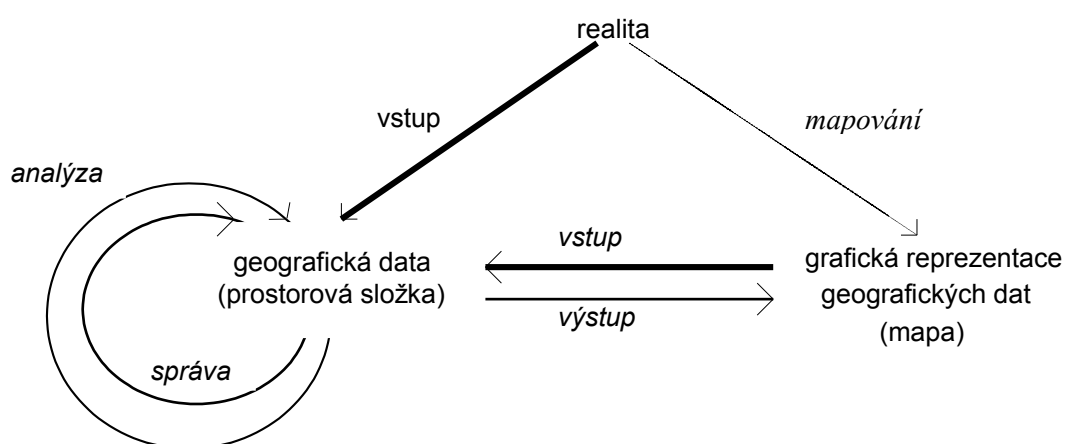


Lekce 8 - Vstup dat do GIS

1. Cíle lekce.....	1
2. Cíle vstupu dat.....	1
3. Kvalita vstupních dat.....	2
3.1 Hledisko mikroúrovně.....	2
3.2 Hledisko makroúrovně.....	2
3.3 Hledisko uživatelské úrovně.....	2
4. Metody vstupu dat do GIS.....	2
4.1 Vstup dat z klávesnice.....	2
4.2 Měření a geodetické výpočty.....	3
4.3 Manuální digitalizace.....	3
4.4 Skenování.....	4
4.5 Konstrukce nad digitálním podkladem.....	4
4.6 Vstup existujících digitálních souborů.....	5
5. Další zpracování pořízených dat.....	5
6. Některé zásady při pořizování prostorových dat.....	5
7. Obsah metodiky pořizování prostorových dat.....	6
8. Zkušební otázky.....	6

1. Cíle lekce

- Provést diskusi o vstupu geografických dat jako jedné ze základních funkcí GIS:
 - vstup
 - management
 - analýza
 - výstup



2. Cíle vstupu dat

Vstup geografických dat do GIS zabezpečuje převod geografických dat do počítačově přístupné formy a do databáze GIS, a to

- převod prostorové složky geografických dat
- převod neprostorové složky geografických dat
- spojení obou složek geografických dat
- vytvoření topologie z geometrických vlastností pořízených dat

Při jednorázovém pořizování rozsáhlých souborů prostorových dat (počet prvků > 10 000) je vhodné navrhnout speciální datový model pořizovaných dat, který může být odlišný od datového modelu cílového GIS. Je možné zvolit rozšíření cílového datového modelu nebo jeho úpravu s tím, že při konverzi dat nedojde ke ztrátě informace. Datový model pro jednorázový vstup dat je volen tak, aby podporoval efektivní pořízení dat a aby umožňoval provádět kontroly jednotlivých kroků technologické linky vstupu dat.

Diskuse o metodách vstupu digitálních geografických dat má velký význam vzhledem k ceně a životnosti geografických dat - geografická data jsou řádově nejcennější součástí GIS.

3. Kvalita vstupních dat

Kvalitu vstupních dat lze posuzovat z několika hledisek.

3.1 Hledisko mikroúrovně

- přesnost určení souřadnic (očekávaná odchylka souřadnic od skutečné polohy)
- přesnost určení atributů (hodnota atributu odpovídá číselníku nebo se nachází v zadaném intervalu)
- logická konzistence dat (jednoznačné relace mezi daty)
- rozlišení (nejmenší reprezentovatelná jednotka, nejmenší objekt zobrazený na mapě)

3.2 Hledisko makroúrovně

- úplnost (úplnost pokrytí území, úplnost klasifikace - jak kvalitně zvolená klasifikace reprezentuje data, úplnost verifikace dat)
- aktuálnost dat (různá frekvence aktualizace podle druhu geografických dat)
- rodokmen dat (historie dat - zdroje a postup zpracování dat)

3.3 Hledisko uživatelské úrovně

- přístupnost dat (privátní data, veřejné databáze, ochrana dat)
- přímá a nepřímá cena dat (nepřímá cena dat způsobená zpracováním nekvalitních dat může být vyšší než cena kvalitních dat)

4. Metody vstupu dat do GIS

K základním metodám vstupu dat patří:

- vstup z klávesnice
- měření a geodetické výpočty
- manuální digitalizace
- skenování
- konstrukce nad digitálním podkladem
- vstup existujících digitálních souborů¹

4.1 Vstup dat z klávesnice

Alfanumerický vstup dat z klávesnice počítače je stejně jako v jiných informačních systémech nejčastějším vstupem dat. Z klávesnice vstupují především předmětná data (fyzikální nebo klasifikační složka geografických dat), ale z klávesnice mohou vstupovat i prostorová data (například kódováním podle kódového systému při geodetickém měření dat).

Důležitou oblastí vstupu dat z klávesnice je vstup jednoznačných klíčů při digitalizaci nebo jiné metodě pořizování prostorových dat. Identifikace prostorových dat jednoznačným klíčem přímo při pořizování dat (samozřejmě jen tehdy, když takový klíč existuje) je jedním ze základních metodických principů pořizování geografických dat.

Výhody:

- možnost vstupu atributů mimo (drahé) grafické prostředí

Nevýhody:

- nelze efektivně pořizovat prostorovou složku

¹ Pokud jde o vstup velkého množství stávajících dat, kterému předchází jejich předzpracování, mluvíme o migraci dat.

4.2 Měření a geodetické výpočty

- automatizace výpočtu souřadnic ve zvoleném kartografickém systému (S-JTSK) - "totální stanice", výstup přes sériový port do textového souboru
- moderní metody pořizování geodetických dat (lokalizace pomocí družice - GPS, fotogrammetrické vyhodnocení snímků dálkového průzkumu Země - DPZ)
- třídy přesnosti 1 - 3 podle střední chyby měřených dat
- základní geodetické konstrukční úlohy
- metoda "hraničního katastru" - zaměření identických bodů a korektní transformace do kartografického souřadného systému
- geodetické zvyklosti při pořizování a údržbě prostorových dat
 1. pořízení digitální mapy - ne prostorových dat (nevhodný datový model na výstupu, odlišení shora neviditelných objektů),
 2. nekorektní aktualizace dat metodami "klasické" geodézie (aktualizace do nepřesného souřadného systému, aktualizace metodou průsvitek)
 3. subjektivní doplňování dat (pravé úhly na rozích budov, kruhové oblouky)

Výhody
přesnost určení souřadnic

Nevýhody
drahá data
geodetické zvyklosti

4.3 Manuální digitalizace

Manuální digitalizace je založena na využití digitizéru - to je digitalizační stůl, který se skládá ze dvou základních zařízení: z tabletu a kurzoru (pera, pointing device). Kurzor je opatřen tlačítky, která umožňují vstup jednoduchých popisných dat - například vstup typu prostorových dat při jejich kategorizaci.

Mezi základní vlastnosti digitizéru patří kromě rozměru aktivní plochy (A1, A0) rozlišovací schopnost (například 1200 ppi²) a především přesnost digitizéru (přesnost opakovaného snímání souřadnic) s typickými hodnotami 0.01", 0.005", 0.001". Přesnost zásadně určuje cenu digitizéru, a proto je výhodné koupit tak přesný digitizér, který stačí na pořízení podkladu v přesnosti dané kvalitou podkladu (síla čáry, kvalita kresby, kvalita lícovacích bodů).

Digitizér může pracovat v několika režimech:

- snímání významných bodů
- sledování čáry a záznam bodů v průsečíku čáry a myšlené mřížky
- sledování čáry a záznam bodů v časových intervalech

Nejčastější je práce v prvním z uvedených režimů.

Kvalitu a cenu dat pořízených digitizérem podstatně ovlivňuje podpůrný software, který musí být připraven tak, aby efektivně podporoval konkrétní způsob pořízení dat. Podpůrný software je závislý na datovém modelu digitalizovaných dat. V této souvislosti je nutné si uvědomit, že cena pořízených dat se skládá z ceny hardware (digitizéru), software, podkladů a práce při pořízení dat, při pořizování většího množství dat se tedy vyplatí nešetřit na podpůrném software.

Důležitým prvkem této i jiných metod pořízení geografických dat je operátor digitalizace, který musí být pečlivý, musí ovládat specifické dovednosti týkající se geografických dat a musí mít jisté (obvykle několikaměsíční) zkušenosti s pořizováním geografických dat.

Výhody

- při snímání se nepohybuje médium
- možnost identifikace resp. klasifikace prvků prostorových dat přímo při digitalizaci
- výběr snímaných dat operátorem
- výsledkem je relativně malé množství dat
- efektivní snímání velkých předloh
- snímání v konstantním měřítku (nemusí se vynucovat zvětšování obrazu jako při obrazovkové digitalizaci)

Nevýhody

- méně přesná data než měření
- nároky na zkušenosti a přesnou práci operátora

² Points per inch, bodů na palec.

4.4 Skenování

Skenováním se pořizují rastrová prostorová data, která lze případně dále vektorizovat. Pro skenování se využívají skenery dvou základních typů:

- válcové skenery, ve kterých se pohybuje skenované médium i skenovací hlava
- stolové skenery, ve kterých se pohybuje pouze skenovací hlava

Protože při pohybu skenovaného média (papírová mapa, fólie) dochází k náhodným změnám rovnoměrného pohybu, výsledkem jsou nelineárně zkreslená data (zkreslení nelze napravit lineární transformací). Skenování na válcových skenerech nelze tedy pro vstup přesných dat do GIS doporučit.

Skenování se provádí v rozlišovací schopnosti 300, 400, 600, 800, 1200 dpi³. Při větší rozlišovací schopnosti dochází k rapidnímu vzrůstu rozsahu výsledných dat.

Rastrová data získaná skenováním lze podle potřeb dále vektorizovat (konverze raster → vektor):

- manuálně - například tzv. obrazovkovou (screen-based) digitalizací (tento způsob pořízení dat je popsán v následujícím odstavci)
- automaticky - některé z metod jsou uvedeny v kapitole 4 (zužování čar a vektorizace skeletu, trasování kontur a vektorizace kontury)

Výhody skenovaných dat

- levná data
- možnost přímého užití skenovaných dat v rastrovém datovém modelu

Nevýhody skenovaných dat

- velké množství dat
- nároky na kvalitní software
- nároky na přesnou práci operátora a přesný postup zpracování
- velké množství zbytečných dat, pokud není zpracováván celý podklad
- při následné manuální vektorizaci nevýhody konstrukce dat popsané v následujícím odstavci

Výhody pořízení dat stolovými skenery

- při snímání se nepohybuje médium
- snímání i zničených originálů
- při správném zpracování přesnější data než manuální digitalizace

Nevýhody pořízení dat stolovými skenery

- dražší než data pořízená na válcových skenerech

Výhody pořízení dat válcovými skenery

- levná data (levnější než stolové skenery)

Nevýhody pořízení dat válcovými skenery

- nelineární zkreslení (pohybuje se snímáný podklad)

4.5 Konstrukce nad digitálním podkladem

Konstrukce dat nad digitálním (rastrovým nebo vektorovým podkladem) je nejčastější způsob vstupu dat. Používá se při aktualizaci dat například v těchto krocích:

- vstup měřených souřadnic a vytvoření digitálního podkladu ve formě bodů
- konstrukce nového stavu dat nad měřenými body
- výběr starého stavu dat (pokud starý stav existuje)
- transformace nového stavu do souřadného systému starého stavu
- ošetření kontextu nového stavu dat ve starém stavu

Kromě měřených bodů lze jako digitální podklad použít:

- skenovaná rastrová data - například při vektorizaci rastrových dat tzv. obrazovkovou (screen-based) digitalizací: rastrová data jsou zobrazena na obrazovce počítače a operátor s větší nebo menší pomocí programového vybavení konstruuje vektorové prvky podle tvaru rastrových dat
- vektorová data - například při doplnění další vrstvy dat do existujících dat nebo při konstrukci generalizovaných dat (manuální konstrukce dat menšího měřítka z dat velkého měřítka)

Výhody pořízení dat konstrukcí

Nevýhody pořízení dat konstrukcí

³ Dots per inch, pixelů (bodů) na palec.

- při kvalitní práci operátora jsou data stejně přesná jako data podkladu
- možnost identifikace resp. klasifikace prvků prostorových dat přímo při digitalizaci
- výběr konstruovaných dat operátorem
- výsledkem je relativně malé množství dat

- nároky na přesnou práci operátora
- nutnost vynutit zvětšení obrazu kvůli dosažení přesnosti⁴

4.6 Vstup existujících digitálních souborů

Existuje velké množství digitálních prostorových dat, která lze využít v různých GIS, například:

- data státního informačního systému (SIS): střední měřítka - ZABAGED (Základní báze geografických dat), velká měřítka - digitální katastrální mapa (DKM), hydroekologický informační systém (HEIS), informační systémy okresních úřadů
- městské informační systémy: digitální technické mapy (DTM) velkého měřítka, územní plány, cenové mapy, dopravní geografické systémy
- informační systémy správců inženýrských sítí
- snímky dálkového průzkumu Země (DPZ) - skenované letecké snímky, družicové snímky
- speciálním případem je vstup (migrace) dat ve stávajícím fyzickém datovém modelu do nově navrženého datového modelu

Při převzetí existujících dat se objevují tyto problémy:

- obsah dat - je obtížné kontrolovat, zda data mají požadovaný obsah
- kvalita dat - často je nemožné zjistit kvalitu dat (přesnost, původ, apod.)
- formát dat - při konverzi dat mezi různými formáty může dojít ke ztrátě nebo zkreslení dat

Z uvedených problémů vyplývá nutnost standardizace při předávání dat.

Výhody vstupu existujících dat

- data jsou již v digitální formě

Nevýhody vstupu existujících dat

- ztráta informace při konverzi dat
- pokud jsou data bez rodokmenu je nejistý stav dat

5. Další zpracování pořízených dat

Před vstupem do GIS je vhodné pořízená data dále zpracovat s cílem data zkontrolovat a zlepšit kvalitu pořízených dat. Další zpracování se skládá z těchto kroků:

- vizuální kontrola pořízených dat (vykreslení pořízených dat na plotru a kontrola shody kresby s pořizovaným podkladem)
- kontrola dodržení datového modelu prostorových dat (kontrola obsahu a topologických vlastností dat)
- řešení styků mezi mapovými listy (mapový podklad je často rozdělen do mapových listů - při údržbě takového podkladu vzniká mnoho chyb především na styku mapových listů, které jsou při pořízení dat přeneseny do výsledku; tyto chyby je vhodné odstranit dřív, než data vstoupí do GIS)
- kontrola připojených identifikací a atributů (pokud jsou prostorová data pořizována s jednoznačnými klíči, případně s dalšími atributy, je vhodné provádět jejich kontrolu v rámci pořizování dat před vstupem do GIS)

6. Některé zásady při pořizování prostorových dat

Při pořizování prostorových dat je výhodné dbát několika zásad:

- volit co nejlepší podklady (je lepší pořizovat data z kvalitnějších podkladů, i když jsou výsledná data někdy dražší)
- provádět postupnou transformaci souřadného systému:
 - digitalizovaná data -> souřadný systém mapy
 - souřadný systém mapy -> identické body v terénu
 (rozložení transformace do dvou kroků umožní kontrolovat přesnost digitalizace a tím kvalitu podkladu a práce operátora)

⁴ Pokud není vynuceno zvětšení obrazu při konstrukci dat - zejména při obrazovkové digitalizaci (například důsledným prováděním kontrolních kreseb a jejich vizuální kontrolou), má operátor tendenci zjednodušovat si práci a konstruovat (vektorizovat) obraz zobrazený na obrazovce v příliš malém měřítku.

- uchovávat parametry transformace (při pořizování dat je nutné provádět transformaci dat ze systému podkladu do kartografického systému, který je na podkladu zobrazen; parametry transformace je dobré uchovávat, aby bylo možné v budoucnu zjistit kvalitu pořizovaných dat)
- minimalizovat zásahy do podkladů před digitalizací (každou operací se do dat vnáší chyby, lepší je pořídít data z neupraveného podkladu a případné korekce provádět až do pořizovaných dat)
- digitalizovat v přesnosti síly čáry (nemá smysl pořizovat data přesněji - a tím i dražší - než je síla čáry, kterou jsou pořizovaná data v podkladu zakreslena)
- navrhovat datový model tak, aby sloužil k verifikaci pořizovaných dat (často je vhodné upravit datový model prostorových dat, aby umožnil verifikovat pořizovaná data a kontrolovat jednotlivé kroky technologické linky pořízení dat)
- formalizovat předávání výsledků digitalizace a jejich kontrolu (formální předávání dat a výsledků jejich kontrol formou předávacích protokolů zdůrazňuje osobní odpovědnost operátora za pořizovaná data a v důsledku vede k vyšší kvalitě dat)
- při digitalizaci rozsáhlých souborů postupovat metodou pilotního projektu (pilotní projekt a jeho vyhodnocení slouží ke korekci metodiky pořizování dat, což vede k levnějším a kvalitnějším datům)

7. Obsah metodiky pořizování prostorových dat

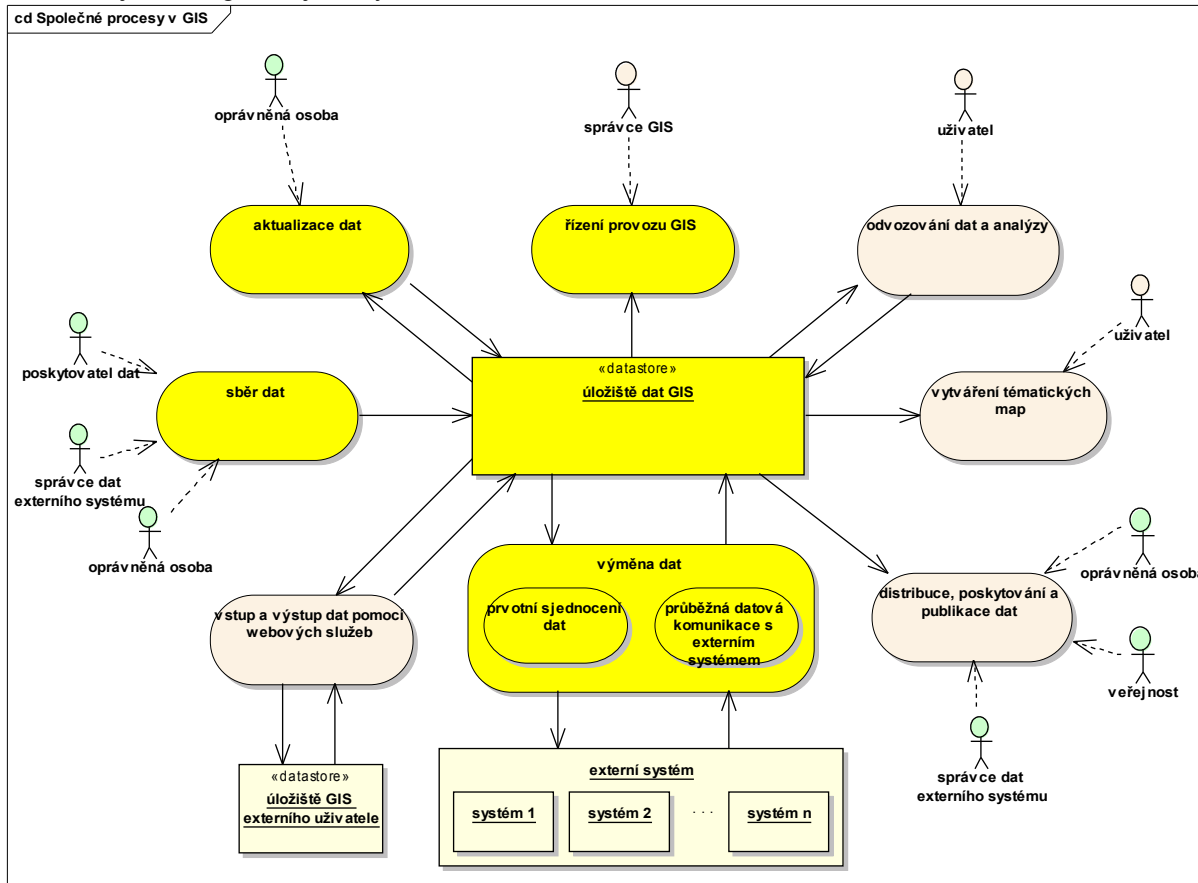
Při pořizování prostorových dat velkého rozsahu je vhodné vytvořit metodiku pořizování speciálně přizpůsobenou konkrétní úloze. Metodika je vytvořena před realizací pilotního projektu a v jeho průběhu - případně v rámci vyhodnocení pilotního projektu - je aktualizována. Metodika pořizování prostorových dat by měla obsahovat:

- popis metody pořizování prostorových dat
- datový model (přizpůsobený metodě pořizování dat)
- způsob přípravy podkladů
- postup digitalizace (pořizování)
- zpracování digitalizovaných (pořizovaných) dat a jejich kontroly
- vstup pořizovaných dat do GIS
- organizační zabezpečení pořízení dat
- hardwarové a softwarové zabezpečení pořízení dat
- počáteční odhady pracnosti a dalších nákladů
- návrh dalšího postupu (vymezení území pro pilotní projekt, nastavení parametrů metodiky pro pilotní projekt)

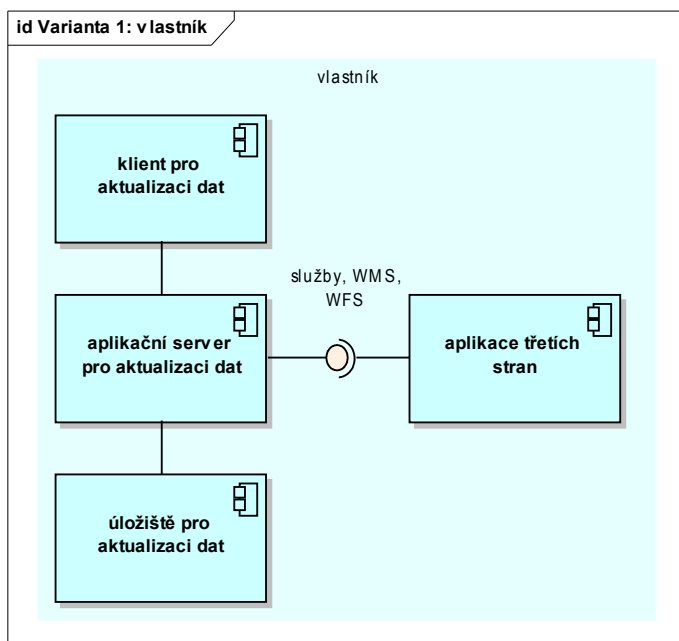
8. Strategie při vstupu a aktualizaci dat

Procesy GIS při vstupu a aktualizaci dat

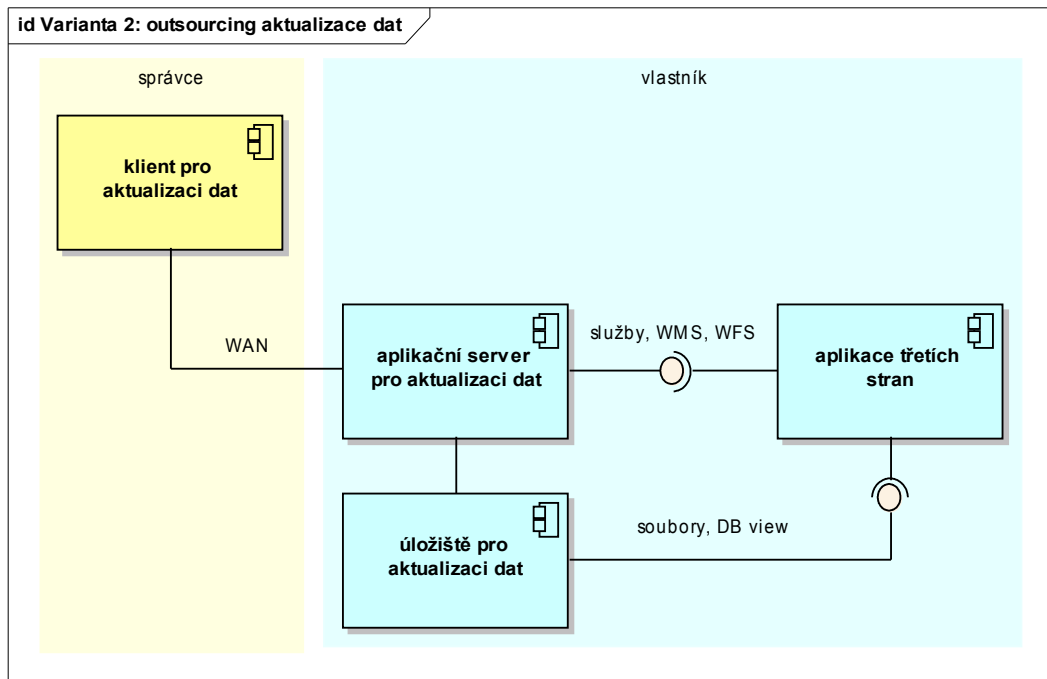
V následujícím diagramu jsou vyznačené žlutě.



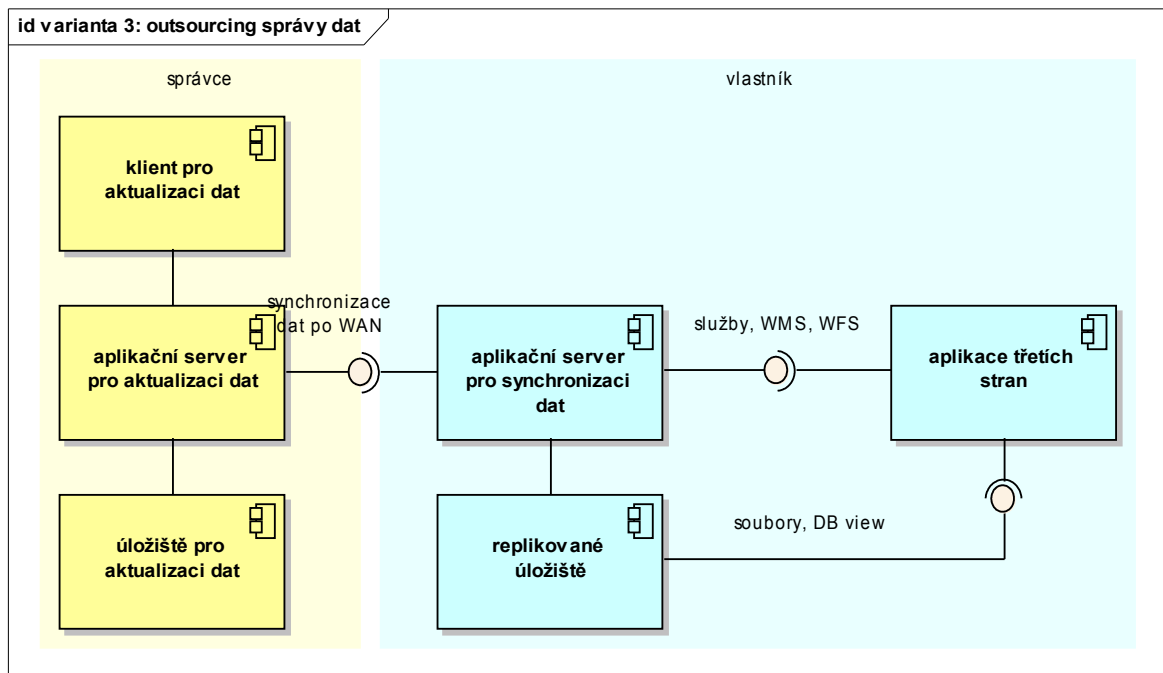
Varianta 1: zpracování dat vlastníkem a provozovatelem GIS



Varianta 2: outsourcing aktualizace dat



Varianta 3: outsourcing správy dat



9. Zkušební otázky

1. Popiš kritéria kvality prostorových dat?
2. Lze dosáhnout skenováním větší přesnosti než manuální digitalizací? Pokud ano, za jakých podmínek?
3. Popiš výhody a nevýhody jednotlivých metod pořízení dat.
4. Jaké důvody vedou k rozhodnutí mezi pořízením prostorových dat skenováním a manuální digitalizací?

Připomínky a dotazy k obsahu lekce posílej, prosím, na adresu:
Rudolf Richter, richter@fi.muni.cz