můžeme mapy vytvářet a upravovat. Výsledná mapa je pak souborem námi vybraných vrstev, které chceme v mapě zobrazit.

Mapa jako obraz nám nebude stačit. Potřebujeme ji doplnit o konkrétní údaje. Tato data jsou uložena v různých databázích, které se často označují jako geodatabáze.

Pokud máme k dispozici obraz i data, můžeme prostřednictvím různých analytických nástrojů získávat další informace potřebné pro rozhodování. Výběrem dat, různými dotazy, aplikací statistických a analytických funkcí a dalších vhodných nástrojů získáme nová data, která nám umožní získat komplexnější pohled na studovaný problém.

Geografické informační systémy lze proto definovat jako prostředek zobrazení, manipulace a analýzy prostorových dat.

Typy dat, se kterými GIS pracují

Geografické informační systémy pracují s různými typy dat. K nejpoužívanějším z nich patří vektorová a rastrová data. Vektorová data mohou být vyjádřena jako polygony (typické pro lesy, území krajů a velkých měst, vodní plochy), čáry pro zobrazení silnic, železnic, řek, ulic nebo jako body (stromy, obchody, malá města a obce). Rastrová data jsou vyjádřena jako matice bodů. K rastrovým datům patří družicové a letecké snímky, zobrazení terénů a sítí.

Co můžeme zjišťovat prostřednictvím GIS?

- Zjišťovat počty (KOLIK?): počet nemocnic, počet úřadů.
- Zjišťovat hustotu (JAK MNOHO?): počet lékařů připadajících na 1 000 pacientů.
- Zjišťovat, co je uvnitř určité oblasti: počet obchodů v určitém regionu.
- Zjišťovat, co je poblíž: nejbližší nemocnice, nejbližší rekreační středisko.
- Modelovat změny (obvykle předpovědi počasí).

Praktické ukázky

Další část příspěvku bude popisovat možné aplikace geografických informačních systémů (nebo lépe řečeno ne přímo jejich podoby) v předmětu Informatika ve veřejné správě. Jelikož se studenti v tomto předmětu seznamují také s nejpoužívanějším informačním zdrojem ve veřejné správě, Portálem veřejné správy České republiky, bude potřeba jejich znalosti rozšířit i o aplikaci mapových služeb. Příspěvek bude charakterizovat ilustrativní příklady, které byly vytvořeny Portálem veřejné správy [1] a prostřednictvím internetové stránky <http://www.mapy.cz/> [2].

Protože se již podruhé konference STAKAN koná v krásném prostředí Rusavy, rozhodla jsem se ukázat možnosti, které geografické informační systémy nabízejí, právě v okolí Rusavy.

Při pokusu najít přesnou polohu rekreačního zařízení Rusava (chata Jestřábí) budeme úspěšní přes internetové stránky <http://www.mapy.cz/>. Pokud se o totéž pokusíme přes Portál veřejné správy, zobrazí se jen turistická mapa. Práce s mapovými službami není vůbec složitá. Mapové služby spustíme kliknutím na odkaz Mapy v pravé horní části Portálu veřejné správy. Největší část této aplikace tvoří prostor pro zobrazení mapy. Postupným kreslením obdélníků pomocí myši specifikujeme region, který chceme prozkoumat. Jakmile jsme s výběrem zcela spokojeni, vybereme z části Funkce aplikace odkaz na obrázek tří map, který představuje seznam tematických úloh. K dispozici je jich celkem padesát.

Velmi důležitou součástí aplikace jsou karty Vrstvy a Legenda. Na kartě Vrstvy lze vybrat, co konkrétního si přejeme zobrazit, na kartě Legenda vidíme vysvětlení popisků mapy a vybrané statistické údaje. Zrušením zaškrtnutí před příslušnou vrstvou odstraníme z mapy všechno, co v ní vidět nechceme. Změna se však provede pouze překreslením mapy (zakulacená šipka nacházející se vlevo od karty Vrstvy). Podle barevného rozlišení jsme schopni zjistit potřebné informace. Rozsah využívaných dat je však omezen pouze na údaje Českého statistického úřadu a na informace týkající se životního prostředí. Přesto můžeme zjišťovat velmi zajímavé informace.

Budeme-li se zajímat o hustotu zalidnění v okolí Rusavy, vybereme tematickou úlohu Hustota zalidnění. Z karty Vrstvy necháme zobrazeny hranice územních jednotek, obce a komunikace a hustotu zalidnění (obr. 1). Z obrázku 1 je patrné, že v okolí Rusavy je hustota obyvatel asi 2 až 14 obyvatel na km².

Autorku příspěvku také zajímalo, zda má obec Rusava veřejnou knihovnu. Přímo v Rusavě i v jejím okolí se nachází jedna knihovna (obr. 2). Další zajímavou otázkou je vybavení obce Rusava kanalizací. K velkému překvapení bylo zjištěno, že v okolí Rusavy není kanalizace vybudována vůbec (obr. 3).

K dalším součástem mapových služeb patří i vektorová mapa pozemních komunikací. Tato mapa umožňuje zkoumat, jaké objekty se v dané lokalitě vyskytují. Z obrázku 4 je patrné, že v Rusavě je pouze 7 mostů (malá kolečka na silnici).

Jinou, velmi zajímavou možností, je měření vzdálenosti mezi dvěma náhodně zvolenými místy. Protože máme zájem vidět dvě konkrétní místa v okolí Rusavy, musíme si vybrat jiný, způsob zobrazení mapy. Zobrazení podrobnějších informací nejlépe nabízí topografická mapa Armády České republiky. Hledáme možnosti, jak se dostat z místa A (zelený text RUSAVA pod kótou 536 vlevo nad obcí Rusava) do místa B (část Ráztoka, kóta 422). První možností je chůze po silnici (obr. 5). Délka trasy je delší než 4 km (přesněji 4 406 m). Druhou možností je vydat se přímo bez ohledu na překážky v terénu, což bude činit přes 2,5 km (přesněji 2 566 m). To znázorňuje obrázek 6.

To, co mapové služby Portálu veřejné správy nedovedou, je plánování tras. Zde byly použity údaje internetové stránky <http://www.mapy.cz/>. Můžeme si naplánovat trasu ze Zlína do Rusavy a zjišťovat jednak nejrychlejší trasu, která trvá 40 minut a její délka je 30,88 km a jednak nejkratší trasu, která trvá 45 minut a její délka je téměř 30 km (přesněji 29,48 km). Můžeme si detailně prohlédnout celou trasu.

Internetová stránka <http://www.mapy.cz/> umožňuje také vyhledávat nejbližší objekty v okolí. Pokud se budeme zajímat o nejbližší ubytování v Rusavě, najdeme celkem tři možnosti, které jsou označeny číselně. V pravé části obra-

zovky si můžeme prohlížet i detailní informace o daném místě. Chata Jestřabí však nebyla v seznamu ubytování na Rusavě nalezena.

Po zadání klíčového slova „rekreační zařízení“ se chata Jestřabí již objeví. Nyní si z tohoto místa naplánujeme cestu do Zlína. Výsledkem bude popis trasy s uvedením počtu kilometrů na jednotlivých úsecích trasy a odhadovaný čas přepravy. Celková délka trasy bude cca 31 km (přesněji 30,88 km) a cesta bude trvat 40 minut, což je vlastně nejrychlejší cesta.

Jak již bylo řečeno, je rozsah zjišťovaných informací, které jsou k dispozici na Portálu veřejné správy, dost omezený. Můžeme čerpat podklady o životním prostředí, údaje z integrovaného registru znečišťování, hranice územních jednotek, volebních obvodů a příslušných úřadů, lze využívat vojenské i staré mapy, údaje Českého statistického úřadu týkající se obyvatelstva a vybavenosti obcí, údaje České pošty, zjišťovat informace o kvalitě koupacích vod a omezené informace o dopravě. I když není tématických úloh tak velký počet, umožňují i tyto informace pomoci naléznout odpovědi na otázky, které nás zajímají. Uživatelé, kteří nedovedou ještě pracovat s geografickými informačními systémy, získají alespoň základní přehled možností mapových služeb Portálu veřejné správy i informační stránky <http://www.mapy.cz/>. Dalším krokem pro ně již bude naučit se pracovat s konkrétním geografickým informačním systémem.

Další možnosti výuky předmětu Informatika ve veřejné správě

Z výše uvedených popisů je patrné, že geografické informační systémy si pozornost určitě zaslouhují. Ukázky uvedené na přednášce STAKAN 2007, které doplňuje krátká videoukázka, jsou první vlaštvou pro inovaci předmětu Informatika ve veřejné správě.

Přesto je vhodné ukázat a naučit studenty oboru Veřejná správa a regionální rozvoj pracovat prakticky s konkrétním produktem geografického informačního systému minimálně v rozsahu dvou až tří cvičení. Tento krok bude podle názoru autorky velkým oživením výuky.

V souvislosti s vypracováním seminární práce na téma analýzy informačních toků se nabízí otázka rozšíření obsahu tohoto předmětu o analýzu rizik informačních toků a informačních systémů v oblasti veřejné správy. Předmět lze rozšířit i o aplikaci metod kontroly a auditu ve veřejné správě, jednoduchých metod pro hodnocení veřejných projektů a veřejných zakázek, pomocí nichž by bylo možné řešit studenty vybraný problém týkající se veřejné správy. Zvažuje se i výuka krizového řízení ve veřejné správě v podobě kon-

krétního softwarového produktu. Nabízí se celá řada možností, které obohacují obsah tohoto předmětu ve všech jeho stránkách.

V neposlední řadě bude tento předmět inspirací pro vypracování bakalářských i diplomových prací a současně jednou z oblastí vědeckovýzkumné činnosti Ústavu informatiky a statistiky.

Autorka příspěvku přivítá jakékoliv podněty i kritické připomínky týkající se geografických informačních systémů a problematiky veřejné správy.

Internetové zdroje: Odkazy byly funkční k 15. červnu 2007.

Reference

- [1] Ministerstvo vnitra. *Portál veřejné správy České republiky* [online]. 2003-2007 [cit. 2007-06-15]. Dostupný z WWW:
<<http://geoportal.cenia.cz/mapmaker/cenia/portal/>>.
- [2] Seznam.cz. *Mapy.cz – mapa Evropy, České republiky, plány měst a obcí v ČR* [online]. 1996-2007 [cit. 2007-06-15]. Dostupný z WWW:
<<http://www.mapy.cz/#>>.

Adresa: Ing. Miroslava Dolejšová, Ph.D., ÚSKM, Fakulta managementu a ekonomiky, Univerzita Tomáše Bati ve Zlíně, Mostní 5139, 760 01 Zlín

Telefon: +420 576 032 845

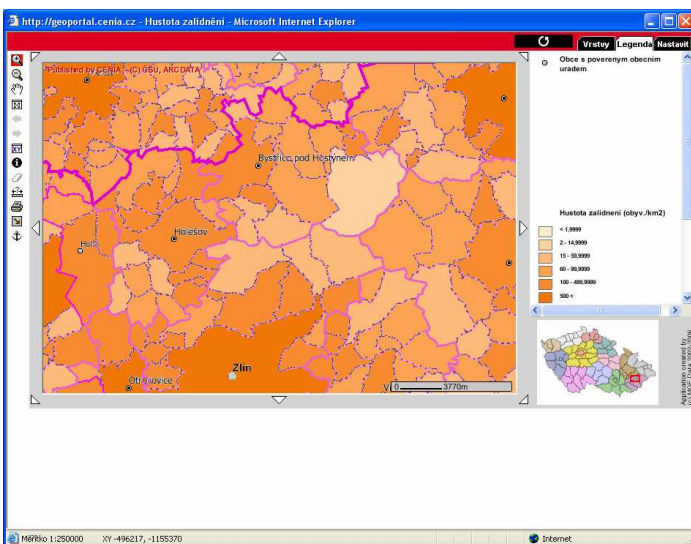
E-mail: dolejsova@fame.utb.cz

Stojí za prozkoumání!

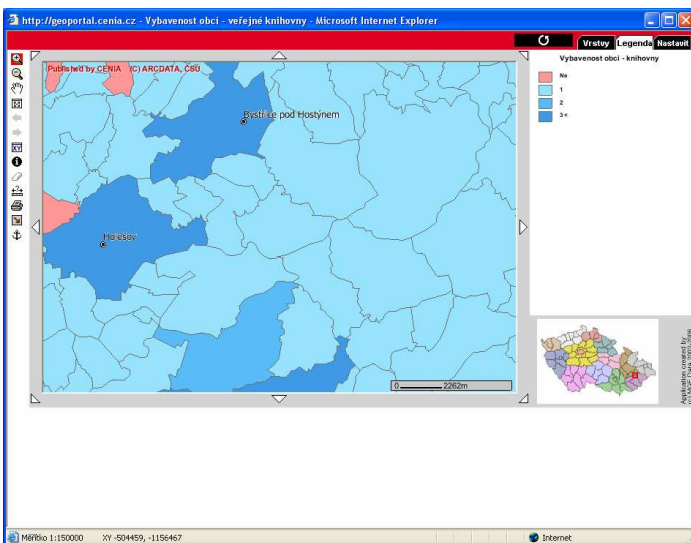
Vítejte ve světě přednášek a *Open Source Software*...

<http://www.linuxrsp.ru/win-lin-soft/table-eng.html>
<http://sf.net/> <http://gnu.org/> <http://www.oss.cz/>
<http://sk.openacademy.eu/> <http://www.aplimat.com/>
<http://tex.mendelu.cz/> <http://live.contextgarden.net/>
<http://www.video.muni.cz/> <http://www.avc-cvut.cz/>

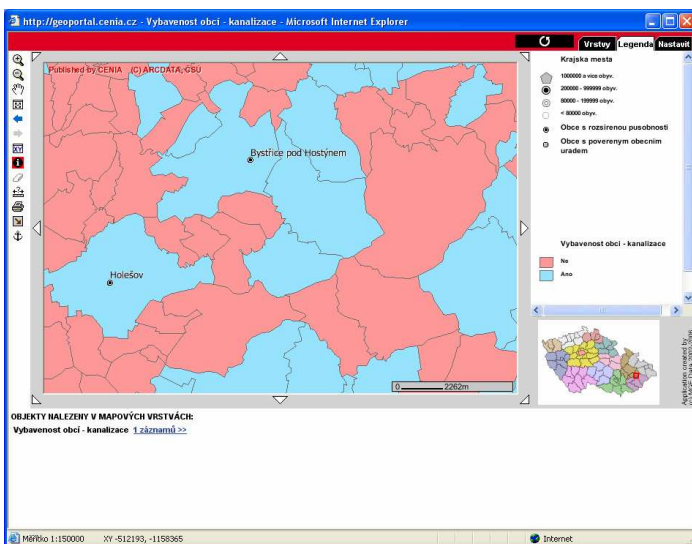
Přejeme Vám hezké zkoumání!



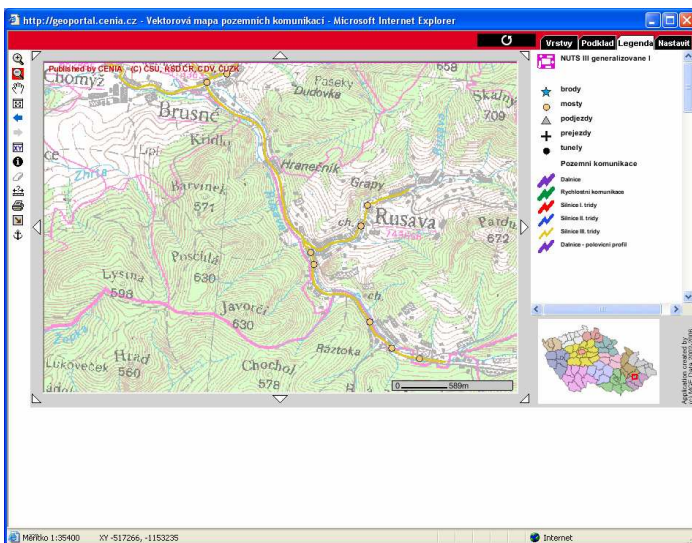
Obrázek 1: Hustota zalidnění v obci Rusava [1]



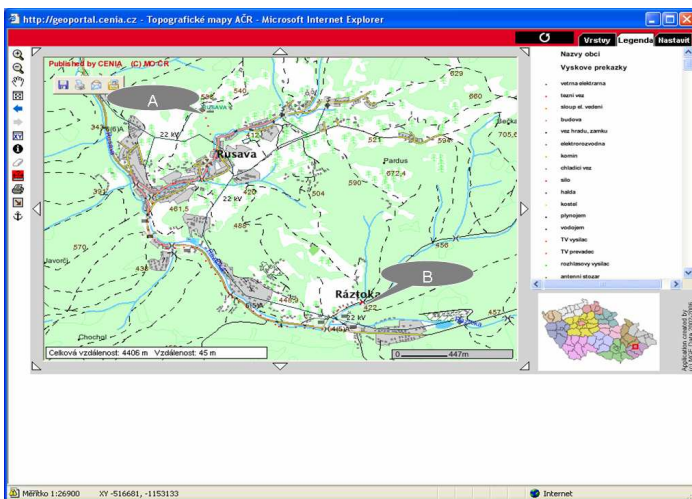
Obrázek 2: Vybavenost obce Rusava veřejnou knihovnou [1]



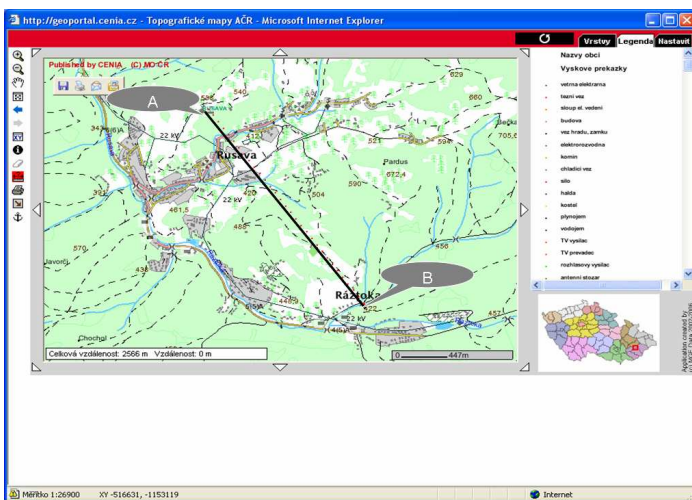
Obrázek 3: Vybavenost obce Rusava kanalizací [1]



Obrázek 4: Vektorová mapa pozemních komunikací [1]



Obrázek 5: Měření vzdálenosti mezi dvěma zvolenými místy: delší trasa [1]



Obrázek 6: Měření vzdálenosti mezi dvěma zvolenými místy: přímá trasa [1]

Pozvánky na konferencie

COMPSTAT 2010

The 19th COMPSTAT symposium will take place in Paris, France, on August 22–27, 2010. The official website is <http://compstat2010.fr/>.

Keynote speakers (taken on March 9, 2010):

- Luc Devroye, School of Computer Science, McGill University, Montreal: „Generating Random Variables on a Computer“
- Lutz, Division of Biostatistics, German Cancer Research Center, Heidelberg: „Computational Statistics Solutions for Modern Biomedical Research: A Challenge and Chance for Both“
- David Hand, Statistics Section, Imperial College, London: „The Laws of Coincidence“

Tutorials:

- Lynne Billard, University of Georgia, United States: „Statistical Approaches for Complex Data“
- Jean-Michel Marin, Université Montpellier II, France: „Bayesian Computation Method“
- Petr Berka, Jan Rauch, University of Economics, Czech Republic: „Machine Learning and Association Rules“

Invited sessions are as follow: ABC methods for genetic data; Algorithms for Robust statistics; Brain imaging; Computational Econometrics; Computer-intensive actuarial methods; Data stream mining; Functional Data Analysis; Kernel methods; Monte Carlo methods in system safety, reliability and risk analysis; Optimization heuristics in statistical modelling; Spatial statistics and epidemiology; KDD Session: Topological learning; ARS Session: Financial Time Series, and IFCS Session: New developments in two or higher mode clustering; model based clustering and reduction for high dimensional data.

MEDSTAT 2010

Residence Hotel Donovaly, 26. – 28. marca 2010.

Konferencia je určená pre pracovníkov rezortu zdravotníctva, ktorí sa zaujímajú o praktické využitie štatistických metód a softvérových produktov vo svojej výskumnej, či každodennej práci. Program sa skladá z výučbových prednášok a z prednášok ilustračných.

Oficiální stránky viz <http://medstat.meu.zoznam.sk/>.

Evropsko-česko-slovenská T_EXová událost roku 2010

Blíží se 20. výročí založení Československého sdružení uživatelů T_EXu a oslava při této příležitosti bude ve mlýně na Brejlově (<http://www.brejlov.cz/>) při spojených, částečně se překrývajících se konferencích, tedy čtvrtého ročníku setkání uživatelů CONTEXTu a třetího ročníku konference T_EXperience, která se snaží navázat na konferenční řadu SLT z dob dřívějších.

Na setkání CONTEXTistů, od pondělí do soboty, 13. – 18. září 2010, lze očekávat příjezd Hanse Hageny, Taco Hoekwatera, Arthura Reutenauera a dalších evropských T_EXových špiček.

Oficiální web setkání je <http://meeting.contextgarden.net/2010/>.

Na druhou konferenci, od čtvrtka do neděle, 16. – 19. září 2010, tamtéž, jsou již nyní přislíbené přednášky a tutoriály:

- Petr Olšák: Matematika a vnitřní algoritmy T_EXu (tutoriál).
- Jan Šustek: Využití T_EXu při přípravě *Mezinárodní matematické soutěže Vojtěcha Jarníka*.
- Karel Horák:
 - Cvičení s MetaType (tutoriál).
 - Pořád je co zlepšovat.
- Zdeněk Wagner: Daňová evidence pomocí L^AT_EXu, XML, Tcl/Tk a Subversion.
- Roman Trušík: Sazba knihy *Bibliografie americké literatury v českých překladech: 2000 & 2010*.
- David Březina: O písmu Scholar.
- Michal Mádr, Pavel Stríž: Zkušenosti s dokončením českého překladu *The Not So Introduction to L^AT_EX 2_ε* po 10 letech (zpráva).

Detaily budou v nejbližším Zpravodaji a webové stránky konference jsou <http://striz9.fame.utb.cz/texperience/>.

Již nyní prozradíme, že on-line registrace na první konferenci je zveřejněna na <http://meeting.contextgarden.net/>, na T_EXperience 2010 se lze registrovat kdykoliv emailem u Pavla Stríže na striz@fame.utb.cz.

Obsah

<i>Hana Šlachťová, Hana Tomášková, Pavla Polaufová, Ivan Tomášek, Jiří Michalík, Anna Šplíchalová</i>	
Příklady použití GIS při hodnocení zdravotních dat	1
<i>Jan Růžička, František Klímek</i>	
Prostorová data	12
<i>František Klímek, Jan Růžička</i>	
GISAK LiveCD	18
<i>Miroslava Dolejšová</i>	
Zařazení geografických informačních systémů do výuky předmětu Informatika ve veřejné správě	25
<i>Redakce</i>	
Pozvánky na konference	35

20 let České statistické společnosti

V letošním roce slaví naše společnost 20 let od svého založení v roce 1990. Výbor společnosti si dovoluje pozvat všechny členy na Brněnské statistické dny, které se budou konat 3.–4. září 2010 a v rámci kterých proběhne i slavnostní zasedání ČStS pod názvem ‘*Quo vadis Stochastica?*’ Na oba dny je připraven nejen odborný, ale i večerní společenský program. Přihlášky a podrobnější informace jsou na stránce společnosti <http://statspol.cz/>.

Informační Bulletin České statistické společnosti vychází čtyřikrát do roka v českém vydání. Příležitostně i mimořádné české a anglické číslo.

Časopis je zařazen na Seznamu Rady, více viz <http://www.vyzkum.cz/>.

Předseda společnosti: doc. RNDr. Gejza DOHNAL, CSc.

ÚTM FS ČVUT v Praze, Karlovo náměstí 13, Praha 2, CZ-121 35

E-mail: gejza.dohnal@fs.cvut.cz

Redakční rada: prof. Ing. Václav ČERMÁK, DrSc. (předseda), prof. RNDr. Jaromír ANTOCH, CSc., doc. Ing. Josef TVRDÍK, CSc., RNDr. Marek MALÝ, CSc., doc. RNDr. Jiří MICHÁLEK, CSc., doc. RNDr. Zdeněk KARPÍŠEK, CSc., prof. Ing. Jiří MILITKÝ, CSc.

Technický redaktor: ing. Pavel Stríž, Ph.D., striz@fame.utb.cz

Informace pro autory jsou na stránkách <http://www.statspol.cz/>

ISSN 1210–8022