

Testování formátu GML

Jan Smékal

Katedra geoinformatiky, Přírodovědecká fakulta, Univerzita Palackého v Olomouci,
tř. Svobody 26, 771 46, Olomouc, Česká republika
smekalho@seznam.cz

Abstrakt. Hlavním cílem této bakalářské práce bylo na základě sady provedených testů popsat a zhodnotit možnosti exportu a importu dat ve formátu GML verze 2 v GIS aplikacích ArcGIS 9.1, GeoMedia 6.0 a Autodesk Civil 3D 2007. Důraz byl kladen zejména na zkoumání přístupu jednotlivých aplikací k této problematice a jejich předností a nedostatků ve využití tohoto formátu. Část práce byla věnována vypracování rešerší volně dostupných prohlížeček, které umožňují zobrazovat geodata uložená ve formátu GML.

Klíčová slova: GML, interoperabilita, testování, prohlížečky

Abstract. Main purpose of this bachelor thesis was describe and evaluate possibilities of export and import of data in GML version 2 format in GIS applications ArcGIS 9.1, Geomedia 6.0 and Autodesk Civil 3D 2007 on the base of set of experted tests. Emphasis was laied especially on examination of admission of particular applications to this issue and their advantages and absences in exploitation of this format. Part of this thesis was dedicated to elaboration of searches of freeware viewers, which enabled vizualization of geodata stored in GML format.

Keywords: GML, interoperability, trstiny, viewers

1 Úvod

V současnosti se lze setkat s mnoha datovými formáty, které jsou používány k uložení vektorových geodat, a které jsou vázány na konkrétní softwarové produkty. Pokud má ale dojít k použití geodat v jiné aplikaci, dochází často k problémům s převodem dat do formátu, který by tato aplikace podporovala. V zájmu geoinformatické komunity je tedy standardizovat pro uložení vektorových geodat takový formát, který by byl jednoduchý, nebyl vázaný na konkrétní GIS software a přitom byl co nejvíce aplikacemi podporován.

Posun v této problematice nastal s příchodem značkovacího jazyka XML (eXtensible Markup Language), jehož modifikací byl vytvořen značkovací jazyk GML (Geography Markup Language), který umožňuje vektorová geodata ukládat do jednoduché strukturované textové podoby. Tento otevřený datový formát vytvořený konsorciem OGC (Open Geospatial Consortium) je v současnosti hlavně používán pro přenos vektorových geodat dat mezi GIS aplikacemi nebo pro přenos dat z mapových serverů.

Otevřenost formátu má ale i svá úskalí. Volnost v implementaci tohoto standardního formátu do GIS aplikací vyúsťuje v problémy s kompatibilitou mezi geodaty, jež jsou v nich v tomto formátu vytvořeny a v různé možnosti jejich využití.

2 Testování formátu GML

Testování bylo prováděno tak, že byla testovací vrstva ve formátu SHP exportována do formátu GML a poté naimportována zpět do do systému z něžž byla exportována. Testování bylo podrobeno šest různých GML profilů, které lze vytvořit v systémech ArcGIS 9.1 (profily FME, ESRI, FIXED), Geomedia 6.0 (GEOMEDIA) a Autodesk Civil 3D 2007 (CREATE, A_FIXED). Všechny profily byly vytvářeny podle specifikace GML verze 2.x.

Testování bylo uskutečněno na sestavě s procesorem Intel Core 2 Duo 2,13GHz, 2x512MB DDR RAM (Dual Chanel), OS Windows XP Professional SP2.

2.1 Kompatibilita

Tento test měl za cíl ověřit, zda je software schopný načítat GML soubory vyexportované v testovaných GML profilech. Všechny testované programy načítali GML soubory v nativních profilech, což by měla být samozřejmost. Systém Geomedia toto ale až do ledna 2007 nedokázal. Teprve až poté na jeho domovských internetových stránkách objevily ke stažení opravené verze všech extenzí nutných pro práci s GML, které problém nekompatibility odstranily.

Program ArcGIS načítal všechny testované profily. Jediný problém nastal u profilů z programu Autodesk Civil. ArcGIS u GML souborů vytvořených v těchto profilech měl problémy s českou diakritikou. Některá písmena se v atributové tabulce zobrazovaly chybně.

Aplikace Geomedia nenačítala GML soubory v obou FIXED profilech, protože neměly přiřazený příslušný soubor XSD s aplikačním schématem, s čímž si neporadila. U profilu CREATE špatně zobrazovala písmena s českou diakritikou. U profilů FME a ESRI vyvstal problém se souřadnicovými systémy. Nedokázala importovat ani vizualizovat GML soubory, které jsou vytvořeny v souřadnicových systémech používaných v České republice tedy S-JTSK, S-42 a WGS84. Pokud ale nebyl souřadnicový systém definován, GML soubor byl importován korektně.

Program Autodesk Civil nedokázal importovat GML soubory vytvořené v programu Geomedia. V modech FME a ESRI nezobrazoval korektně českou diakritiku. U GML souborů vytvořených v profilu FIXED nedošlo k importu atributové části geodat, přičemž geometrickou složku zobrazoval korektně.

Tabulka 1. Kompatibilita GML profilů s testovanými programy.

Profil GML	Testovaný program		
	ArcGIS 9.1	Geomedia 6.0	Autodesk Civil 3D 2007
FME	ANO	ANO (s problémem)	ANO (s problémem)
ESRI	ANO	ANO (s problémem)	ANO (s problémem)
FIXED	ANO	NE	ANO (s problémem)
GEOMEDIA	ANO	ANO	NE
CREATE	ANO (s problémem)	ANO (s problémem)	ANO
A_FIXED	ANO (s problémem)	NE	ANO

2.2 Komplety vrstev

Všechny testované programy umožňovali export více vrstev do jednoho GML souboru (v případě ESRI profilu byly vrstvy exportovány zvlášť do jedné složky). Cílem testu bylo zjistit, zda po importu dokáží testované programy vytvořit znovu korektně jednotlivé vrstvy. Byly použity bodové, polyliniové a polygonové vrstvy z souboru ArcČR 500.

Systém ArcGIS načítal korektně všechny testované GML profily. Pouze u profilů z programu Autodesk Civil nezobrazoval správně atributová data. Tento problém byl ale způsoben už při exportu do GML formátu, kdy nedošlo k přiřazení atributových tabulek k jednotlivým vrstvám, ale namísto toho k jejich zkombinování do jediné tabulky, která byla poté přiřazena ke všem vrstvám.

Program Geomedia nabízel možnost výběru vrstev k importu, tudíž nebylo nutné načítat všechny vrstvy z kompletu. Nespolupracoval ani s jedním z FIXED profilů. U profilů FME a ESRI docházelo k problému se souřadnicovými systémy popsáným v části 2.1. V CREATE profilu nedokázal rozdělit GML soubor do jednotlivých vrstev, ale načítal jej jako vrstvu kombinovanou, tudíž s ní nebylo možné provádět žádné prostorové analýzy či jiné operace. Zcela korektně načítal pouze nativní profil. Aplikace Autodesk Civil umožňovala načíst vybrané vrstvy z kompletu jako program Geomedia. Zcela korektně nedokázala zpracovat žádný z testovaných profilů. U nativních profilů nedošlo k rozdělení na jednotlivé vrstvy. U profilů vytvořených programem ArcGIS došlo pouze k nedostatkům s atributovou částí dat popsáných v části 2.1.

Tabulka 2. Kompatibilita kompletů vrstev s testovanými programy.

Profil GML	Testovaný program		
	ArcGIS 9.1	Geomedia 6.0	Autodesk Civil 3D 2007
FME	ANO	ANO (s problémem)	ANO (s problémem)
ESRI	ANO	ANO (s problémem)	ANO (s problémem)
FIXED_ESRI	ANO	NE	ANO (s problémem)
GEOMEDIA	ANO	ANO	NE
CREATE	ANO(s problémem)	NE	NE
A_FIXED	ANO(s problémem)	NE	NE

2.3 Datové typy

Datové typy atributů významně ovlivňují velikost atributové části jakékoliv vrstvy. Ve formátu SHP je atributová část vrstev uložena v samostatném souboru DBF. Pro účely testu byla v systémech ArcGIS a ArcView vytvořena bodová vrstva s jedním geoprvkem, jehož atributová tabulka obsahovala sedm atributů s různými datovými typy (SHORT, LONG, FLOAT, DOUBLE, TEXT, DATE a BOOLEAN). Velikost vytvořeného DBF souboru vrstvy byla 390B. Test spočíval ve vyexportování tohoto souboru do GML, jeho zpětném importu a následné konverzi na soubor typu SHP. Sledována byla změna velikosti mezi původním a novým souborem DBF a změna datového typu polí, která byla zjištěna ze souboru GML nebo příslušného souboru XSD.

Provedený test prokázal, že při konverzi dat došlo v mnoha případech ke změnám datových typů, které vedly ve všech případech ke zvětšení DBF souborů. Nejmenší změna byla zjištěna u GML profilu A_FIXED programu Autodesk Civil, kdy došlo k nárůstu velikosti DBF souboru na 925B. Největší změnu naopak vykázal profil FIXED ze systému ArcGIS. V tomto případě vzrostla velikost až na 2,58kB, tedy o více než 600% oproti původní. Změna velikosti závisí vždy na atributové části konkrétní vrstvy. Navíc v některých GML profilech vznikají po exportu nová atributová pole, která také ovlivňovala výsledné hodnoty. Velikost změny tudíž nelze obecně kvantifikovat. Lze ale konstatovat, že v důsledku konverze mezi SHP souborem a GML dochází ke zvětšení velikosti nového SHP souboru v důsledku změny datových typů.

Profil GEOMEDIA nedokázal zpracovat atribut v datovém typu DATE. Datový typ sice převedl, ale hodnota atributu byla po exportu ztracena.

Tabulka 3. Změna datových typů po exportu do formátu GML.

Datový typ	Profil GML					
	FME	ESRI	FIXED	GEOMEDIA	CREATE	A_FIXED
SHORT	short	short	xml_int 16	integer	integer	integer
LONG	double	double	xml_real 64	integer	double	double
FLOAT	double	double	xml_real 64	double	double	double
DOUBLE	double	double	xml_real 64	double	double	double
TEXT	string	string	xml_char	string	string	string
DATE	string	string	xml_char	datetime	string	string
BOOLEAN	short	short	xml_int 16	boolean	integer	integer

2.4 3D souřadnice

Ve specifikacích k formátu GML 2.x není zachování 3D souřadnic garantováno. Tento test měl za cíl tuto skutečnost ověřit. V programu ArcView byl vytvořen 3D soubor SHP, který byl vyexportován do formátu GML, zpětně načten a posléze převeden opět do SHP. Výsledek testu byl poté zkoumán pomocí funkce 3DScene v systému ArcView.

3D vlastnosti zůstaly zachovány pouze u GML souboru vytvořeném v nativním profilu programu Geomedia. Při exportu bylo ale třeba zvolit možnost „Export 3D Coordinates“, protože v základním nastavení byl test neúspěšný.

2.5 „Díra“ v polygonu

V datech se občas vyskytují nekompletní polygony, které mají ve své ploše „díru“ nebo jiný polygon. V souborech GML je tato problematika řešena značkami `innerBoundaryIs` a `outerBoundaryIs`. Na testovacích datech vytvořených v programu ArcView bylo sledováno, zda po exportu těchto dat do GML formátu a jejich opětovném načtení zůstane u polygonu „díra“ zachována. Ve všech profilech byla zpracována data korektně. Program Geomedia ale nedokázal „díru“ správně vizualizovat, tudíž jsou taková data v tomto systému dále nezpracovatelná.

2.6 Souřadnicové systémy

Geodata v sobě obsahují informace o souřadnicovém systému v němž byla vytvořena. Ve struktuře GML souboru je tato informace zakódována pomocí atributu `srsName` značky `gml:Box`. Cílem tohoto testu bylo zjistit, zda si dokáže formát GML poradit s exportem geodat, tak aby po opětovném importu, měla nadále tuto informaci v sobě zahrnutou. Test byl prováděn na datech v souřadnicových systémech obsažených v ArcČR 500 tedy S-JTSK, S-42 a WGS84.

Systém ArcGIS převáděl podle parametrů souřadnicového systému na co nejpodobnější podle OGC. Systém S-42 byl převeden na Pulkovo 1942/Gaus Krüger zone 3 s naprosto totožnými parametry. S-JTSK nedokázal exportovat, takže po importu neměla tato vrstva přiřazený žádný souřadnicový systém. WGS84 byl převeden na Lat Long WGS84.

V programu Geomedia byl souřadnicový systém definován pomocí souboru CSF (coordinate system file), který musel být přiřazen k exportované vrstvě a být pojmenovaný ve tvaru EPSG kódu. Tyto soubory byly umístěny implicitně v adresáři extenze Geomedia GML Export Command. Aby mohl být takto vytvořený GML soubor zpět naimportován bylo nutné, aby byl stejný CSF soubor umístěn i v adresáři extenze Geomedia GML Data Server. Některé CSF soubory byly již obsaženy v instalaci výše zmíněných extenzí. Systému WGS84 odpovídal soubor EPSG4326. Pro zbývající dva souřadné systémy musely být vytvořeny nové pomocí funkce GeoWorkspace Coordinate System z nabídky View.

Program Autodesk Civil nedokázal zachovat informaci o souřadnicovém systému S-JTSK a S-42 ani v jednom ze svých GML profilů. U WGS84 zůstaly korektní. Před exportem bylo ale nutné souřadnicový systém zvolit příkazem Určit globální souřadnicový systém z nabídky Map.

3 Rešerše volně šiřitelných GML prohlížeček

V kapitole věnované řešením byla zkoumána jejich kompatibilita s testovanými GML profily a s komplety vrstev a popisovány jejich nadstandardní funkce. Konkrétně se jednalo o prohlížečky GML Viewer, Gaia, iSMART Explorer, FME Universal Viewer, Degree Viewer, JUMP, Aristoteles 3D. Prohlížečka FME Universal Viewer je základem extenze Data Interoperability pro ArcGIS 9.1.

4 Závěr

Na základě provedených testů byla zjištěna řada nedostatků, kvůli kterým je velmi obtížné využívat formát GML pro přenos vektorových geodat mezi jednotlivými systémy. Samotný standard ale ve své specifikaci obsahuje všechny náležitosti, které jsou k jeho korektnímu využívání třeba. Tyto nedostatky jsou tedy způsobeny odlišnou implementací standardu GML v aplikacích a z ní vyplývajícím odlišným chováním testovaných GML profilů.

5 Reference

- [1] Aristoteles Viewer [online]. Poslední aktualizace 2007-03-21 [cit. 2007-03-28]. <131.220.71.208>
- [2] Carbon Project [online]. Poslední aktualizace 2007-04-24 [cit. 2007-04-03]. <<http://www.thecarbonportal.net>>
- [3] eSpatial [online]. Poslední aktualizace 2007-04-17 [cit. 2007-03-25]. <<http://www.espatial.com>>
- [4] Lake Ron.: *Geography mark-up language (GML)*. John Wiley & Sons, Chichester, 2004. ISBN 0-470-87154-7
- [5] OGC. Geography Markup Language (GML) Implementation Specification, version 2.0 [online]. 2002 [cit. 2007-04-01]. <http://portal.opengeospatial.org/files/?artifact_id=1034>
- [6] OGC. Geography Markup Language (GML) Implementation Specification, version 2.1.2 [online]. 2002 [cit. 2007-04-01]. <http://portal.opengeospatial.org/files/?artifact_id=11339>
- [7] Safe Software [online]. 2007 [cit. 2007-04-10]. <<http://www.safe.com>>
- [8] Snowflake Software [online]. Poslední aktualizace 2007-03-20 [cit. 2007-04-01]. <<http://www.snowflakesoftware.co.uk>>
- [9] Source Forge [online]. Poslední aktualizace 2005-02-03 [cit. 2007-04-05]. <<http://deegree.sourceforge.net>>