

PODPORA VÝMĚNNÉHO FORMÁTU RÚIAN V OPEN SOURCE GIS APLIKACÍCH

Martin LANDA

Katedra geomatiky, Fakulta stavební, ČVUT v Praze, Thákurova 7, 166 29, Praha, Česká republika
martin.landa@fsv.cvut.cz

Abstrakt

Cílem tohoto příspěvku je shrnout současný stav podpory Výměnného formátu (VFR) Registru území identifikace, adres a nemovitostí (RÚIAN) v open source aplikacích jako je QGIS, GRASS GIS či MapServer. Klíčovým stavebním prvkem je v tomto ohledu implementace podpory VFR v knihovně GDAL, která tento formát podporuje od verze 1.11. Díky tomu většina open source GIS aplikací formát VFR podporuje nativně včetně jeho konverze do jiných GIS formátů jako je Esri Shapefile, souborová geodatabáze Esri, OGC GeoPackage či geodatabáze PostGIS. Na základě toho lze tato data dále jednoduše poskytovat přes mapový server jako např. MapServer ve formě webovým služeb jako WMS či WFS apod.

Abstract

The aim of this paper is to summarize the current support of Czech Exchange Format for RUIAN (VFR) in Open Source software like QGIS, GRASS GIS or MapServer. The key point is an implementation of VFR support in GDAL library which supports this data format since version 1.11. Thanks to GDAL implementation the most of Open Source applications enable the users to read VFR data including conversion to other GIS formats like Esri Shapefile, Esri File Geodatabase, OGC Geopackage, or PostGIS geodatabase.

Klíčová slova: RÚIAN; VFR; open source; GIS

Keywords: RUIAN; VFR; Open Source; GIS

1 REGISTR ÚZEMNÍ IDENTIFIKACE, ADRES A NEMOVITOSTÍ

Registr územní identifikace, adres a nemovitostí (RÚIAN) je jedna ze čtyř komponent systému základních registrů veřejné správy ČR, který byl spuštěn v polovině roku 2012. Podstatnou vlastností tohoto registru je jeho *otevřenost*. Data z registru jsou poskytována zdarma bez nutnosti registrace a to přes internetovou aplikaci *Veřejného dálkového přístupu (VDP)* široké veřejnosti samozřejmě i včetně komerční sféry [1]. Data poskytovaná VDP lze označit jako *data otevřená*. V českém kontextu je projekt z pohledu otevírání dat jedinečný jak svým rozsahem, tak především kvalitou vydávaných dat [2]. Aplikace VDP (<http://vdp.cuzk.cz>) je provozována Českým úřadem zeměměřičským a katastrálním (ČÚZK).

VDP poskytuje data registru RÚIAN jako předgenerované soubory ve vlastním formátu označovaném jako *Výměnný formát RÚIAN (VFR)*, viz kapitola 1.2. Data jsou členěna podle územních prvků a dále poskytována ve dvou základních formách jako *data stavová*, tj. celá datová sada platná k poslednímu dni daného měsíce anebo jako *data změnová* ke každému kalendářnímu dni. Tento přístup publikování dat je pro cílové uživatele velmi výhodný, neboť jim umožňuje jednoduše udržovat začleňováním změnových souborů datovou sadu aktuální. Díky kvalitní dokumentaci postupně vznikla kolem těchto dat celá řada nástrojů a služeb ať už na proprietární či otevřené bázi. Jako příklad proprietárního nástroje můžeme uvést *VFR Import* od společnosti ARCDATA [3], který je poskytován ve třech verzích od základní bezplatné instalace až po placené verze umožňující automatické stahování stavových dat či začleňování denních změnových vět do geodatabáze. Tento příspěvek je zaměřen na druhou skupinu nástrojů patřících do sféry volně dostupného open source software.

1.2 Výměnný formát RÚIAN

Z technického hlediska je *Výměnný formát RÚIAN (VFR)* odvozen z technické normy ISO 19136 (OGC Geography Markup Language; GML [4]). Formát VFR nicméně není technickou normou, pouze z ní vychází.

K formátu existuje kompletní dokumentace, jeho licence není restriktivní, což z něj dělá formát *otevřený*, i když stále *proprietární*.

Formát VFR je na straně aplikací poměrně snadno zpracovatelný, neboť již většina moderních geografických informačních systémů (GIS) formát OGC GML, ze kterého VFR vychází, nativně podporuje. Výsledkem je, v porovnání např. s Výměnným formátem katastru nemovitostí (VFK) [5], jeho relativně rychlé začlenění do nejrůznějších GIS.

2 PODPORA FORMÁTU VFR V OPEN SOURCE GIS

Dle dostupných informací je historicky prvním (červenec 2012) open source nástrojem podporujícím datový formát VFR skupina konverzních skriptů *Ruian2gis*, které vznikly pro účely komunity okolo projektu OpenStreetMap [6]. Jedná se o jednoduché XSL transformační skripty umožňující konverzi dat ve formátu VFR do GML, který lze dále zpracovat v běžných GIS nástrojích. Vzhledem k účelovosti projektu se tyto nástroje nerozšířily mimo OpenStreetMap komunitu.

Text příspěvku dále odkazuje na klíčovou implementaci nativní podpory formátu VFR přímo v knihovně GDAL, která otevírá přístup k datům poskytovaných v tomto formátu široké komunitě uživatelů nejrůznějších GIS nástrojů od open source jako je QGIS, GRASS GIS či MapServer, ale i proprietárních jako je Esri ArcGIS právě v kombinaci s výše uvedenou knihovnou.

2.1 GDAL

Podpora formátu VFR byla začleněna do verze knihovny GDAL 1.11 (duben 2014) formou rozšíření podpory formátu GML [7]. V souvislosti s tím knihovna GDAL ve stejné verzi přináší pro formát VFR klíčovou podporu pro prvky s vícenásobnou geometrií [8], neboť geometrická složka popisu geoprvků je v tomto formátu běžně vyjádřena jak definičním bodem, tak hranicí originální či generalizovanou. Podrobné informace jsou dostupné na portálu FreeGIS [9].

Knihovna načítá ze souboru VFR (soubor s příponou *.xml.gz*) jednotlivé geoprvky v logických blocích (GMLFeatureClass), tzv. vrstvách (*layers*). Např. soubor VFR typu ST_UKSH (tj. územní prvky stát až základní sídelní jednotky, úplná kopie, kompletní datová sada, originální hranice) dostupný z VDP¹ je prezentován nástrojem knihovny GDAL *ogrinfo* poskytujícím základní informace o datech ve formě 14 vrstev od „Staty“ až po „Zsj“:

```
$ ogrinfo /vsicurl/http://vdp.cuzk.cz/vymenny_format/soucasna/20141130_ST_UKSH.xml.gz
1: Staty (Point, Multi Polygon, Multi Polygon)
2: RegionySoudrznosti (Point, Multi Polygon, Multi Polygon)
...
14: Zsj (Multi Point, Multi Polygon)
```

U každé vrstvy je uveden seznam navazujících typů geometrií, např. u vrstvy „Staty“ se jedná o „MultiPoint“ (definiční bod), „MultiPolygon“ (originální hranice) a „MultiPolygon“ (generalizovaná hranice). Obdobně jsou načítány soubory pro území prvků obcí. Příklad pro obec „Litoměřice“ (kód obce 564567):

```
$ ogrinfo
/vsicurl/http://vdp.cuzk.cz/vymenny_format/soucasna/20141130_OB_564567_UKSH.xml.gz
1: Obce (Multi Point, Multi Polygon, Multi Polygon)
2: SpravniObvody (Point, Multi Polygon)
...
11: AdresniMista (Point, Point, Point)
```

1 V textu jsou použita stavová data ze dne 30.11.2014, http://vdp.cuzk.cz/vymenny_format/soucasna/20141130_ST_UKSH.xml.gz. Knihovna GDAL umožňuje načítání online dat přes rozhraní */vsicurl/*.

Knihovna GDAL kromě stavových dat podporuje i data *změnová*, která jsou poskytována přes VDP na denní bázi. Příklad načtení změnových dat typu UKSH ze dne 18.12.2014:

```
$ ogrinfo /vsicurl/http://vdp.cuzk.cz/vymenny_format/soucasna/20141218_ST_ZKSH.xml.gz
1: Staty (Point, Multi Polygon, Multi Polygon)
2: RegionySoudrznosti (Point, Multi Polygon, Multi Polygon)
...
18: AdresniMista (Point, Point, Point)
19: ZaniklePrvky
```

Konverze dat VFR do jiných GIS formátů je pro knihovnu GDAL, resp. její nástroj *ogr2ogr*, již běžnou rutinou. Následuje příklad konverze definičních čar ulic pro obec „Litoměřice“ do formátu Esri Shapefile:

```
$ ogr2ogr -f 'ESRI Shapefile' ulice.shp
/vsicurl/http://vdp.cuzk.cz/vymenny_format/soucasna/20141130_OB_564567_UKSH.xml.gz Ulice
```

Výsledkem bude soubor *ulice.shp* s geometrií definičních čar ulic a příslušnými popisnými informacemi.

Od připravované verze 1.11.2 bude GDAL podporovat i speciální formát volebních okrsků, tj. typ souboru ST_UVOH.

Dále GDAL ve své vývojové verzi 2.0.0dev nově podporuje i *křivky* [10]. To je klíčové především pro vrstvu „Parcely“ a „Ulice“, které mohou obsahovat geometrickou složku popisu geoprveků tvořenou kružnicemi či kruhovými oblouky, viz obr. 1.



Obr. 1. Příklad parcel jejichž hranice je tvořena kruhovými oblouky. Vlevo je reprezentace hranic ve verzi GDAL 1.11, která křivky nepodporuje. Vpravo je reprezentace hranic ve verzi knihovny GDAL 2.0.0dev, tj. včetně kruhových oblouků.

2.1.1 Konverzní nástroje VFR

Nad knihovnou GDAL vznikla skupina konzolových nástrojů, které umožňují dávkové stahování dat z VDP a dále korektní zpracování změnových vět. Nástroje byly implementovány v programovacím jazyce Python a jsou podobně jako knihovna GDAL šířeny pod licencí MIT [11]. Balíček obsahuje tři příkazy:

- *vfr2ogr* – obecný nástroj pro konverzi dat z formátu VFR do libovolného GIS formátu podporovaného knihovnou GDAL;

- *vfr2ogr* – specializovaný nástroj pro konverzi dat z formátu VFR do open source geodatabáze PostGIS;
- *vfr2oci* – specializovaný nástroj pro konverzi dat z formátu VFR do proprietární geodatabáze Oracle Spatial.

Tyto nástroje umožňují z VDP soubory ve formátu VFR stahovat automaticky. V níže uvedeném příkladu vytvoříme soubor ve formátu Esri Shapefile (*--format*), který bude obsahovat polygony (*--geom OriginalniHranice*) okresů v ČR (*--layer Okresy*). Soubor ve formátu Esri Shapefile se uloží do aktuálního adresáře jako *okresy.shp*. Dle typu souboru (*--type ST_UKSH*) se odvodí URL pro stažení dat přímo z VDP. Pokud není definováno specifické datum (*--date*), tak skript najde poslední dostupná stavová data, v našem případě z 30.11.2014 (příklad se vztahuje k 19.12.2014):

```
$ vfr2ogr --type ST_UKSH --format ESRI_Shapefile --layer Okresy --geom OriginalniHranice
Processing 20141130_ST_UKSH.xml.gz (1 out of 1)...
Processing layer Okresy                ...                77 features added
```

Kromě jednotlivých souborů umožňují skripty stahovat dávkově i více souborů najednou uložených jako seznam v textovém formátu, např. na základě souboru *seznam_linku.txt*:

```
OB_554979_UKSH
OB_531367_UKSH
OB_535826_UKSH
```

Příkaz uvedený níže stáhne z VDP následující soubory:

- 20141130_OB_554979_UKSH.xml.gz
- 20141130_OB_531367_UKSH.xml.gz
- 20141130_OB_535826_UKSH.xml.gz

Na základě nich vytvoří v aktuálním adresáři soubor *obce.shp*, který bude obsahovat polygony (*--geom OriginalniHranice*) výše uvedených obcí (*--file*):

```
$ vfr2ogr --file seznam_linku.txt --format ESRI_Shapefile --layer Obce --geom
OriginalniHranice
3 VFR files will be processed...
Processing 20141130_OB_554979_UKSH.xml.gz (1 out of 3)...
Processing layer Obce                ...                1 features added
Processing 20141130_OB_531367_UKSH.xml.gz (2 out of 3)...
Processing layer Obce                ...                1 features added
Processing 20141130_OB_535826_UKSH.xml.gz (3 out of 3)...
Processing layer Obce                ...                1 features added
Summary
Layer                Obce                ...                3 features
```

Import dat do geodatabáze PostGIS. V následujícím příkladě si ukážeme komplexní přístup k importu dat z VDP do open source geodabáze PostGIS včetně zpracování změnových souborů. Nejprve do prázdné geodatabáze PostGIS (*--dbname ruian*) naimportujeme kompletní stavová data typu UKSH za celou ČR (balíček konverzních skriptů obsahuje předgenerovaný seznam linků v souboru *test_suite/db_uksh.txt*):

```
$ vfr2ogr --file test_suite/db_uksh.txt --dbname ruian
6254 VFR files will be processed...
Processing 20141130_ST_UKSH.xml.gz (1 out of 6254)...
Processing layer Staty                ...                1 features added
```

```

...
Processing layer Zsj ... 0 features added
Processing 20141130_OB_554979_UKSH.xml.gz (2 out of 6254)...
Processing layer Obce ... 1 features added
...
Processing layer AdresniMista ... 544 features added (79 without geometry)
...
Processing 20141130_OB_534510_UKSH.xml.gz (6254 out of 6254)...
Processing layer Obce ... 1 features added
...
Processing layer AdresniMista ... 398 features added (12 without geometry)
Summary
Layer Staty ... 1 features
...
Layer AdresniMista ... 2897100 features

```

Stavová data se stáhnou k poslednímu dni předešlého měsíce, v našem případě tedy k 30.11.2014. Dalším příkazem stáhneme dávkově všechny změnové soubory (pro typ UKSH mají změnové soubory prefix ST_ZKSH) od 1.12.2014 do 19.12.2014 (`--date :20141219`)² a data v geodatabázi k tomuto dni aktualizujeme:

```

$ vfr2pg --type ST_ZKSH --date :20141219 --dbname ruian
Processing 20141201_ST_ZKSH.xml.gz (1 out of 19)...
Processing layer Staty ... 0 features ( 0 added, 0 updated, 0 deleted)
...
Processing layer AdresniMista ... 158 features (131 added, 27 updated, 25 deleted)
...
Processing 20141219_ST_ZKSH.xml.gz (19 out of 19)...
Processing layer Staty ... 1 features ( 0 added, 1 updated, 0 deleted)
...
Processing layer AdresniMista ... 143 features (109 added, 34 updated, 33 deleted)
Summary
Layer Staty ... 1 features
...
Layer AdresniMista ... 2898388 features

```

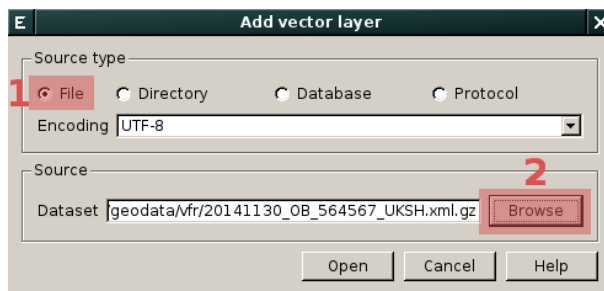
2.2 QGIS

Díky podpoře formátu VFR v knihovně GDAL umožňuje QGIS načítat tato data stejným způsobem jako např. data ve formátu Esri Shapefile či jiná vektorová data (viz obr. 2).

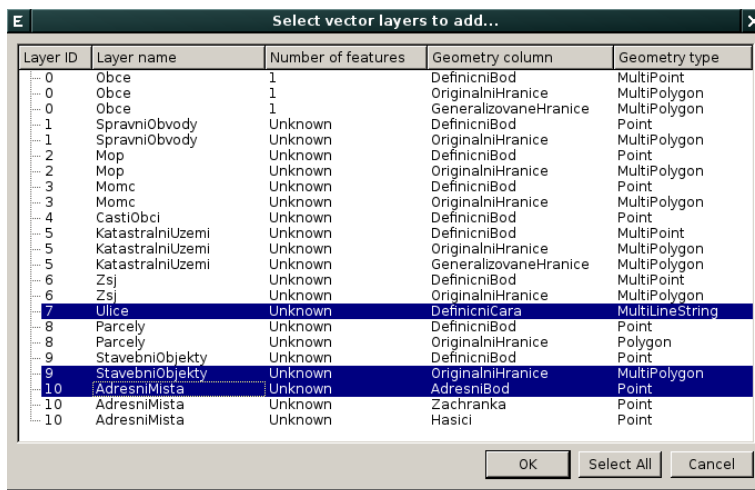


Obr. 2. Přidání vektorové vrstvy z nástrojové lišty aplikace QGIS

² Řetězec ':20141219' je zkratkou pro '20141201:20141219', tj. rozsah od 1.12.2014 až do 19.12.2014 včetně.

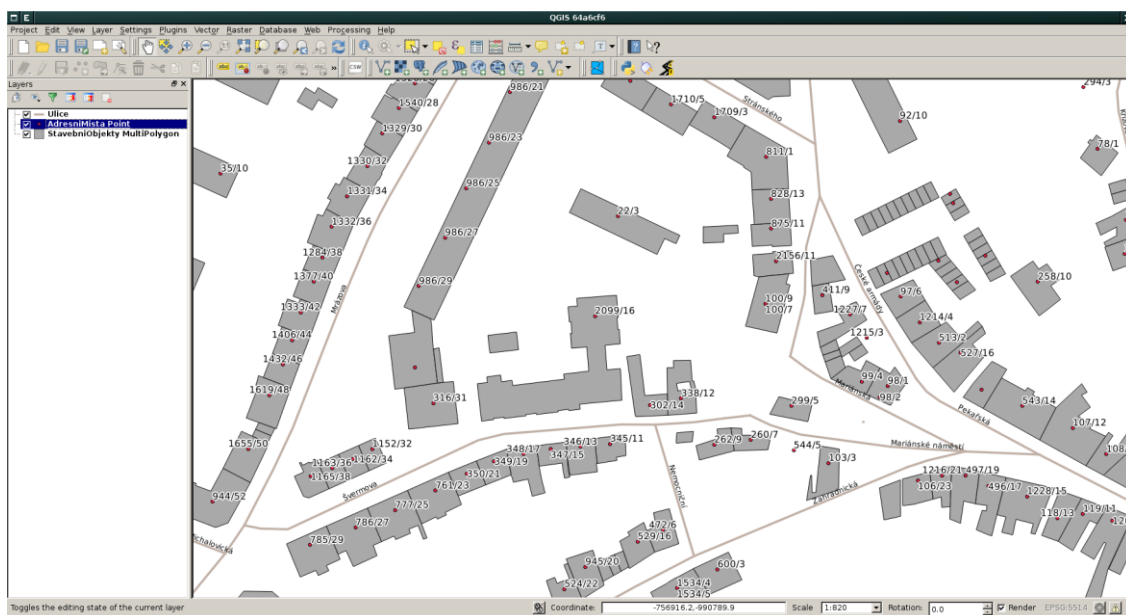


Obr. 3. V následujícím dialogu vybereme načtení dat ze souboru (1) a zvolíme vstupní soubor ve formátu VFR (přípona .xml.gz), který chceme do QGISu načíst



Obr. 4. V posledním dialogu vybereme vrstvy, které chceme zobrazit. V našem případě zvolíme definiční čáry ulic, originální hranice stavebních objektů a adresní body³

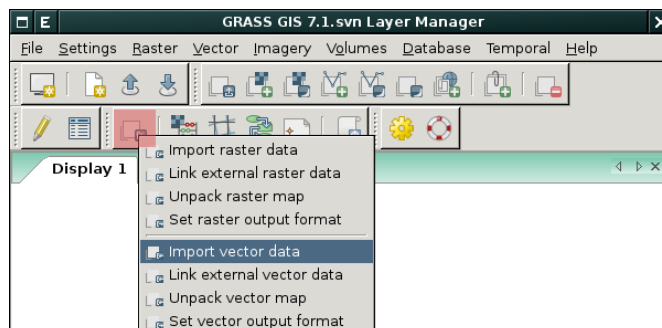
Příklad vizualizace dat je uveden na obr. 5.



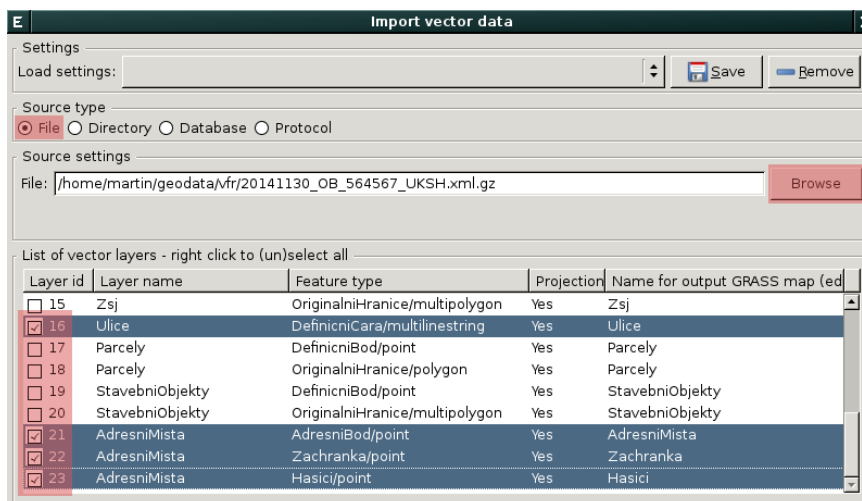
Obr. 5. Vizualizace ulic, stavebních objektů a adresním míst v aplikaci QGIS

³ Dialog pro výběr vrstev ve verzích QGIS 2.6 a nižších nabízí pouze první geometrii z dané vrstvy, např. pro vrstvu „Obce“ se zobrazí pouze geometrie „DefiniciBod“. V těchto verzích QGISu je nejprve nutné vytvořit pomocí konzolového nástroje *vfr2ogr* např. Shapefile s požadovanou geometrií a ten potom do QGISu načíst.

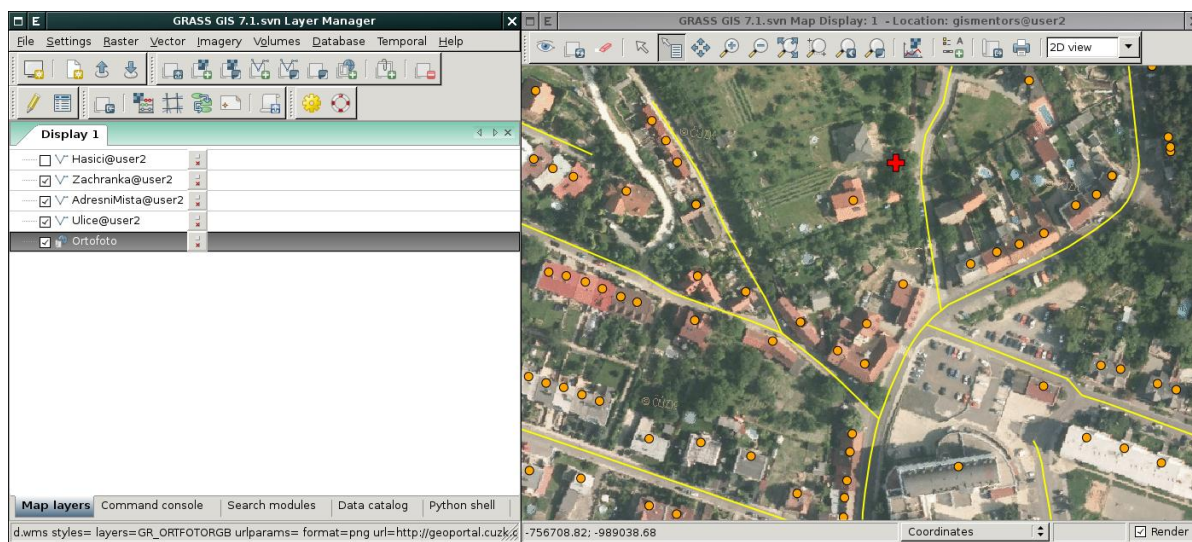
Do systému GRASS lze načíst data ve formátu VFR obdobně jako ostatní vektorová data a to buď tak, že data převedeme do nativního topologického formátu systému GRASS příkazem *v.in.ogr* anebo vytvoříme pomocí příkazu *v.external* odkaz na původní data ve formě jednoduchých geoprvků (*simple features*). Tyto nástroje jsou dostupné z menu grafického uživatelského prostředí systému GRASS, viz obr. 6.



Obr. 6. Import vektorových dat z nástrojové lišty grafického uživatelského rozhraní systému GRASS



Obr. 7. V následujícím dialogu vybereme vrstvy, které chceme naimportovat ⁴



Obr. 8. Příklad vizualizace ulic, adresních bodů a záchranky (červený kříž) na podkladě ortofota (ČÚZK WMS) v systému GRASS

4 Systém GRASS podporuje vícenásobnou geometrii (viz dialog na obr. 7) při importu vektorových dat pomocí knihovny GDAL od verze 7.0.0RC1.

2.4 MapServer

K datům ve formátu VFR lze přistupovat podobně jako k ostatním datům podporovaných knihovnou GDAL. V následující ukázce nastavíme mapový soubor MapServeru tak, aby nabízel vrstvu „Okresy“ (*NAME*) s geometrií originálních hranic, kódem a názvem okresů v ČR (*DATA*). Dále nastavíme typ prvků (*TYPE*), druh připojení (*CONNECTIONTYPE*) k vektorovým datům (*OGR*) a cestu k souboru ve formátu VFR na disku (*CONNECTION*).

```
LAYER
  NAME "Okresy"
  TYPE POLYGON
  CONNECTIONTYPE OGR
  CONNECTION "/ruian/20141130_ST_UKSH.xml.gz"
  DATA "SELECT Kod,Nazev,OriginalniHranice from Okresy"
  STATUS ON
END
```

3 ZÁVĚR

Implementace podpory formátu VFR v knihovně GDAL je, vzhledem k jejímu rozšíření ve většině open source a řadě proprietárních GIS aplikací, pro přístup k těmto datům bez závislosti na proprietárním řešení zcela klíčová. Podpora tohoto formátu v knihovně GDAL umožňuje snadný přístup k těmto datům z desktopových GIS programů jako je např. QGIS či GRASS GIS, publikaci dat na mapovém serveru jako např. MapServer a také poměrně jednoduché dávkové zpracování dat poskytovaných online z VDP včetně korektního zohlednění změnových souborů s cílem jejich importu například do geodatabáze PostGIS.

Z pohledu geometrické složky poskytovaných dat je zásadním rozšíření podpory pro křivky ve vývojové verzi knihovny GDAL 2.0.0dev [10]. Další informace o celém projektu včetně instalace vývojové verze knihovny GDAL pro operační systém MS Windows najdete na portálu FreeGIS [9].

Další vývoj bude směřován k vylepšení komfortu uživatelů při práci s daty poskytovanými přes VDP. Programy QGIS i GRASS ve svých připravovaných verzích (QGIS 2.8 a GRASS 7.0) již podporují vícenásobnou geometrii a umožňující tak současně vybrat jak např. definiční body, tak polygony geoprvků dat poskytovaných ve formátu VFR. Přínosné by bylo doplnit do QGIS nový zásuvný modul, který by umožňoval z grafického uživatelského rozhraní stahovat data přímo z VDP, začleňovat automaticky změny apod. Tyto nástroje jsou v současnosti již implementovány, jsou ale dostupné prozatím pouze ve formě konzolových programů spustitelných z příkazové řádky (viz kapitola 2.1.1).

Vývoj podpory formátu VFR v knihovně GDAL včetně podpůrných konverzních skriptů byl finálně podpořen Výzkumným ústavem geodetickým, topografickým a kartografickým (VÚGTK). Případné dotazy směřujte na autora tohoto příspěvku.

4 LITERATURA

- [1] ČÚZK. RUIAN, <http://www.cuzk.cz/ruian/RUIAN.aspx>, 18.12.2014
- [2] OpenGeoLabs. Otevírání geografických dat, <http://opengeolabs.cz/wordpress/wp-content/uploads/2014/12/otevirani-geograficky-dat-2.2.0.pdf>, 19.12.2014
- [3] ARCDATA Praha. VFR Import, <http://www.arcdata.cz/produkty-a-sluzby/software/arcdata-praha/vfr-import>, 19.12.2014
- [4] OGC. Geography Markup Language. <http://www.opengeospatial.org/standards/gml>, 19.12.2014
- [5] Landa M. OGR VFK Driver Implementation Issues. GIS Ostrava 2010. <http://geo.fsv.cvut.cz/~landa/publications/2010/gis-ostrava-2010/paper/landa-ogr-vfk.pdf>, 19.12.2014

- [6] Barta D.: Ruain2gis, <http://freegis.fsv.cvut.cz/gwiki/RUIAN / Ruian2gis>, 19.12.2014
- [7] GDAL. GML driver, http://www.gdal.org/drv_gml.html, 19.12.2014
- [8] GDAL. RFC 41 : Support for multiple geometry fields in OGR,
http://trac.osgeo.org/gdal/wiki/rfc41_multiple_geometry_fields, 19.12.2014
- [9] Landa M. Podpora VFR v knihovně GDAL, <http://freegis.fsv.cvut.cz/gwiki/RUIAN / GDAL>, 18.12.2014
- [10] GDAL. RFC 49: Curve geometries, http://trac.osgeo.org/gdal/wiki/rfc49_curve_geometries, 19.12.2014
- [11] Landa M. GDAL-VFR, <https://github.com/landam/gdal-vfr>, 19.12.2014