

## INFORMAČNÍ SYSTÉMY VEŘEJNÉ SPRÁVY JAKO ZDROJ DAT PRO MIKROANALÝZY V MĚSTSKÉM PROSTŘEDÍ

Jiří HORÁK<sup>1</sup>

<sup>1</sup> Institut geoinformatiky, HGF, VŠB-TU Ostrava, 17. listopadu 15, 70833, Ostrava-Poruba, CZ  
*jiri.horak@vsb.cz*

### Abstrakt

Informační systémy centrálních i místních orgánů veřejné správy představují potenciálně významné datové zdroje pro realizaci mikroanalýz, dovolující podrobně monitorovat a modelovat časoprostorovou situaci v území a využít ji v nových aplikacích. Jsou příležitostí, jak identifikovat významné trendy, lépe pochopit a vysvětlit procesy probíhající v území, poskytovat včasné varování a přijímat cílená lokální opatření, které mohou např. eliminovat hrozby nárůstu nežádoucích sociálních jevů. Real-time propojení na vybrané zdroje dovoluje vytvářet nové formy aplikací, které uživatelům zejména mobilních zařízení poskytuje vysoce aktuální a přesný servis při dotazování na požadované služby např. v dopravě. Využití IS veřejné správy je však spojeno s řešením řady problémů počínaje legislativními a etickými požadavky, řešením harmonizace dat, zajištěním jejich konzistence, zejména při propojování dat z více systémů. Příspěvek charakterizuje hlavní příležitosti, výhody, nevýhody a dokumentuje na příkladech možnosti využití zejména v městském prostředí.

### Abstract

Information Systems of the Public Sector as a data source for micro-analysis in urban environment: Information systems of both central and local governments represent important data sources for micro-analysis enabling to monitor and model spatio-temporal situation and utilize them in new applications. They provide possibilities how to identify trends, better understand and explain processes running in a given territory, trigger early warnings and launch well-targeted measures which can help to i.e. eliminate societal threats. Real-time links to selected sources enable to develop new applications especially for the mobile platform providing highly updated and accurate services i.e. for transport optimisation. The exploitation of public sector information systems faces many issues including legislative and ethical requirements, data harmonisation, and consistency improvements. The paper identifies main plus and contras, opportunities, and describes selected examples of application especially in urban environment.

**Klíčová slova: prostorové mikroanalýzy; veřejná správa; město; informační systémy.**

**Keywords: spatial microanalysis; public administration; urban analysis; information system.**

### ÚVOD

Důsledkem narůstající digitalizace veřejné správy je zvyšující se počet informačních systémů různého rozsahu od základních registrů až po různé místní evidence lokálního významu. Ne vždy musí být přímo udržovány veřejnou správou, ale mohou být provozovány různými servisními organizacemi v důsledku outsourcingu různých forem.

Obecně můžeme mluvit o Public Sector Information PSI, tedy o informacích veřejného sektoru. Tímto termínem se rozumí všechny informace (dokumenty, databáze, mapy, data) které produkuje, sbírá a shromažďuje tzv. veřejný sektor, což jsou instituce, úřady a organizace zajišťující veřejné služby a financované z veřejných prostředků (PSI, 2003). Směrnice 2003/98/EC o opakovaném využívání informací veřejného sektoru ze dne 17. listopadu 2003 vyžaduje, aby informace vyprodukované jednou veřejnou institucí byly použitelné i pro jiné účely, a to jak v samotném veřejném sektoru, tak i mimo něj (prostřednictvím zpřístupnění informací veřejnosti).

Jeden z možných přístupů k rozdělení PSI poskytuje studie projektu MEPSIR, která vymezila následujících 6 tematických domén PSI:

- hospodářské informace
- geografické informace
- meteorologické informace
- dopravní informace
- sociální informace
- právní informace

Bez ohledu na použité členění je zřejmé, že k nejvýznamnějším skupinám PSI patří socioekonomické informace, což je dáno potřebami veřejné správy. Patří k nim např. evidence obyvatel, budov, bytů, dopravní infrastruktury, pozemků, nezaměstnaných, sociálních dávek, mezd/platů, sociálního a zdravotního pojištění, ekonomických subjektů, zdravotní evidence, školská evidence, evidence kulturních zařízení, evidence trestných činů a přestupků apod.

Samozřejmě by bylo možné jmenovat jednotlivé základní zdroje těchto informací, z nichž část patří do soustavy základních registrů (registr obyvatel, registr osob, RUIAN), nicméně je skutečností, že i v této oblasti existuje a je dále rozvíjeno z řady důvodů více evidencí. Poměrně široký přehled geoinformačních databází pro popis a charakteristiku regionů poskytuje Mikulík et al. (2008), řadu z nich lze využít i pro charakteristiky urbánního prostředí.

Pokud je pravdivý výrok, že 80% dat má prostorovou složku (možná je na čase uznat, že podle principu crowdsourcingu se toto mnohokrát opakované prohlášení již stalo pravdou), pak je zřejmé, že propojování, znovuužívání a zpřístupňování veřejných informací je samozřejmě součástí budování geoinformační infrastruktury a rovněž jádrem směrnice INSPIRE.

Pokud se zamyslíme nad zdroji dat, je evidentní, že potenciálně největší rozdíly mezi celkovými a dílčími výsledky (global x local) můžeme zjistit u dat s nejvyšším prostorovým rozlišením. Např. pro hodnocení nezaměstnanosti v kraji mají výsledky získané za části obcí jiný charakter než hodnocení dat za obce. Existuje zde největší nepokrytá poptávka, zejména v souvislosti s územním plánováním pro menší územní celky – ORP a menší. Územní podmínky pro udržitelný rozvoj území se mají hodnotit (aktualizovat) co 2 roky. V takovém případě jsou údaje SLDB nevhodné a jiné údaje statistická služba neposkytuje za menší jednotky především z důvodů finančních a časových. Doplňme, že požadavky na podobně podrobná a aktuální data se objevují rovněž v řadě dalších aplikací, např. v krizovém řízení.

Problém u cenových šetření a u dotazníkových zjišťování spočívá především v jejich pořizování, zejména v jejich ceně, stáří, pracnosti získávání a zpracování. Kromě toho se ale přidává i rostoucí nevěrohodnost, protože stále více lidí (respondentů) není ochotno sdělovat pravdivé údaje o sobě ať již z důvodu ochrany osob nebo se šetřením nesouhlasí a chce svůj nesouhlas manifestovat poškozením výsledků šetření (případně jde o nějakou formu recese). Proto do budoucna tato data budou stále méně spolehlivá.

Někteří vnímají dotazníková šetření jako obtěžování lidí - zejména pro ty údaje, které jsme schopni získat přímou cestou z evidencí. U řady požadovaných údajů teoreticky existují záznamy ve veřejném sektoru – zaměstnání, vzdělání, příjem, registrované vyznání, dlouhodobá migrace (stěhování), část veřejnou službou registrované krátkodobé migrace (evidence slevových průkazů ve veřejné dopravě, čítače), podmínky bydlení, rodinné poměry.

Zbývá zjistit „pouze“ názory, postoje, spokojenost, plány. Cestou k nim mohou být různé formy tzv. sentimental analysis (Yi et al., 2003) s využitím sociálních sítí či jiných internetových služeb.

Vzhledem k nedostatku aktuálních dat s vysokým časovým a prostorovým rozlišením lze očekávat nejvýznamnější využití ISVS právě u prostorových mikroanalýz neboli lokálních prostorových analýz.

Prostorové analýzy jsou souborem technik pro analýzu a modelování lokalizovaných objektů, kde výsledky analýz závisí na prostorovém uspořádání těchto objektů a jejich vlastností (Horák, 2011).

Data ISVS zpravidla neobsahují přímou prostorovou referenci (souřadnice). Proto jejich společným rysem je, že jsou geokódovány adresou.

Základní zpracování zpravidla zahrnuje harmonizaci adres, geokódování a agregaci primárních dat.

## RÁMCOVÝ POSTUP ZPRACOVÁNÍ DAT IS VS

Základní postup při zpracování dat ISVS vychází ze skutečnosti, že zde zpravidla chybí jedinečný prostorový identifikátor (georeference) a/nebo popis prostorové reference obsahuje chyby. Potřebné jsou následující 4 kroky:

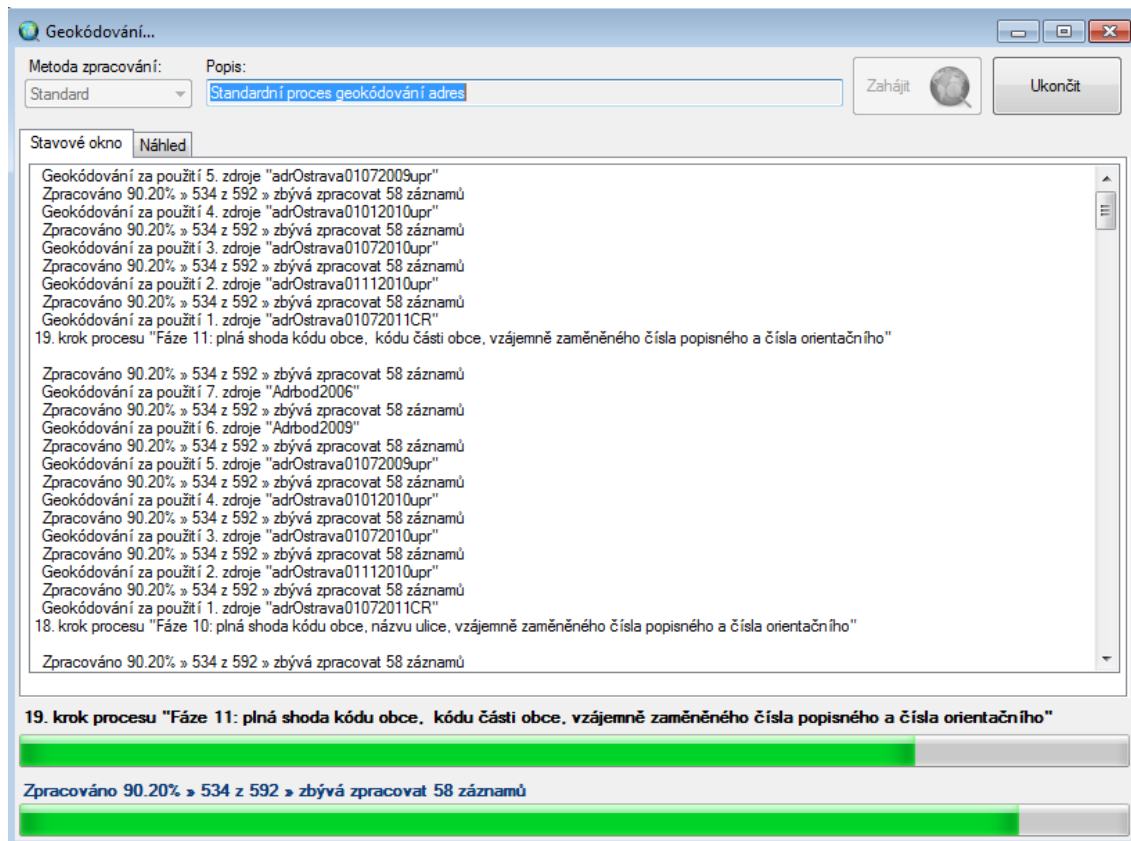
1. Zpravidla je nutné změnit strukturu dat a dekomponovat některá pole (adresa).
2. Harmonizace představuje proces unifikace struktur a unifikace popisu adresy, včetně identifikace a nahrazování abnormálních (resp. v číselnících neznámých) jmen. Samozřejmě přichází do úvahy i více sofistikovaná řešení se schopností se automaticky vypořádat s chybami a mírně odlišnými zápisy (zpracování nestrukturovaných dat).
3. Geokódování jako proces doplnění prostorových identifikátorů s přímou souřadnicovou referencí.

Ke geokódování může být využita aplikace typu AddressGeocoder.

AddressGeocoder je aplikace, vytvořená autory Fojtík a Horák primárně pro potřeby MPSV ČR, určená pro úpravy a identifikaci strukturovaných adres libovolných databázových zdrojů, pro které identifikuje odpovídající objekty v adresních vrstvách a doplní k nim odpovídající souřadnice. Provádí import zdrojových dat do standardizované struktury, přitom je schopna automaticky identifikovat odpovídající sloupce nebo ručně vybrat vhodné sloupce, pokud u některých není automatické mapování atributů možné či vhodné (obr. 1). Při vlastním geokódování iterativně prohledává referenční sady adres, aplikuje sadu pravidel a provádí přiřazování souřadnic v pořadí od nejspolehlivěji určených k nejméně jistým (obr.2). Aplikace proti běžným nástrojům pracuje s více adresními atributy, dokáže zpracovat i neúplné nebo částečně špatně zadané adresy. Ke každému zpracovanému záznamu připojí rovněž informaci o způsobu identifikace adresy (a tedy přiřazení souřadnic).

NázevSloupce	Mapovat jako	Velikost [B]	Numerická přestnost	Datový typ	AllowDBNull	IsReadOnly	IsUnique
PVM_IDCISL	Identifikátor	40	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ZKRATKA		60	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
FIR_NAZEVI		254	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
FIR_ICO		8	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PRAC_NAZEVI		254	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OBEC	Obec	48	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CAST_OBCE	Část obce	48	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ULICE	Ulice	48	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
OKRES		32	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ADR_PSC	PSC	8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
ADR_CISLO_		8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
BDR_CISLO_	Číslo popisné (dom...	8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CDR_CISLO_		1	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PVM_MZDA_M		8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
QVM_MZDA_M		8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PVM_POCET_		8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
QVM_POCET_		8	15	System.Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CZ_ISCO_KO		5	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PVM_CZ_ISC		254	255	System.String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PVM_OD		8	255	System.DateTime	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Obr. 1 Mapování atributů



Obr. 2 Postupné geokódování

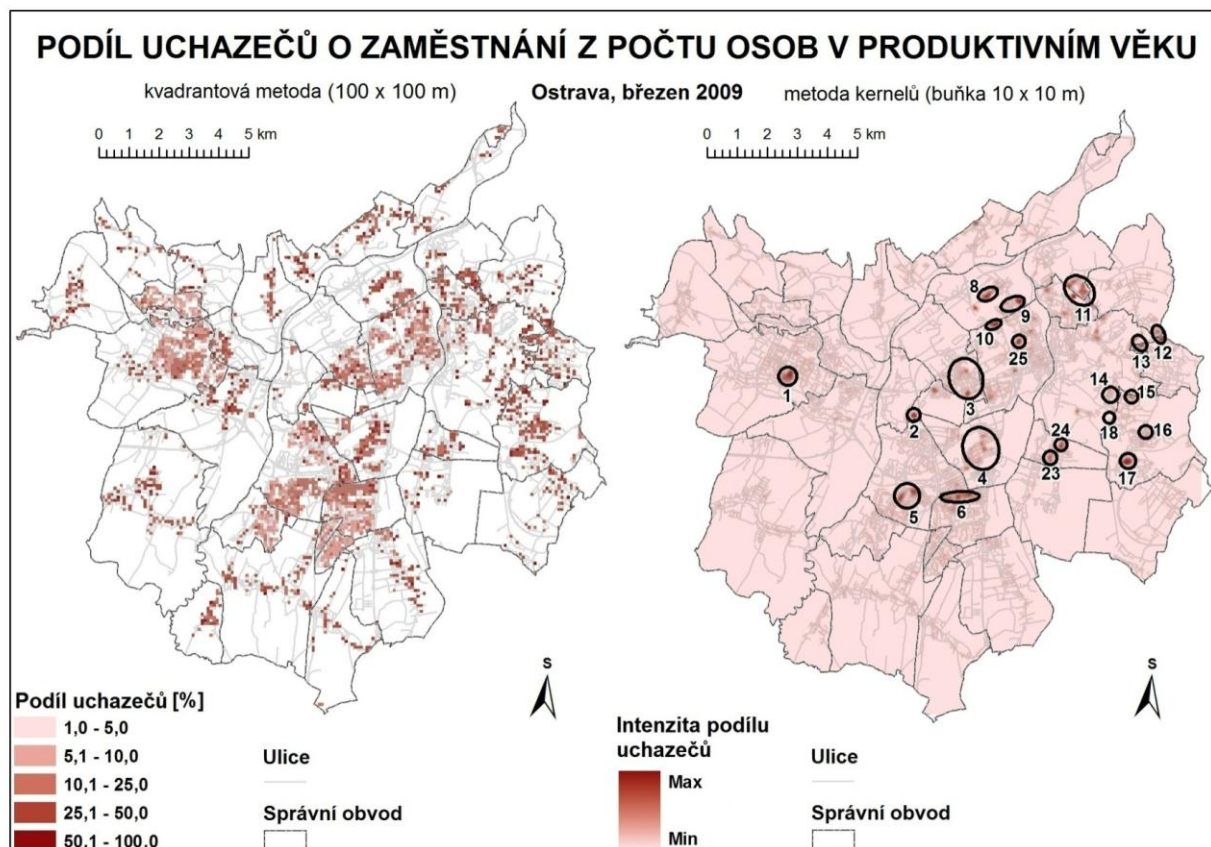
#### 4. Agregace

Geokódovaná data z registrů (a různých evidencí) mohou být snadno agregována do různých prostorových jednotek včetně adresních bodů, administrativních územních jednotek, územních jednotek jiného než administrativního určení, do buněk pravidelné (či nepravidelné) sítě, do lokalit (polygony či areály specifického určení, např. jako prostorové vymezení sociálně vyloučených lokalit). V některých případech se provádí víceúrovňová agregace. Na závěr by měla být provedena kontrola konzistence, včetně analýzy při ní odhalených problémových míst.

Při agregaci do pravidelné sítě buněk je nezbytné použít vhodné nastavení sítě, zejména velikosti buňky. Při volbě velikosti lze jednoduše použít minimální velikost nejmenšího detailu, který se má ve výsledku projevit (resp. jedna polovina velikosti). Pokud není definován takový požadavek, můžeme použít metody, které optimalizují výslednou velikost buněk sítě, jako jsou segregáční míry, Gini koeficient (maximální segregace by měla ukazovat na nejlepší oddělitelnost jednotlivých kategorií) či míry korelace (hledáme optimum mezi vlivem ekologické chyby, malou datovou základnou a zvýšenou korelací při zvyšování velikosti buňky)

Při vytvoření hierarchického agregačního systému lze aplikovat nástroje Business Intelligence k průzkumu dat.

Další možností je agregace pomocí jádrových odhadů.



Obr. 3 Srovnání 2 agregačních metod (Ivan et al. 2010)

## PŘEDNOSTI A NEDOSTATKY LOKÁLNÍCH PROSTOROVÝCH ANALÝZ DAT ISVS

### Výhody lokálních prostorových analýz

- Vysoké prostorové a časové rozlišení
- Aktuálnost
- Detailní studium vybraných údajů dokumentujících socioekonomické jevy
- Popis reálného geografického či urbánního prostředí, různých omezení, bariér. Snadnější určování příčin studovaného jevu.
- Vhodná aplikovatelnost pro veřejnou správu ve formě iniciace cílených místních opatření
- Vyhnutí se problému ekologické chyby (Longley et al., 2005)
- Možnost využití běžně nedostupných, ale cenných indikátorů typu hodnocení jevu v úzce vymezených věkových kategoriích (např. migrace, nezaměstnanost atd.)

Hlavní efekt je v možnosti získání rychlé zpětné vazby pro posouzení účinnosti či jakékoliv reakce na nově aplikovaná opatření. Např. zřízení nové policejní služebny a dopady na kriminalitu v okolí.

Neméně významné je již avizované usnadnění navrhování včasných a vhodně přizpůsobených místních opatření např. v oblasti prevence kriminality a vymáhání práva, prevence insolvence, zdravotnických a sanitárních opatření, zvýšení počtu terénních pracovníků či spuštění nových sociálních programů.

## Nevýhody lokálních prostorových analýz

Primární problémy – problém výběru správných dat:

- Neurčitost kvantity – registry jsou „živými“ systémy, kde kontinuálně probíhá editace dat, korekce a doplňování údajů. Tyto úpravy znamenají přidávání nebo odstraňování některých záznamů, ale také změny některých údajů v evidenci, např. adres. Data se mohou lišit v závislosti na termínu, kdy je proveden jejich export či zpracování. Např. počet uchazečů o zaměstnání v Ostravě ke stejnému rozhodnému dni vzroste během půl roku třeba i o 200.
- Lokalizační neurčitost – typicky existuje více než 1 adresa u objektu (např. kontaktní adresa a trvalá adresa osoby). Nemusí být jednoznačné, kterou adresu vybrat pro lokalizaci osoby. Navíc se tento výběr může lišit podle účelu. Např. pokud bychom použili údaje pro mapování počtu osob skutečně se vyskytujících na jistém území, bude vhodnější kontaktní adresa, ale pokud budeme údaje propojovat s jiným registrem, vedeným podle trvalého bydliště (např. registr obyvatel), bude vhodnější realizovat výběr adres trvalého bydliště a dosáhnout tak vyšší konzistence výsledků při integraci dat ze dvou systémů (a možné výpočty podílů apod.). U jiných registrů nejsou jednotlivé adresy rozlišeny, mají všechny stejný popis a prioritu. To je případ např. sériových trestných činů, kdy se v Evidenci trestního řízení jako 1 záznam evidují případy, které se staly v průběhu jistého období na více místech.

Sekundární problémy – problémy správného zpracování:

- Nejistota identifikace správné adresy (geokódování), resp. správného párování adres (mezi zdrojem a referenčními daty se souřadnicově vyjádřenou polohou) – nejistota je tím vyšší, čím je nižší úroveň strukturovanosti (ve smyslu dekompozice) zápisu adresy a čím nižší je míra využití jednoznačných územních identifikátorů. Volný text samozřejmě způsobuje řadu problémů s variabilitou zápisu, přirozeným výskytem chyb a ani v případě zcela správného a přesného zápisu nemusí být automatizovaný proces dekompozice adresy zcela spolehlivý. Problémy způsobují rovněž rozdílné číselníky (resp. spíše jejich časová nekonzistence). Výsledné geokódování může proběhnout s téměř 100% pokrytím, ale nedostatečná schopnost rozpoznání správné adresy může vést v nejasných případech k selhání ve výběru jedné adresy a následně k náhradnímu použití nadřazeného geografického objektu, jako je ulice či dokonce část obce. Některé postupy zpracování (např. jádrové odhady) pak vyžadují, aby nedocházelo k umělé koncentraci v těžišti nositele (geografického objektu) a proto se provádí náhodné rozmitání (jittering, Gorr, Kurland 2012) bodů v rozsahu nositele. Tedy např. v případě ulice se provede náhodné umístění bodu podél ulice (solistikovanější postup pak vybere náhodně jeden ze skutečně existujících adresních bodů podél ulice). Obecně pro některá území může být identifikace adres ztížena rychlými změnami adres, zejména nově urbanizované plochy s intenzivní plošnou výstavbou, ale důvody mohou být i opačné např. u poddolovaných území v souvislosti s rychlou demolicí domů (desítky domů demolovaných v částech obcí na Karvinsku za rok), kdy úspěšnost z tohoto důvodu může klesnout i na 30% (Horáková et al. 2004).
- Narušení konzistence dat – důvodem jsou rozdíly v čase, sémantice a metodice, které se projeví při spojování dat z více registrů. Časový nesoulad je snadno doložitelný, více skryté a o to nebezpečnější jsou uvedené další důvody nekonzistence. Nekonzistence se může projevit např. abnormálními podíly vypočtenými z nekonzistentních dat – typickým příkladem je v daném místě větší počet nezaměstnaných než obyvatel či abnormální index kriminality v místech bez rezidentů. Je vhodné podotknout, že nekonzistence nemusí být jen negativní jev, protože může sama o sobě představovat zajímavý indikátor jiného jevu. Např. výše uvedená nekonzistence u nezaměstnaných je typická pro úřady místní samosprávy, pro ubytovny, ale i pro některé sociálně problémové lokality.
- Ochrana soukromí – práce s daty s vysokým rozlišením vyžaduje odpovídající postupy pro zajištění ochrany soukromí osob, případně organizací, kterých se data týkají. Zejména tvorba výstupů musí podléhat adekvátní úrovni ochrany. Základním požadavkem je, že z publikovaných výsledků nesmí

být možné odvodit individuální data o 1 domácnosti, 1 zařízení či 1 firmě. K základním metodám pro zajištění ochrany soukromí patří filtrování dat (např. publikaci pouze těch údajů, jejichž datová základna je větší než 50 osob nebo 16 domácností pro Small Area Statistics uvádí Martin, 1995), náhodné přidávání či odčítání malých hodnot simulací (k hodnotě buňky se náhodně přidává  $<-1, 0, +1>$  dle Kraak, Ormeling, 1996), permutace dat ("data scramble" viz Batista et al. 2011), zaokrouhlování výsledných statistik (v případě citlivých zdravotnických dat pro malé areály, vymezené populací menší než 150000 osob dle Quinn, 1996) či použití jadrových odhadů, některých indexů a kohortových studií (např. Horák 2010). Způsob znepečnění údajů je důležitou informací pro uživatele a měl by se vždy uvádět. Rovněž vlastní správa dat musí být řízena tak, aby se minimalizovala rizika úniku či zneužití citlivých údajů. Problém ochrany individuálních dat uvádí Mikulík et al. (2008) jako hlavní důvod nemožnosti pracovat přímo se zdrojovými daty.

- Chudost a monotónnost dat – registry veřejné správy jsou vytvářeny pro určité úlohy a na ně je zaměřen jejich datový obsah. Úzce vymezený účel vede přirozeně k úzkému spektru zaznamenávaných údajů. Nemohou být tak komplexní jako např. výsledky specializovaného dotazníkového šetření. To se projevuje monotónností dat, která může vést k jednostrannému pohledu na studovaný jev. Dalším projevem může být nutnost náhrady běžně používaných indikátorů, jež není možné z dat registrů odvodit, za jiné, více nepřímé indikátory. Nepřímé indikátory však zvyšují nejistotu popisu jevu. Např. pro nedostatek primárních dat o vzdělání je použit ukazatel o nejvyšším dosaženém vzdělání uchazečů o zaměstnání, což může být dostatečné jen pro extrémní lokality s velmi vysokou nezaměstnaností (a vysoce pravděpodobně s velmi nízkou úrovní vzdělání).
- Riziko chyb - menší počet údajů vede k vyššímu riziku chyb v důsledku vyšší závislosti na kvalitě lokalizace, záznamu faktů, aktuálnosti, konzistence dat z různých zdrojů

## PŘÍKLADY

Využití ISVS pro prostorové mikroanalýzy je aktuální i ve světě. Např. Kugler (2012) integruje data environmentální s údaji o obyvatelích ve světě s rozdílnou úrovní rozlišení. Je pozoruhodné, že přestože aplikace se připravuje celosvětově, vychází se z dat s největším možným rozlišením dostupným v daném prostoru. V USA využívá Population Microdata, která obsahují individuální data a rovněž data o rodinách s množstvím významných atributů (např. i adresa matky pro posouzení předchozí migrace). Základem zpracování je integrace dat na základě lokalizace (locatio-based integration) a následná agregace.

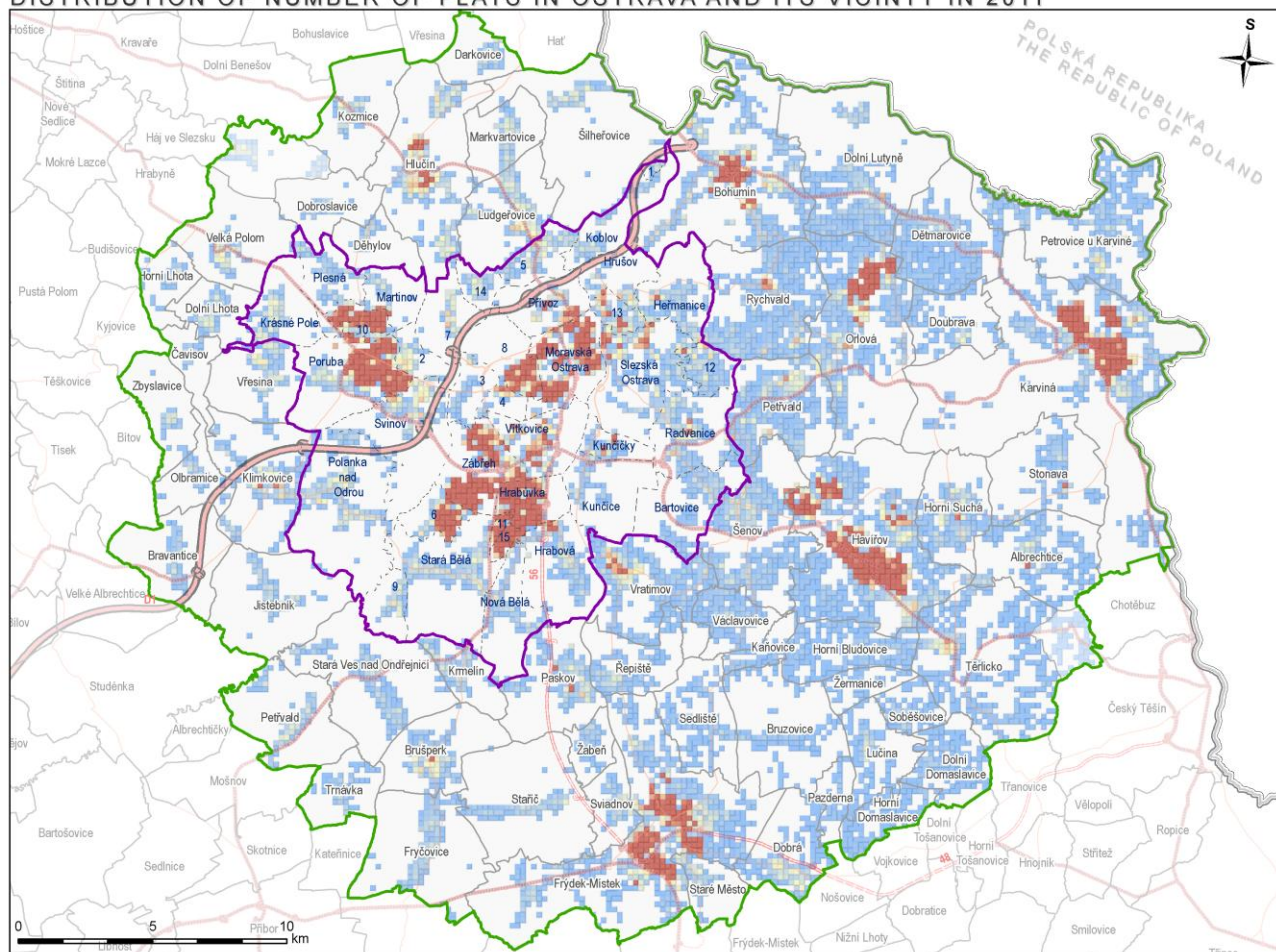
Vlastnosti budov a jejich prostorová konfigurace je předmětem hodnocení a modelování v mikroměřítku jako základ pro model, pomáhající urbanistům v rozhodování (Al-Sayed, 2011).

U nás jako příklad zpracování 1 zdroje můžeme uvést mapu na obr. 4, která využívá data Registru sčítacích objektů a budov ČSÚ a je součástí socioekonomického atlasu Ostravy. Návazně se data propojovala s evidencí vlastníků budov získané z Magistrátu města Ostravy.

Jednou z oblastí, kde se princip spojování dat z různých informačních systémů může výhodně uplatnit, je sociální problematika, zejména sociální vyloučení, vznik a vývoj problémových lokalit (SVL), jejich prostorová, sociální, kulturní a mentální segregace. Je nezbytné monitorovat v území sociální vývoj, sledovat místa, kde se vyostřuje situace z pohledu sociální problematiky (zejména nedostatek příjmů, nezaměstnanost, chudoba, finanční problémy, výskyt sociálně-patologických jevů zvláště spojených s užíváním drog a kriminalitou).

V případové studii SVL v Ostravě (Horák et al, 2010) se studoval problém segregace a sociálně-prostorového oddělení s využitím prostorových mikroanalýz. Jako základní indikátory byly doporučeny ukazatele nezaměstnanosti, nízké vzdělání a extrémně nízký index stáří. Pomocí těchto analýz byly odhaleny další problémové lokality, které nejsou experty dosud jako sociálně vyloučené lokality sledovány.

## ROZLOŽENÍ POČTU BYTŮ V OSTRAVĚ A OKOLÍ V ROCE 2011 DISTRIBUTION OF NUMBER OF FLATS IN OSTRAVA AND ITS VICINITY IN 2011



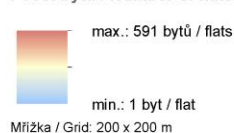
### Administrativní hranice Administration boundaries:

- hranice ČR  
boundary of the Czech Republic
- Obec  
Municipality
- hranice obcí  
boundaries of municipalities
- Část obce  
Municipality part
- hranice části obce  
boundaries of municipality parts
- hranice obce Ostrava  
Ostrava city
- užší jádro aglomerace  
narrower core of agglomeration

### Doprava / Transport:

- Dálnice  
highway
- Rychlostní komunikace  
expressway
- Silnice I. třídy  
1st class roads
- Silnice II. třídy  
2nd class roads

### Počet bytů / Number of flats:



### Číselník územních jednotek / Code list:

- |              |                  |                 |
|--------------|------------------|-----------------|
| 1 Antošovice | 6 Výškovice      | 11 Bálský les   |
| 2 Třebovice  | 7 Hošťálkovice   | 12 Michálkovice |
| 3 Nová Ves   | 8 Mariánské hory | 13 Muglínov     |
| 4 Hulváky    | 9 Proskovice     | 14 Lhotka       |
| 5 Petřkovice | 10 Pustkovec     | 15 Dubina       |

Zdroj dat / Data sources: ČSÚ - Adresní místa k 1.1.2011, RSD 2010  
Institút geoinformatiky VSB-TU Ostrava  
Institute of Geoinformatics Technical University of Ostrava  
Administrativní hranice platné k 1.7.2010  
Administration boundaries valid as of 1.7.2010  
Souřadnicový systém S-JTSK  
Coordinate system S-JTSK

**Obr. 4** Rozložení počtu bytů v Ostravě a okolí v roce 2011 (Hruška-Tvrđý a kol., 2011)

Podobně v případě města Nový Jičín byla použita data z informačního systému města a místního úřadu práce (Glacová, Horák, 2002). Z IS úřadu práce byly použity agregované údaje o uchazečích o zaměstnání - počet, počet žen, počet dle věkového složení a počet v kategorii změněné pracovní schopnosti. Z IS Magistrátu byly použity agregované údaje o počtu obyvatel, počtu žen, početnosti jednotlivých věkových tříd a počtu neplatičů v obecních bytech. Rovněž byly získány souřadnice adres (centroidy budov). Údaje z Magistrátu byly přeneseny na ÚP, zde propojeny s evidencí uchazečů a následně agregovány na čtvercové buňky o rozměrech 600 m, 300 m a 200 m. Následně byly vypočteny vhodné indikátory situace a vizualizovány.

Jednou z důležitých výhod prostorové orientace prováděných mikroanalýz je možnost sledovat změny uvnitř lokalit. Je možné monitorovat a rozlišit lokality s homogenní či variabilní strukturou, s rozdílným vývojem (zlepšování či zhoršování situace).



Jinou oblastí příležitostí pro mikroanalýzy představuje využití dopravních informací. Hromadné automatizované zpracování jízdních řádů dovoluje velmi detailně studovat dopravní obslužnost, její parametry a podmínky (Horák et al., 2007). Hodnocení podmínek dojíždění veřejnou linkovou dopravou může být využito při hodnocení území, analýzách dostupnosti služeb, optimalizaci jejich rozmístění či analýze situace na trhu práce. Pro analýzy je vhodnější využití předzpracované databáze dopravních spojení, která může obsahovat parametry pro každý pár obcí či částí obcí (např. čas, cena, počet přestupů v daný časový interval, včetně cesty zpět, počty spojů v daný interval). Takto byla např. řešena podpora pro Integrovaný portál SSZ MPSV, umožňující vyhledávání volných míst v okruhu dostupném veřejnou linkovou dopravou.

Real-time propojení na zdroje aktuálních dopravních informací umožňuje vytvářet nové formy aplikací, zejména typu LBS, které uživatelům mobilních zařízení poskytují vysoce aktuální a přesný servis při dotazování na požadované služby v dopravě (Raper 2012). Taková data mohou být ale využita i pro zpřesnění hodnocení dostupnosti a mikroanalýzy dopravní situace.

## ZÁVĚR

Využití ISVS pro lokální prostorové analýzy je podle mého názoru nevyhnutelnou budoucností. Je potřebné pracovat na podmínkách a opatřeních, které povedou k jejich efektivnější tvorbě a zejména vytěžování pro nejrůznější potřeby. K těm základním by měly patřit:

- Zajištění kontinuální aktualizace základních registrů (zejména adres a evidence budov) co nejbližší skutečnému stavu.
- Propojit všechny ISVS na soustavu základních registrů, aby se minimalizovaly harmonizační problémy.
- Výstavba vylepšených geokódovacích systémů pro kvalitní a rychlé automatizované georeferencování s minimálními ztrátami dat a co nejmenší prostorovou (lokalizační) neurčitostí.
- Zajistit více operativní a konzistentní propojování dat napříč ISVS tak, aby bylo možné snadno propojovat data z různých systémů a usnadnit řešení problémů rozdílů v sémantice či metodice.
- Zajištění lepších vnitřních kontrol uvnitř ISVS a napříč mezi nimi a ověřování dat. To může zahrnovat primárně doplnění dalších integritních omezení na data a sekundárně sofistikovaná řešení zjištěných nekonzistencí.
- Přiměřený systém kontroly ochrany osobních dat a důvěrnosti dat, který by zajišťoval dostatečnou úroveň ochrany subjektů, ale současně neomezoval zpracovávání dat a jejich analýzy.

V aplikační úrovni může další výzkum směřovat do oblasti podrobného pozorování jevů na mikroúrovni, sledování dopadů vnějších podnětů na místní situaci, studium vzájemného ovlivňování, šíření či předávání atributů, kompenzační a retardační vlivy sousedství apod. To by mělo vést k lepšímu pochopení těchto jevů, jejich podmíněnosti a tedy i možnosti jejich predikce, vhodné prevence či represe, profilování a kategorizace pro snadnější pochopení pravděpodobných scénářů vývoje a vhodných koncepčních řešení.

Závěrem několik rad. Zvažte, jaká data a metody potřebujete. Nevyhýbejte se různorodým datům, mohou vám poskytnout více pohledů na zkoumaný problém, odhalit nedostatky či jednostranné vychýlení 1 zdroje dat. Počítejte však s náročnějším zpracováním, které přináší více úskalí.

## LITERATURA

Batista F., Gallego J., Lavvare C. (2011). The effect of ancillary data in population dasymetric mapping. In proceedings of EFSG 2011, Lisabon.

Glacová M., Horák J. (2002). Analýza situace na trhu práce v měřítku okresního města. [CD-ROM] Sbor. ref. konference GIS Seč 2002 12.-14.6.2002, Seč, 2002, 18 s.

Gorr W.L., Kurland K.S. (2012). GIS Tutorial for Crime Analysis. ESRI press, Redlands, ISBN 978-1-58948-214-2.

Horák J. (2011), Prostorová analýza dat. VŠB-TU Ostrava.

Horák J., Ivan I., Inspektor T., Hruška-Tvrđý L. (2010). Mapping and Monitoring of Socially Excluded Localities in Ostrava City. In Horák et al. Advances in Geoinformation Technologies 2010. 1.vydání. VŠB-TU Ostrava, p. 117-131. ISBN 978-80-248-2145-0.

Horák J., Šeděnková M., Ivan I., Fojtík D. (2007). Databáze dopravních spojení pro Českou republiku a příklady využití. In Sborník semináře „Geoinformačné systémy v doprave“ v Bratislavě, 27.11.2007, 7 stran. ISBN 978-80-227-2767-9.

Horáková B., Horák J., Chrobák P. (2004). Výsledky propojení RSO a UIR-ADR a potenciál pro analýzy trhu práce In Sbor. ref. konference GIS Seč 2004 9.-11.6.2004, Seč, 2004, 14 s. ISBN: 80-86143-28-7 , ISSN: 1213-4163.

Ivan I., Horák J., Inspektor T., Hruška-Tvrđý L. (2010). Využití registru nezaměstnaných pro lokalizování sociálně vyloučených lokalit – případová studie Ostravy. In: Geografie pro život ve 21. století: Sborník příspěvků z XXII. sjezdu České geografické společnosti pořádaného Ostravskou univerzitou v Ostravě 31. srpna - 3. září 2010. Ostrava: Ostravská univerzita v Ostravě. ISBN 978-80-7368-903-2.

Kraak, M., Ormelling, F. (1996). Cartography. Visualisation of spatial data. Londýn, Longman Scientific & Technical. 222 s., ISBN 0-582-25953-3.

Kugler, T. (2012). Terra Populus: integrated data on population and environment. In proceedings of EFSG2012, Praha 2012.

Longley, P.A., Goodchild M.F., Maguire D.J., Rhind D.W. (2005): Geographical Information Systems and Science. Wiley.

Martin, D. (1995). Censuses and the modelling of population in GIS. In Longley, P., Clarke, G. (ed.): GIS for Business and Service Planning. Cambridge, Geoinformation International, s.48-71, ISBN 1-899761-07-1

Mikulík O., Voženílek, V., Vaishar A. a kolektiv (2008). Studium rozvoje regionu založené na vizualizaci geoinformačních databází. UP Olomouc. ISBN 978-80-244-1928-2.

Quinn, M. (1996). Confidentiality. In Elliott P., Cuzick J., English D., Stern R. (ed.). Geographical & Environmental Epidemiology. New York, Oxford University Press, s. 132-140, ISBN 0-19-262235-8.

Raper J.F. (2012), Research challenges and opportunities in open transport data. In proceedings of 1st InDOG conference, Olomouc 2012.

Yi, J., Tetsuya N., Bunescu R., Niblack W. (2003). Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In IEEE Intl. Conf. on Data Mining (ICDM).